Semantic Interoperability for Enterprise-wide Sharing of Geospatial Information

Manoj Paul, S. K. Ghosh
School of Information Technology
Indian Institute of Technology
Kharagpur, India
{mpaul,skg}@sit.iitkgp.ernet.in

Abstract: Although large amount of Geospatial data is being made available across several organizations, inter-organization sharing of geospatial Information (GI) is a major challenge due to the complexity that arises from the proprietary nature of these spatial data sources. The growing need in several business applications and the underlying heterogeneity of spatial data present the need for interoperable and distributed GI services. Enterprise-wide sharing of GI is one of the requirements in the Geospatial domain today. As a consequence, increasing number of GI services is being available on the web, and the problem arises in the semantic heterogeneity in service description itself. Thus it is necessary for to assess the semantic nearness of the services. The quality of service (QoS) depends on resolving the proper semantic of the requested information. Besides other QoS parameters also required to be considered. The paper proposes an approach for GI services discovery and access model at the Enterprise level in which the semantics of the GI services has been taken into account and services ranked in the order of degree of similarity matching with respect to the service request and service advertised. Ontological thesauri at domain level have been used for service-independent matching of request parameters. On the other hand, application level ontology for individual application is used for service-dependent matching. An ordered list of services is prepared based on similarity score, which is further refined based on the other QoS factor. In this paper it has been shown that these two QoS issues need to be considered for optimum service performance in GI domain. An efficient design for *Geospatial Service Provider* (GSP) and *Geospatial Service Broker* (GSB) in SOA (*Service Oriented Architecture*) based GI paradigm has been proposed. The OGC and ISO/TC standards for geospatial services have been adhered.

Introduction and Motivation

Spatial information is a kind of basic information resource that is used in almost all areas of business decision support applications, such as organization planning, land management, weather forecasting, disaster/crisis management. The data required for these purpose have been collected by several organizations in huge amount but in isolation in several proprietary vendor specific formats. Over the past decade, geospatial technologies have found a paradigm shift. It has evolved from the traditional model of stand-alone systems, in which spatial data is tightly coupled within the systems in which they have been created, to an increasingly distributed model based on independently provided, specialized, interoperable geospatial services (henceforth referred as GI services). This evolution is fueled by the factors such as GIS's growing role in today's organizations, spatial data's increasing availability and inherent conduciveness to reuse, the maturity of Web and

distributed computing technologies, and GIS's key role in the promising location-based services market.

The services-based GIS model is rapidly materializing, owing in part to the advancements in general web service technologies, and in part to focused efforts by the Open Geospatial Consortium (OGC; www.opengis.org) to sponsor consensus-based development of interoperable GI service interfaces. The goal is to dynamically assemble spatial applications in the form of GI services for use in a variety of client applications. But then a major challenge evolves within this GI service paradigm itself. Most often users face difficulty in finding the appropriate services from the huge sea of available services, which can provide the necessary information according their requirement, the services the context of which is in accordance users need. This situation is overwhelmed due to diverse semantics followed for the data features and service annotations. The present problems existing in applications of spatial information are enormous distributed data, complicated processing, heterogeneous system structures etc.

Enterprise-GIS

Recent developments of geomarketing, territory management, and scientific discovery increase the need for integration of available highly heterogeneous spatial data resources spreading across organizations. As the web accessibility of GISs is increasing tremendously, there is the huge requirement from users' for online access and analyze data from those systems. With the proliferation of geographic and spatial data on the Internet, a much larger audience is entitled to access and share data currently available in various GIS. Interoperability is the key issue for a public or private organization that wishes to consolidate its disparate spatial data into an integrated GIS that would serve as a basis for any decision making process [PG05].

Inter-organization sharing of geospatial data will directly benefit business organizations by providing better decision support. The reuse of the existing geospatial data in the form of GI services will further enhance the interoperability in business information among corporate and government sectors. In this context it is worth to mention that Enterprise Geographic Information System (E-GIS) [PGa06] is an organization-wide approach to GIS implementation, operation, and management. An E-GIS provides access to shared geo-spatial information and analysis resources for a large number of concurrent users located in different parts of an institution. EGIS can also be defined as an effort to design integrated geospatial data management system to serve a complex institution. It can also be seen as a global virtual geospatial data repository, which provides a single point access for diverse spatial repositories spreading over large geographical location. In the past, numerous technological roadblocks hampered the successful implementation of E-GIS. Refs. [PGa06, PGb06, PGc06] gives more detail on E-GIS. One of the objectives of the paper is ensure effective search of GI services for the efficient development of the E-GIS system by considering the semantics and the QoS of the available GI services.

#

GIS Interoperability#

Interoperability#is the major issue that needs to be considered for the business# organizations getting benefit from the spatial information in heterogeneous domain# like GISs. Let us first try to find the exact meaning of the term INTEROPERABILITY# and then discuss the methodologies for ensuring interoperability among GI services.# There are numerous definitions for interoperability in the literature. Instead of# trying to add one more, we quote some of them [PT06]:#

- IEEE defines it as the ability of two or more systems or components to# exchange information and to use the information that has been exchanged.#
- According to#IDABC, European Interoperability Framework, interoperability# means the ability of information and communication technology (ICT)# systems and of the business processes they support to exchange data and to# enable sharing of information and knowledge.#
- DARPA defines interoperability as##
    - the ability of systems, units, or forces to provide services to and# accept services from other systems, units or forces and to use the# services so exchanged to enable them to operate effectively together,#
    - the condition achieved among communications-electronics systems or# items of communications electronics equipment when information or# services can be exchanged directly and satisfactorily between them# and/or their users.#
- TOGAF (The Open Group Architecture Framework) defined interoperability# as:##
    - the ability of two or more systems or components to exchange and# use shared information, and##
    - the ability of systems to provide and receive services from other# systems and to use the services so interchanged to enable them to# operate effectively together.#

#

In the context of GIS, it can be said that interoperability is the methodology for# defining standard interfaces for accessing Geospatial Information, a mechanism for# disparate access of information from the heterogeneous GISs, irrespective of the# underlying complexity in the data provider's proprietary data formats and access# mechanisms. Achieving interoperability by handling syntactic and semantic# heterogeneities among large scale of geospatial data repositories has been widely# discussed in the literature. Refs. [GM99, KS98] give more detail on the# interoperability issue for GISs.##

#

In service based GI methodology, the service providers need to advertise the# service descriptions in some standard means (e.g. WSDL). Interoperability approach# seeks that a central authority must take the responsibility of publishing the GI# services by maintaining single/distributed catalog/metadata for the service# providers for a uniform service description method. The central registry, which is# also acting as a service, should understand the semantics of the user's query and# use the available metadata information to provide semantically close service# descriptions. Such a system can then support high-level, context sensitive queries# over multiple data sources, irrespective of the underlying heterogeneities. This# issue, termed as semantic interoperability, is the major focus of this paper and#

#

whenever the term interoperability is used in the rest of the paper, it will mean semantic interoperability.

The proposed system follows Service Oriented Architecture (SOA) (next section details on SOA and GI services) model for geo-spatial services that utilizes the semantic based QoS for interpreting users' semantics. Our main concern is to provide interoperability and greater extensibility of GI services by adhering to the OGC web service framework for designing the proposed model. The framework uses WMS, WFS, and WCS for building various kinds of geo-spatial servers.

Web Service and SOA

Recently, web services have evolved to be an efficient way for inter-organization sharing of heterogeneous data. XML, SOAP [SOAP01], UDDI [UDDI02], WSDL [WSDL02], SOA [ET04] has become major tools for realization of service based system. Adoption of Service Oriented Architecture (SOA) is expected to allow enterprises to contract-out their non-critical functions to several parties. In the new world economy, business processes typically transcend departmental as well as organizational boundaries. Web services are expected to provide the ideal platform to automate these processes as they allow integration of disparate platforms and systems. These advantages of web services can be taken for the GI services as well. A great deal of attention in the web service technology has been directed towards efficient service integration among heterogeneous distributed servers [DFK02, BB99, BE01]. The feature of web services, which allows sharing the dynamic content across organization, is especially important the enterprise business community, because it provides the opportunity for sharing information faster and efficiently. However, it is really challenging to provide geospatial information in the form of services on web environment on account of diversity and complexity [LT92] of geo-spatial data themselves. Detection of location of services on the basis of its capabilities is essential in this perspective. Although several organizations have started offering the spatial data as services for the users' convenience, the efficient design of Geospatial Service Provider (GSP) and the Geospatial Service Broker (GSB) is still undergoing heavy development and standardization [BN03, KRH01] in service based integration method for GI services.

In order to serve geospatial information to the users, we have adopted Service Oriented Architecture (SOA) principles [ET04] for designing the framework for the E-GIS system. SOA supports a global directory (distributed or centralized) in which service providers can advertise their services in a form that enables potential clients to find and invoke them over the Internet. For service interface definition purpose, we follow the Open Geospatial Consortium (OGC) specified standard specifications. OGC has introduced specifications for geospatial web services, such as Web Feature Service (WFS) [OWF02], Web Map Service (WMS) [OWF02], Web Coverage Service [OWC02], Web Registry Service [OWR03] and Catalog Service [OCS01]. Nevertheless the SOA and OGC approach for Enterprise-base GI service integration has several limitations. The shortcomings are: WFS and WMS only define the interface requirement for defining services i.e. it only standardizes the interface for service invocation, little information about the formal way of advertising the services and subsequent discovery methodology. Thus standard

means for service description will enhance the discovery of GI services. As# mentioned earlier, with the increasing number of available GI services, an# immerging need has also appeared for methodologies to locate desired services# that provide information in close proximity of the users' need. It has been observed# that in most of cases the usability of information created in one context is often of# limited use in another context, as the semantics of service description is not known# to the other party. Key-word based search appears to be incomprehensible for# discovering the right service [PKPS02, MSZ01]. This seems to be the major# disadvantage for Catalog# Service (CS) [OCS01] defined by OGC for service# registering and discovering. The objective of the proposed service discovery# mechanism is to provide the requesters with a collection of matched service offers# ordered on the basis of similarity score.#

Semantic Service Discovery#

The similarity matching approach for services based on ontology have been studied# extensively in recent times [MSA05] for matching a service description with service# request. Identifying similarity between services is challenging [MSA05] because the# terminology used to describe the services may not be identical to the one used in a# query. The problem is similar to that of document retrieval from a library of# documents. As in document retrieval, searching for services should be on the# notion of similarity rather than identical matches as we are unlikely to find exactly# matching descriptions. The solution to this problem requires a language for# semantic expression of service, and the specification of a matching algorithm# between service requests and service advertisements that recognizes when a# request matches a service advertisement.#
#
In this paper, an approach for the semantic classification and discovery of GI# services has been proposed. In particular, in the discovery process services are# firstly looked for on the basis of their functional conformance to the user request# and then filtered on the basis of their quality features (popularly known as QoS for# services). Identifying appropriate GI services for performing a required function# from a large collection of available GI services has been addressed in this paper. A# similarity-matching-based-filtering of services based on the matchmaking of# attributes and operations and rank the services in order of greater degree of# matching. The usability of domain and application ontologies has been discussed# for finding the similarity among the services.#
#
The contribution in of the proposed work can be pointed out as follows:#

- Defining a framework for service based Enterprise-GIS and its requirements,# especially focusing on the requirement of geospatial information retrieval.#
- Addressing the structural heterogeneity among the geospatial data# repositories (very brief description citing previous works).#
- Identifying the requirements for the services mainly#concentrating on the# semantic aspects and QoS requirement.#
- A similarity matching approach for discovering services for increasing# trustworthiness of information and subsequent precision recall study.#

- Incorporating other popular QoS measurements and further enhancing the partially ordered list of services obtained from the semantic matchmaking approach for satisfying user specified non-functional requirement.

The rest of the paper is structured as follows. Section 2 gives a discussion on the related works for semantic based service discovery. Section 3 proposes the service based E-GIS system. In section 4 the requirements for the GI to be made available as services on the web has been identified. Section 5 gives a detail discussion of the methodologies followed for semantic based ranking based on similarity matching of the GI service description. Finally, we conclude in section 6.

Related Work

Several approaches have been proposed and applied towards the goal of sharing and integration of spatial information repositories [DFK02, JMS88, BB99]. Migrating to OGC specified web services for spatial services and semantic based service advertising and discovery have become important research issue in recent times. In this section we briefly overview the technologies closely related to our work considering system architecture and service-based collaboration. It now appears that the problem of service description mapping for efficient service discovery is similar to the well-studied problems of component retrieval and information retrieval. The major tool for services interface description is WSDL. WSDL specification is the description of a Software Component including a description of its interface and a description of where the actual implementation exists and how it can be used. This usually includes a set of natural-language description of the service itself and comments on its elements. Some related work on the semantic based services similarity mapping is discussed in this section.

GIS Interoperability Approach

The need of sharing/integrating GISs has been addressed from quite some time [TG01]. Several ways of integrating the spatial data repositories have been proposed in the literature. The database community has done extensive research for the formalization of an integration architecture [WG92]. Several solutions for integrating heterogeneous repositories have been proposed by the research community, ranging from schema integration approach of federated to mediation approaches using ontologies, which embrace the idea of metadata standards as the key to information sharing and analysis. These include the *Federal Geographic Data Committee* (FGDC) and the *National Spatial Data Infrastructure* (NSDI), *Geospatial One-Stop*, and *U.S. Geological Survey's The National Map* as well as standards from the *International Standards Organization* (ISO) for geospatial metadata [RBPA03]. The NSDI attempts to bring together geographical information sources from all levels of government and other organizations into a single point of entry for easier access to data. Some of these recent approaches have been analyzed here and their relative advantage and disadvantages have been brought out. The proposed approaches for the realization of an enterprise level integrated GIS can be classified into the following categories:

#

#

a. Data Warehousing;#
b. Mediator-based systems;#
c. Cataloging Geographic sources;#

#

Most current approach to geographic data integration is building a warehouse of# spatial data from the available sources after being converted into a common# standard format. Ref. [RP97] proposes a spatial data warehouse based technique# for integrating spatial data repositories and employing middleware technology for# data exchange. In the datawarehouse-based technique data from multiple# repositories are cleaned up and maintained in a warehouse. The warehouse can be# composed of a central data repository or distributed data repositories depending on# the design of the warehouse. Ref. [CDSS98] proposes a mediation-based approach# to facilitate data integration.#Most of these approaches rely on the wrappers# [RP97, CDSS98] for accessing the data sources and then translating the data in# some standard format before sending it to the requester. The main disadvantage of# the warehouse-based approach is that a local administration is required to maintain# data. The currency of data may not be preserved well, unless the datawarehouse is# updated quite frequently. The frequency of updation is limited by the huge size of# spatial data. It also makes the data providers tightly#coupled to each other, in the# sense that configuring a new data source to the system requires a substantial# amount of overhead. Mediator based approach, on the other hand, handles the# queries directly through wrappers and integrates the data locally, which can be# costly.#

#

The major initiative towards geospatial interoperability has come from the Open# geospatial Consortium (OGC). It is an international organization standardizing the# geoprocessing activities and geospatial data formats for ensuring interoperability.# OGC specified specifications provide methodologies to integrate geospatial data# and geo-processing resources into mainstream computing. This leads an effort to# provide a) standardized protocols for accessing geospatial information and services# and#b) standardized service metadata for enhancing the service discovery in spatial# domain. OGC introduces a) Catalog Specifications (v.2) [OCS01] which defines# discovery and retrieval of metadata that describes geospatial data and geo-processing services and#b) Web Registry Service (WRS) [OWR03] Specifications as# implementation specification of the OGC Catalog Specifications. WRS Specifications# define a standard mechanism to discover/publish operation and attribute# information of geospatial services and presents a domain-specific registry# capability for geospatial information. As mentioned earlier there is no methodology# specified which addresses the issue of semantic heterogeneity for discovery and# subsequent retrieval of spatial Information.#

#

Although OGC provides interface specifications for GI service standardization, they# do not provide any standard framework for geospatial service-based computing.# World Wide Web Consortium establishes Web Services Architecture to define# components and their relations of service framework. WSCA (Web Services# Conceptual Architecture) [KRH01] provides foundation of building and deploying GI# service applications. OGC together with committee of ISO TC211 proposed a#

#

#
framework of OGC Web Services (OWS), which is an online spatial#information# service framework enabling seamless integration of spatial information services.#

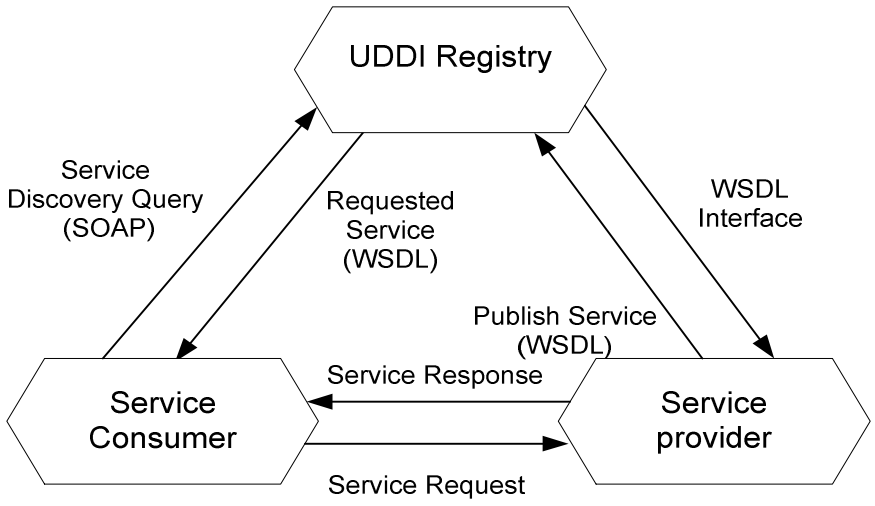
Semantic Service Matching#

The major objective of the paper is to consider the semantic heterogeneity in# service descriptions for spatial services. The semantic aspect of service# description, which poses a challenge for discovering services, has been discussed in# the literature [PKPS02, MSA05]. Ref. [PKPS02] proposes the use of ontologies for# matching service descriptions based on the meanings of the query parameters# rather than exact matching. It also proposes a mechanism for sorting the matching# services on the basis of degree of matching. WordNet [MGA83] has been used for# finding the synonyms and hyponyms of the requested parameters. WordNet is a# lexical database,#inspired by current psycholinguistic theories of human lexical# memory. English nouns, verbs, adjectives and adverbs are organized into synonym# sets, each representing one underlying lexical concept. The use of WordNet in# information retrieval [RS05] has been incorporated for service discovery as well. In# performing the search to identify existing services that can be used to implement a# required service, we need to address two fundamental issues [MSA05]. We need a# way to match the requirements and capabilities of services i.e, identify similarity# between services. Our objective is to apply the advancements in web service# technology in GI domain and more importantly resolving the various semantics of# user request and available services and preparing an ordered list on the basis of# this semantic QoS for increasing users confidence in retrieving spatial information.#
#
A close approach is#*Signature#Matching*#[ZAWJ95], which is an efficient way for# component retrieval form software libraries. The method is efficient in the sense# that function signatures are automatically generated from the function code.# Furthermore, signature matching efficiently prunes down the functions and/or# modules that do not match the query, so that more expensive and precise# techniques can#be used on the smaller set of remaining candidate components.# Although this methodology can be incorporated in service discovery also, signature# matching considers only function types and ignores their behaviors; and two# functions with the same signature can have completely opposite behaviors.##
#
The industrial world has developed a number of XML based standards to formalize# the specification of Web services, their flow composition and execution, such as# BPEL4WS [CGK02]. There is also a DARPA project DAML-S#[ABH02] that has taken# the procedural approach to service composition. As part of the DARPA Agent# Markup Language (DAML) program, an ontology description language has been# developed, called DAML-S, which is used for semantically describing web services# for#better machine understandability. Although this approach has advanced in the# general information technology domain, their use is still to gear up in the GI# service domain.#
#
#
#

#

Service#Oriented Paradigm##

Web services are "self-contained, self-describing, modular applications that can be# published, located, and invoked across the Web. …Once a Web service is deployed,# other applications (and other Web services) can discover and invoke the deployed# service" [TD05]. They provide access to sets of operations through one or more# standardized interfaces. The development of the service-oriented web-based# applications architecture style, as implied by the web-services stack of standards,# defines how reusable components should be specified (through the Web-Service# Description Language#-#WSDL [WSDL02]), how they should be advertised so that# they can be discovered and reused (through the Universal Description, Discovery,# and Integration API#-#UDDI [UDDI02]), and how they should be invoked at run time# (through the Simple Object Access Protocol API#-#SOAP [SOAP01]). Services are# independently published by different providers and, according to the Service# Oriented Architecture (SOA), users can request and access a wide variety of# services by choosing among the published ones.#Services satisfying user preferences# and constraints are selected, possibly by means of brokers, looking for them in a# Service Registry. To facilitate service discovery two main research issues have to# be addressed: (i) services organization and classification according to functional# properties; (ii) service discovery on the basis of QoS demand. A critical step in the# process of reusing existing WSDL-specified components for building web-based# applications is the discovery of potentially relevant components. Figure 1 shows# the general publish-find-bind methodology for web service discovery and retrieval.#
#



#
##

Figure 1: General publish-find-bind methodology for web service discovery and retrieval#

OGC Web Services#

As we turn our focus for the usability of services in geospatial domain, we see that# GI services are mostly researched and standardized by OGC. OGC has specified# many implementation specifications for GI service of various kinds of geospatial# data. In the beginning, OGC announced WMS [OWM02], WFS [OWF02], WCS# [OWC02], and Catalog services specification [OCS01] for service based access of#

#
spatial data over the web. GIS services can be grouped into the following# categories:#
#
Data services#are#typically#tightly coupled with specific data sets and offer access# to customized portions of that data. Examples of data services include the Web# Mapping Service (WMS), which produces maps as two-dimensional visual portrayals# of geospatial data; the Web Coverage Service (WCS), which provides access to# unrendered geospatial information as needed for client-side rendering; and the# Web Feature Service (WFS), which lets a client retrieve geospatial data encoded in# Geography Markup Language (GML).#
#
Registry, or catalog services#allows#users and#applications to classify, register,# describe, search, maintain, and access information about#GI#services. They store# information on data types, online data instances, service types, and online service# instances.#These services can be thought of as an extension of UDDI for general# service-based architecture.#
#
We mainly adopt the WFS and WMS specification for ensuring interoperability# among the GI services.###

Web Feature Service#

Web Feature Service is one of the GIS web service interoperable specifications# defined by OGC#[OWF02]. It is the most powerful data service of OGC Web# Services. Web Feature Service allows a client to retrieve geospatial data from# multiple geospatial data servers.##
WFS defines three primary operations:##

- GetCapabilities operation describes capabilities of the web feature service# using XML, it indicates which feature types it can service and what# operations are supported on each feature type#-#can be used for the# discovery of a service.#
- DescribeFeatureType operation describes the structure of any feature type it# can serve#-#helps to get knowledge of the feature data and subsequent# formation of spatial query.##
- GetFeature operation services a request to retrieve feature instances. In# addition, the client should be able to specify which feature properties to# fetch and should be able to constrain the query spatially and non-spatially.#

#*Web Map Service*#

The OGC WMS#[OWM02]#is capable of creating and displaying maps that come# simultaneously from multiple sources, in standard image formats such as#Scalable# Vector Graphics (SVG), Portable Network Graphics (PNG), Graphics Interchange# Format (GIF) or Joint Photographic Expert Group (JPEG).#
It proposes the following main operations:##

- GetCapabilities allows a client to instruct a server to provide its mapping# content and processing capabilities and return service-level metadata.#

#

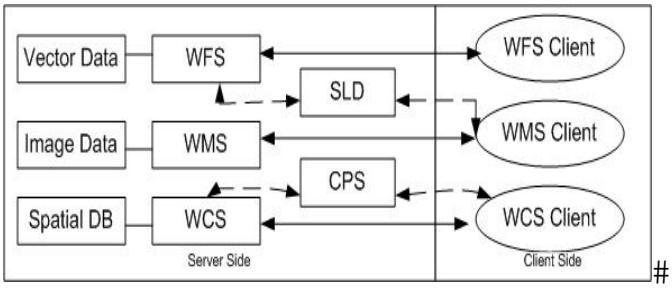
#

- GetMap enables a client to instruct multiple servers to independently craft# "map layers" that have identical spatial reference system, size, scale, and# pixel geometry. The client#can then display these overlays in a specified# order and transparency such that the information from several sources is# rendered for immediate human understanding and use#
- GetFeatureInfo enables a user to click on a pixel to inquire about the# schema and metadata values of the feature(s) represented there.#

#

OGC began to cooperate with W3C recently, and OGC is announcing new# specifications such as WRS [OWR03], GML [OGM02], and OGC Reference Model# (ORM) [ORM03] for standard service architecture. Because interoperability is# critical to the GIS Web services architecture, the GIS community, with OGC# sponsorship, is focusing on defining interoperable interfaces to the basic services.# There are additional specifications for Styled Layer Descriptor (SLD) [OSL02] and# Coverage Portrayal Service (CPS), which define methods to change client-side# mapping into server-side mapping in case of WFS and WCS servers. Fig. 2 shows# relations between OGC specifications that serve various kinds of geospatial data.#

#



#

#

Figure 2: Relations among various components in OGC specified implementation specification#

#

Due to the lack of proper architecture and#suitable#methodology for spatial# information retrieval from the large number of GI services, spatial domain is yet to# get the advantages#of service based computing. On the other hand even standard# UDDI seems to be inefficient (discussed in subsequent sections) for semantically# heterogeneous GI domain. Thus semantic web can be a promising solution. It# requires that data be not only machine#readable, but also machine understandable.#

Requirement for GI Service Computing#

In this section the applicability of present service oriented computing in GI domain# has been studied and specific requirements specifically need to be considered for# GI domain, have been identified.##

#

The major components of service architecture are SOAP, WSDL and UDDI. A critical# step in the process of efficient utilization of WSDL-specified components for# building web-based applications is the discovery of potentially relevant# components. UDDI is the means for publishing services in a uniformed manner.# General UDDI servers are essentially catalogs of published WSDL specifications of#

#

reusable components. This catalog-based service discovery method is clearly insufficient as does not consider the semantic aspects of the service description.

In this paper it is proposed that the GI services be discovered and ranked in order of degree of matching for retrieving services to the requester, which are close in context to the service request. This semantic QoS for services has been exploited in this paper. The methodology follows an ontological matching of the request parameters to that of the available services. Reference [MS04] proposed a similar ontology-based framework for dynamic Web service selection.

Semantic Web

There is huge hype at present on the Semantic Web aspect. We briefly discuss about this comparatively new paradigm. With the growth of the World Wide Web has come the insight that currently available methods for finding and using information on the web are often insufficient. Today's retrieval methods are typically limited to keyword searches or sub-string matches, offering no support for deeper structures that might lay hidden in the data or that people typically use to reason; therefore, users may often miss critical information when searching the web. At the same time, the structure of the posted data is flat, which increases the difficulty of interpreting the data consistently. There would exist a much higher potential for exploiting the web if tools were available that better match human reasoning. In this vein, the research community has begun an effort to investigate foundations for the next stage of the Web, called Semantic Web [BHL01].

The Semantic Web requires that data be not only machine readable (just like the Web nowadays does), it also wants the data to be machine understandable. Moreover, the developers of end user applications will not need to worry about how to interpret the information found on the web, as ontologies will be used to provide vocabulary with explicitly defined and machine understandable meaning [MDL98]. A great deal of attention has been focused on semantic web services, the aim of which is to describe and implement web services so as to make them more usable to the user. Ontologies can be used to describe services so that services can be advertised and discovered according to a semantic specification of functionality [TBP02].

With the proliferation of semantic web in the Internet, there are also proposals on Geospatial Semantic Web [SMSJ02]. The enormous variety of encoding of geospatial services makes them particularly challenging while processing requests for GI services. In the future, the Geospatial Semantic Web will allow the returning of both spatial and non-spatial resources to simple queries, using a browser. For example, a query lakes in Maine should return all relational resources containing information about lakes in Maine (pictures, text, ...) in different formats (XML, HTML, JPG, PDF, References, ...) [EJ02].

As a first step in realizing the Semantic Web, new standards for defining and using ontologies are already being developed. RDF, which is being developed by the W3C RDF Core working group, is a web markup language that provides basic ontological

primitives [BG03]; DAML+OIL is an ontology language that extends RDF with a much richer set of primitives (e.g., Boolean operators and cardinality constraints), and which is now the basis for the W3C Web Ontology Language working group's development of the OWL ontology language standard [HA02]. Moreover, if applications are to exchange semantic information, they will need to use common ontologies. One such ontology (written in DAML+OIL), which has been designed for the purpose of describing web services, is the DAML-S ontology [BB99].

Semantics for GI Services

In [SMSJ02], semantic heterogeneity is defined as the consequence of different conceptualizations and database representations of a real world fact. The semantic of service requests needs to be resolved properly at the *Geospatial Service Broker* (GSB) so that a user can get what they are looking for. Keyword-based search can have low recall if different terminology is used and/or low precision if terms are homonymous or because of their limited possibilities to express complex queries [BK02]. Existing approaches for service selection [PKPS02, MS04] make no provision for user-specified ranking criteria as part of the service request. Semantic Web Services utilize regular Web service technologies such as SOAP and WSDL (for service description). At the higher level, semantic and more expressive descriptions are used to describe the services. In this paper, Mechanisms for augmenting WSDL to provide semantic descriptions and enhancing UDDI to provide semantic discovery ahs been proposed. Fig 3 illustrates the modified SOA architecture (see fig 1) adapted to suit the needs of *Semantic Web Services* (SWS), which includes Annotated WSDL files, an Enhanced-UDDI registry and the corresponding API's in the Service Registry and Provider.

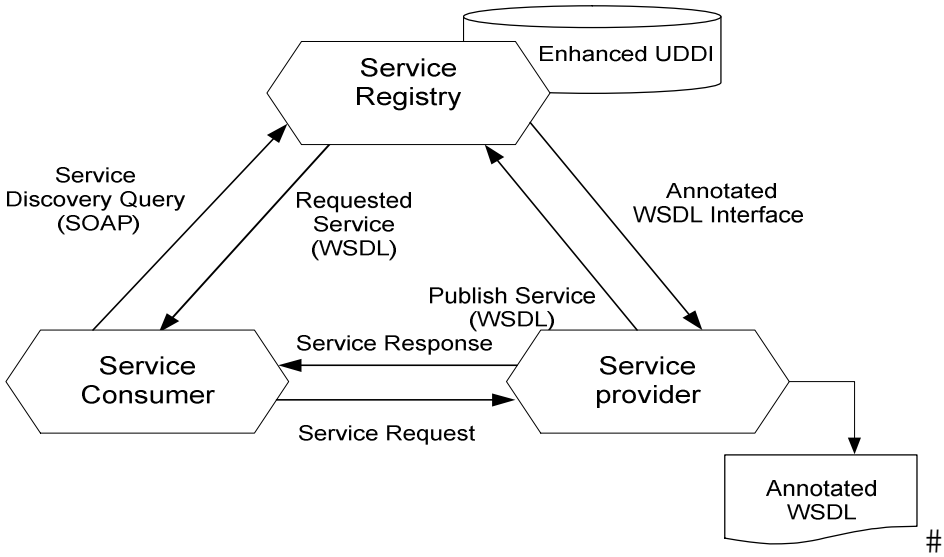


Figure 3: Enhanced SOA architecture to incorporate the semantic aspects for GI Services

The requirement for SWS arises in the distributed business applications like E-GIS, where the provider and consumer has no prior agreement on the service offers. So, it is unlikely that a request will match exactly to that of the service request. What we want is a matching methodology based on the meanings of users' service

#

request parameters.# As# an illustration let us consider two services# **S1::RoadNEtworkService**#and#**S2::SurfaceTransport-Service**. Description of these# two services is shown in figures#4(a) and#4(b)#using RDF service description# language.#Both the services provide information#almost#about the same feature,# the road network#of a place. But as can be seen, there are#taxonomical#differences# in the interfaces of the services. However, if the objective is to find a service that# gives information about the availability of a given part, semantically, both services# should be matched. This type of match cannot be determined using a Key-word# based matching mechanism. Thus we need an approach that can understand the# semantics for the services, rather than blind textual matching. Given a query WSDL# file and a set of WSDL files in a repository, it is reasonable to assume that the best# matching services are those that have a large number of semantically related# attributes.#

#

```
<rdf:RDF xml:base=  "http://www.animals.fake/RoadNetwork#">

<rdf:Description rdf:ID="RoadA">
  <rdf:type  rdf:resource="http://www.w3.org/2000/01/
    rdf-schema#Class"/>
</rdf:Description>

<rdf:Description rdf:ID="RoadB">
  <rdf:type
   rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
  <rdfs:subClassOf rdf:resource="#RoadNetwork"/>
</rdf:Description>

</rdf:RDF>
```

#

Figure#4#(a): RDF description of RoadNetwork service#

#

```
<rdf:RDF xml:base=  "http://www.animals.fake/SurfaceTransport#">

<rdfs:Class rdf:ID=" SurfaceTransport A" />

<rdfs:Class rdf:ID=" SurfaceTransport B">
  <rdfs:subClassOf rdf:resource="#SurfaceTransport"/>
</rdfs:Class>

</rdf:RDF>
```

#

Figure#4#(b): RDF#description of SurfaceTranport service#

#

The semantic heterogeneity arises because of the obvious reason#that the service# interfaces were created for different organizations, over different periods of time# and for different initial purposes.#It is unlikely#that they use a common set of terms# to name services and parameters. Thus there is the need for resolving the semantic# issue at GSB on basis of similarity-based match making.#

#

Service Definition for GI Service#

As already pointed out that services need to be advertised in the common catalog# for the convenience of users.#Web Service framework for#GI#services are supposed# to provide not only geospatial data, but also geospatial functions (e.g.#shortest# distance between two places).#Clients can access geo-spatial data directly or# indirectly using geo-spatial functions. However, the specifications of OGC,# especially WRS, did not provide any formal representation for#service functions. ##

We propose a formal advertisement for service operations using Description Logic# (DL) [BN03], which defines a function in terms of its desired action along with the# pre and post conditions it has to satisfy. A DL reasoner can check whether two# concepts subsume each other. In DL, an interpretation I = (Δ^I#, $¢^I$#) consists of a set# Δ^I#, called the domain of I, and a valuation $¢^I$#which maps every concept to a subset# of#Δ^I#and every role to a subset of#Δ^I#X#Δ^I#such that, for all concepts, roles, and# nonnegative#integers, the properties are satisfied. The use of DAML-S [ABH02] has# been studied for semantically uniform service description in correspondence with# WSDL.#

The problem that we are mainly concerned with is the service procurement# problem. It can be defined as follows [MRDB03]: given a multidimensional space Σ# whose dimensions#are associated to the domains of considered quality#parameters,# some approaches represent the QoS demand as a subspace of Σ (a set of constraints# on the values of#each quality parameter) and the QoS offer#another subspace#in Σ# (expected# value for each quality parameter) so that the# problem becomes# equivalent to check whether the offer#(subspace)#belongs to the demand (the# subspace).#

Semantic Based SOA Framework for GI Services#

This section gives a detail description of the service-based E-GIS system framework# and the enhanced methodology for spatial service computing. The adapted# methodologies for#*Geospatial Service Provider*#(GSP) and#*Geospatial Service Broker*# (GSB) have been discussed#in detail. Figure 5 shows overall architecture of the# suggested framework for semantic-based service discovery in GI Service# environment. The GSB has been designed for the semantic based brokering of# service request and subsequent analysis of the QoS of the discovered services.# Given only a textual description of the desired service, a semantic information-retrieval method can be used to identify and subsequently ranked in the order of# most similar service-descriptions. The set of services that can partially fulfill users# need are also a part of the candidate service sets in this proposed methodology.#

In a SOA, individual users or communities of users are expected to query for# services of interest to them using descriptions that are expressed using terms in# their own ontologies. But with proliferation of independently developed and# deployed services, the semantic correspondences between the user ontology on# which the user queries are based and the domain ontologies on which the service#

descriptions are based, are likely to vary. Consequently, users ought to be able to specify inter-ontology correspondences to facilitate matchmaking between the service requests and service advertisements. Subsequent section gives detail about this method for the ontology-based service discovery.

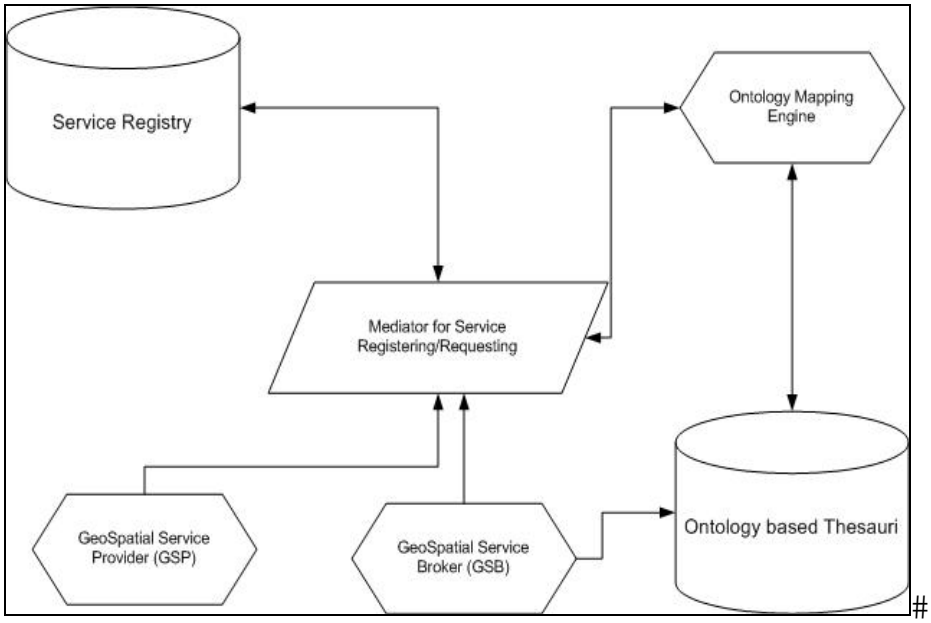


Figure 5: The overall model for the Service Oriented computing in the Geospatial environment

Geospatial Service Provider

In the suggested framework, we provide different GSP components for different kinds of data sources, namely shape file, GML file, Oracle Spatial DBMS, PostGIS etc. The uniformity in heterogeneous GIS data sources is accomplished at the data GSP end. This is accomplished by retrieving data from multiple geospatial data sources based on different platforms, performing on-line overlapping, analysis, visualization of map data, and perform user intercommunication. The *Data Transformation Service* (DTS) is interfaced as wrapper that can communicate service request at multiple data repositories irrespective of their underlying heterogeneities (figure 6). It accomplishes the function of data retrieval from all sorts of GIS platforms and data format transformation into uniform GML (Geography Markup Language). DTS mainly provides the following GML data transformation abilities of GIS data: ArcInfo's SHP files, spatial data stored in Oracle spatial and MapInfo's Tab files. In the proposed framework, the process of analysis of requests, division of requests, delivering small queries such as *GetCapabilities*, *DescribeFeatureType*, and *GetFeature* are actually designed by using WFS interfaces. Refer to [PGa06, PGb06, PGc06] for more detail on the syntactic interoperability issue for geospatial integration.

Map Data storage Format	Transformers
Oracle Spatial Database Storage	ORACLEtoGML
ArcInfo File Storage	SHPtoGML
MapInfo File Storage	MAPINFOtoGML

Table 1:Transformers of various geospatial data formats

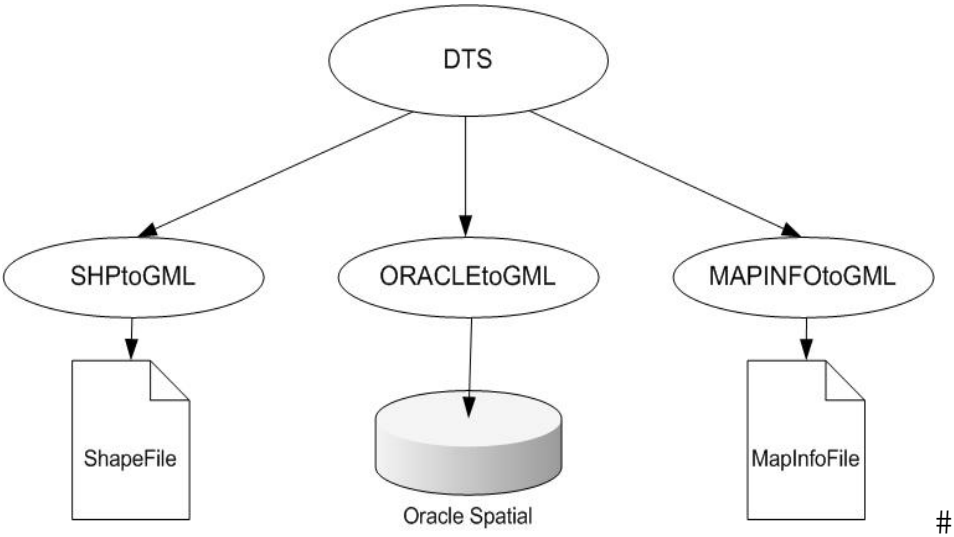


Figure 6: Heterogeneous data integration platform

Service Interface Definition

Although DTS can provide uniformity in data heterogeneity, it does not have the intelligence to process the request based on the meaning/context of the request. The semantic aspect of data request is analyzed at the GSB. To facilitate the semantic analysis at GSB, the service description at GSP has been standardized by bringing semantics in service description. For this purpose there is the need for a description language for expressing the capabilities of services that provide semantic view of the GI service descriptions. The traditional XML based approach (WSDL) for expressing service interface appears to be insufficient to capture the semantics of the services. A limitation of this is the lack of explicit semantics, in the sense that two identical XML documents may mean different things with respect to the context of their use. This is an important consideration because the requester does not know what are the services provided at any given time, otherwise, it could have contacted the providers directly without the need of discovering them. It is unrealistic to expect advertisements and requests to be identical or their exists services that exactly fulfills users request.

We choose DAML-S [ABH02] for service advertising due to its capability of semantic web service description. This language has a well-defined semantics and enables the markup and manipulation of complex taxonomic and logical relations between entities on the web. A fundamental component of the Semantic Web will be the markup of Web services to make them computer-interpretable, use-apparent. DAML-S has been used to make GI services computer-interpretable and enabling automated service discovery, invocation, composition and monitoring. It defines the notions of a *Service Profile* (what the service does), a *Service Model* (how the service works) and a *Service Grounding* (how to use the service). GI services are represented by a semantic signature, which contains *Description Logics* (DL) [FBW03] concepts to represent inputs and outputs, and a specification of pre- and post-conditions.

For enhancing the semantic discovery of services, we adhere to the standard service description methodology where a service is defined with the following components

- A signature of the service $S_{signature}$ consisting of the operation names S_{ops}, input/output parameters S_{inp} and S_{out} respectively. Thus we can say

$$S_{signature} = \left\{ S_{ops} \cup S_{inp} \cup S_{out} \right\}$$

Signature of service gives the static nature of the service.

- Specification of the service - $S_{specification}$ gives the dynamic aspects of a service. Specifications more precisely characterize the semantics of a component, rather than just its signature.

Given two services S and S′ such that S = ⟨$S_{signature}$, $S_{specification}$⟩ and S′ = ⟨$S'_{signature}$, $S'_{specification}$⟩, the matching between the two service descriptions ***Match*(S, S′)** can be defined as

$$Match(S, S') = Match(S_{signature}, S'_{signature}) \wedge Match(S_{specification}, S'_{specification})$$

As with function signature representation [MW95] for *Signature Matching*, we can represent service in the same manner by annotating all its inputs, outputs and operation. As an example we can refer to the example signature of module ***SET*** in [MW95], which defines a module with the signature of its composition functions as follows

```
signature SET =
  sig
     type α T
     val create: unit -> α T
     val insert: (α T, α) -> α T
     val delete: (α T, α) -> α T
     val insert: (α, α T) -> bool
     val union: (α T, α T) -> α T
     val intersection: (α T, α T) -> α T
     val difference: (α T, α T) -> α T
  end
```

Application of DAML-S

The objective of a service profile is to describe the functionalities that a web service wants to provide to the community. A service profile generally contains the following things (figure 1)

- Actors - Information about the provider of the service.
- Functional Attributes - Quality measurements of the services.

#

- Functional Description#-#Description of the capabilities of the services in# terms of its inputs, outputs, pre and post conditions. An input is what is# required for a service in order to produce some output. On the other hand,# output is the result that is expected as a part of service execuition. The# preconditions represent conditions#that has to be true in order to the# sucessful execution of the service.#

#

The terminology for description of operation is chosen from the domain ontology,# which consists of a thesaurus of shared vocabulary for the domain. As mentioned# the operations of a service are advertised in GSB by its inputs, outputs, pre-#and# post conditions. The ability for DAML-S to describe the semantics of a web service# is in stark contrast with emerging XML based standards related with web service. It# supports the needs for semantic representation of GI services through its tight# connection with DAML+OIL [DAML01].#

#

As an example let us consider that there is a web service W1, which provides roads# information for state A and uses domain ontologies to specify its capabilities and# services it offers. W1 may require the name of the state, category of road# (National Highways or Others) and formatting information as inputs, and road# names, road geometry and nearby cities as output. In addition, valid state name,# valid road category etc could be the prerequisites for invoking the service# (preconditions), whereas, displaying the road geometry along with its names and# providing other means of connectivity (bridges etc.) may effect after a successful# invocation of the service (post conditions). Figure 7 shows an extract of the# proposed semantic representation of the service advertisements. An example pre-condition is depicted in Figure 8.#

#

The same service profile can be used to represent a request as well as the service# advertisements. A request description for a service consists of the hypothetical# description of the service that can perform the task needed by the user. These# request service pattern is sent to the registries of services for possible matching of# services, which can perform the#desired task. An example request service profile is# shown in figure#9.#

Geospatial Service Broker

The framework defines a broker (GSB) that acts as a central registry and manages# all metadata for available services [HK03]. It is remarkable for a broker#to be# capable of managing the metadata for geo-spatial functions as well as geo-spatial# data. It is also remarkable to accommodate OGC specifications in GSB in order to# guarantee the conformance to international standards. We have addressed the# following issues while designing the GSB#

- Flexible semantic matching methodology between service advertisement and# service request#
- Ranking the services based on semantic matching score and the QoS of the# services.#
- Matching process has to be efficient to avoid excessive delays in service# response.#

#

The major r QoS parameters for the web services include Performance, Reliability, Security. Ref. [RLBO03] divided the services QoS into three categories: system centric, process centric and information centric. In this paper, we are more concerned on the reliability of the information by ensuring that a user gets what they requested. The semantic similarity based on ontology described in the next section is aimed at providing semantic QoS for the GI services. We also provide a refining method for the services listed by the semantic similarity approach for further improvement in performance based on other QoS measurements.

```
<daml:Profile rdf:ID="RoadNetworkService">
 <Profile:comment> Road Network Service      </Profile:comment>
 <Profile:ServiceName> RoadNetworkService      </Profile:ServiceName>
 <input>
  <Profile:ParameterDescription rdf:ID="StateName">
   <Profile:ParameterName> state
   </Profile: ParameterName>
   <Profile:RestrictedTo
   rdf:resource="Concepts.daml#state"\>
  </Profile:ParameterDescription>
 </input>
 ..............................................................
 <output>
  <Profile:ParameterDescription rdf:ID="Raod_geom">
   <Profile:ParameterName> RoadGeometry
   </Profile: ParameterName>
   <Profile:RestrictedTo
   rdf:resource="SurfaceTransport.daml#road"\>
  </Profile:ParameterDescription>
</output>
..............................................................
   </daml:Profile>
```

Figure 7: DAML-S representation of the Semantic Signature of operations in service description

```
<daml:Pre rdf:ID="RoadNetworkServicePreCond">
 <Pre:comment>RoadNetworkService Pre Condition
 </Pre:comment>
 <input>
    <Pre: ParameterName ="state_name"
    </Pre: ParameterName>
 <Pre:RestrictedTo
 rdf:resource="Variable.daml#string"\>
<daml:Pre>
```

Figure 8: XML representation of the Semantic Signature of an operation

```
<daml:Profile rdf:ID="RoadNetworkService">
 <Profile:comment> Road Network Service      </Profile:comment>
 <Profile:ServiceName> RoadNetworkService      </Profile:ServiceName>
 <input>
  <Profile:ParameterDescription rdf:ID="StateName">
   <Profile:ParameterName> state
   </Profile: ParameterName>
   <Profile:RestrictedTo
   rdf:resource="Concepts.daml#state"\>
  </Profile:ParameterDescription>
 </input>
 ..........................................................
 <output>
  <Profile:ParameterDescription rdf:ID="Raod_geom">
   <Profile:ParameterName> RoadGeometry
   </Profile: ParameterName>
   <Profile:RestrictedTo
   rdf:resource="SurfaceTransport.daml#road"\>
  </Profile:ParameterDescription>
</output>
..........................................................
   </daml:Profile>
```

#

Figure#9: An example service request#

Limitations of UDDI registry#

There is a need for integration between the OGC and UDDI Service information# model. UDDI provides a registry of business and web services by their physical# attributes such as name, addresses and the services they provide. Since UDDI does# not provide the service capabilities information, it is very inefficient for searching# the services on the basis of what they provide. We observe that the adoption of# UDDI Specification in various domains such as GIS is slow, since existing UDDI# specification has following limitations.##

- First, UDDI introduces keyword-based retrieval mechanism. It does not allow# advanced metadata-oriented query capabilities on the registry.##
- Second, UDDI does not take into account the volatile behavior of services# and supports only quasi-static service registries.##
- Third, since UDDI is domain-independent, it does not provide domain-specific query capabilities in particular for GIS domain such as spatial# queries.##

#

Pure keyword based search fails to retrieve services which are described using# synonyms of the search string. Moreover, singular/plural word forms used in the# service description also affect the search result. Employing wild characters (e.g.# '%') for search helps to increase the recall rate, but necessitates human judgment# to filter out relevant services. Ideally, we would like to match two services if their# capabilities and requirements are the same. Even in interface specifications, the# terms used to describe matching services may be semantically related but not# necessarily identical. In other cases, the terms used may not appear to be#

#

semantically related at all. Thus general UDDI registry seems to be inefficient for describing GI services. Service similarity matching should address these seemingly diverse set of questions [AMJM97]:

- Retrieval: How can service be retrieved form a catalog of services based on its semantics, rather than on its syntactic structure?
- Reuse: How might services from a catalog of services fit the needs of a given application or more specifically a service request?
- Substitution: When can one service be replaced with another for a given service request without affecting the observable behavior of the entire system?
- Subtype: When is a service of one type a subtype of another?

The main feature of an enterprise system like the proposed E-GIS system is to locate information from distributed islands of information sources. Matchmaking by classical information retrieval models (e.g., Vector Space, Probabilistic, Boolean) [YRNB99] is based on lexicographic term matching. However, two terms can be semantically similar (e.g., can be synonyms or have similar meaning) although they are lexicographically different. Therefore, matchmaking by classical retrieval methods will fail to retrieve documents with semantically similar terms. Semantic Similarity relates to computing the similarity between concepts which are not lexicographically similar. The task of the *Match Maker* module is to use the knowledge of the world and the semantic understanding of the advertisements and requests to recognize the degree of similarity between them and retrieve those advertisements of services, which closely match the request.

As the Internet use for B2B reached some level of maturity, a significant interest in Internet-wide back-end integration has emerged. At the center of related efforts is the use of XML and other open standards as the foundation of B2B integration. ebXML, started as a UN/CEFACT initiative and supported by OASIS, was tasked with defining the new generation B2B. ebXML Registry/Repository (RR) contains registry as well as repository of elements, although the current emphasis of the specification is on protocol-related pointers rather than semantics [ebXML05]. Although the usability of ebXML has been charaterized in B2B domain, nothing has been specified relating to the semantic aspects of service description or advertisement. We, therefore, look at the other advanced technologies, which are aimed at providing the semantic interoprability in enterprise applications.

Semantic Matching of Service Description

For semantic based search of service, we need to understand the meaning of the service parameters i.e. we need to find similar services for service matching rather than exact matching. Since exact matching is almost impossible, what we need is the service advertisements those are sufficiently similar [PKPS02] to service request. Thus we make service matching method as flexible as possible where instead of searching for exact matches, we try to find service those provide services close in requirement of the requester.

As with *Signature Matching* approach for matching a software function or module from a software library, service matching is the process of determining the services

that matches a query service. This is a common approach in *Information Retrieval* domain. As we have pointed out, expecting a request service will match exactly with services in the registry is too strong a requirement. There may be some services that does not match exactly with the request service, but is similar in some way and hence could match a query if the query semantics are compared with the service parameters. This is what termed as *Relaxed Matching* between query and services in the registry. With this idea service matching can be defined more precisely as follows:

Service Match: Query Service, Match Algorithm, Advertised Service Library → Registry of Services

Service Match (q, M, A) = { a ∈ A : M (a, q)}

This means that given a query service q, a *Matching Algorithm* M, and a collection of service advertisements A, service matching returns a set of service, each of which staifies the service request.

A *Degree of Matching* parameter is used for finding the degree of similarity between services.

Application of Ontology for Matchmaking

Though business organizations are looking for seamless integration solution for GI services, they lack standards to expose expressive representations of their services. In this paper, we suggest means of overcoming this by providing richer descriptions about the services being offered. To facilitate understanding by any third party, these descriptions are expressed as a standardized conceptualization of the application domain (ontology). This is the core concept behind Semantic Web Services (SWS). In the proposed architecture for GSB, we explored the possibilities of using ontology that helps in finding the services based on the semantics of the request. An ontology is an explicit formal specification of a shared conceptualization [SBF98]. The application of ontology for service brokering has gained interest in recent times, especially in GI domain, where the semantics for the services are also of major importance [SBF98]. An ontology can be expressed as a hierarchical structure for deriving the concepts of the domain. The hierarchy specifies the subclass-super class nature of the concepts. Figure 10 shows one of such hierarchy. Formally ontology can be defined as follows:

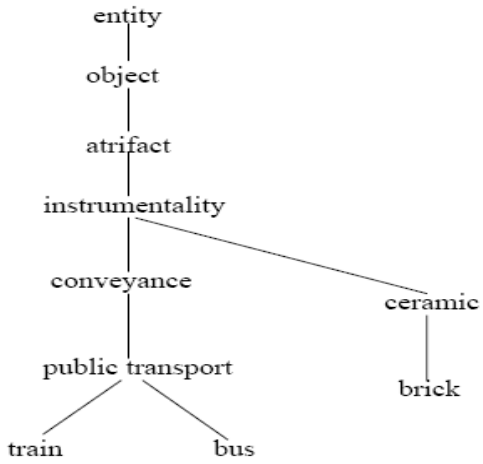
Let Ψ be a partially ordered set under ordering ≤. Then according to [SMSJ02] an ordering ⪯ defines a hierarchy for Ψ if the following three conditions are satisfied:

1. **$x \preceq y \Leftrightarrow x \leq y; \forall x, y \in \Psi$. We say (Ψ, ⪯) is better than (Ψ, ≤)),**
2. **(Ψ, ≤) is the reflexive, transitive closure of (Ψ, ⪯),**
3. **No other ordering satisfies (1) and (2).**

Let Ψ = {*SurfaceCommunication, RoadNetwork, NationalHighways*}. We can define the partial ordering ≤ on Ψ according to an is-a (or sub-class) relationship

#

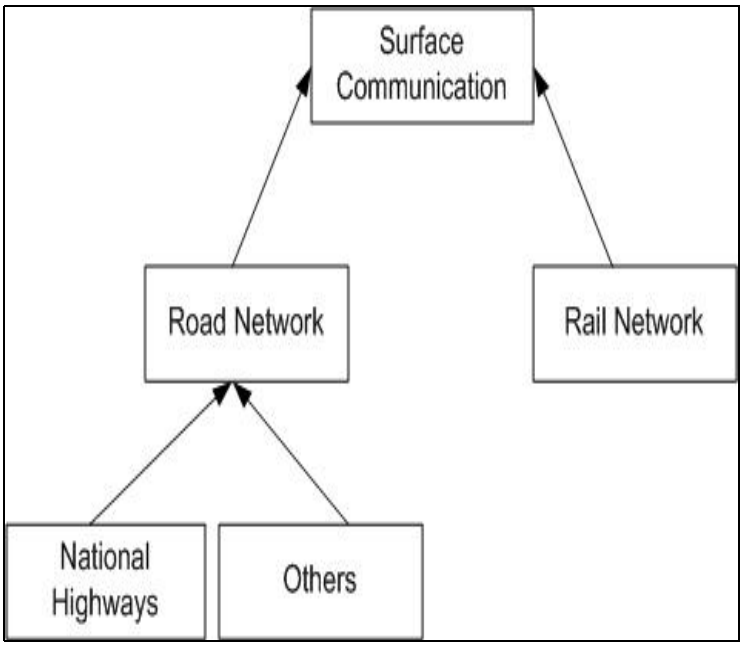
(Figure# 9). For example,# *NationalHighways*# is-a subclass of# *RoadNetwork*,# *RoadNetwork*#is-a sub-class of#*SurfaceTransport*#and, also#*NationalHighways*#is-a# sub-class of#*SurfaceTransport*.#
#



#

Figure#10: Semantic hierarchy among words as perceived by WordNet#
#

Ontology plays a key role in the Semantic Web by providing machine readable# vocabularies used by applications to understand the shared meanings. DAML+OIL# [HA02] is an ontology language that has been designed specifically to be used in the# Semantic Web and it is based on RDF and RDF Schema. DAML+OIL describe the# structure of a domain in terms of classes and properties. It also supports the use of# data types in class description. We propose the use of ontology at two levels#-# domain ontology,#application ontology. As in document retrieval, searching for# services should be on the notion of similarity rather than identical matches as we# are unlikely to find exactly matching descriptions.#
#



#

Figure#11: Partial ordering in Surface Communication Ontology to verify the usability of the service# in the context of the requested information#

#

Domain and Application Ontology

Semantically rich ontology for E-GIS, especially for service discovery purpose, is one of the major focuses of the paper. Here we design web services' domain-specific application ontology in order to achieve an agreement at the semantic level between various parties. On the other hand, domain ontologies for the entire domain are derived using an English thesaurus. These domain ontologies don't provide any information about any particular services. Rather, they use the query tokens for simplifying the query semantics. A domain independent source of clues gives us a breadth of coverage for common terms, while a domain specific ontology gives a depth of coverage by providing clues based on application specific terms and relationships. By combining these sources, we hope to improve both precision and recall.

Sense of the attributes form services are compared to find similar terms. This is achieved using thesaurus of synonyms for English language such as WordNet [MGA83]. WordNet is an on-line lexical reference system developed at Princeton University. WordNet attempts to model the lexical knowledge of a native speaker of English. It can also be seen as an ontology for natural language terms. It contains around 100,000 terms, organized into taxonomic hierarchies. Nouns, verbs, adjectives and adverbs are grouped into synonym sets (synsets). The synsets are also organized into senses (i.e., corresponding to different meanings of the same term or concept). Refer to [MGA83] for more detail on WordNet.

The matching procedures goes through a preprocessing step with help of domain-independent ontology, which includes the following observations in the service request parameters:

- The multi-term attributes in request are simplified. This is achieved by breaking the multiword attributes in simple words called tokens using language thesauri.
- The abbreviation expansion uses domain vocabularies by abbreviation expansion. Several expansion of a single attribute is possible.
- The synonyms of the request parameters are found out and used to enrich the service request for increasing the probability of a matching service.

With the available information after tokenizing the attributes in the input request similarity score is calculated for the matching parameters. The assigned score is based on both semantically matching and ontological matching. Let X is some of the parameters in the query and Y is the advertisement parameter in the available services. These parameters are first preprocessed as described above and all possible tokens are generated. Let p and q are the number of tokens generated from X and Y respectively. Then the semantic similarity **Sem (X, Y)** between X and Y has been taken from [MAS05] as

$$Sem(X,Y) = \frac{2 * Match(X,Y)}{(p+q)}$$

#

where Match (X, Y) is the same as defined in previous section. The services that# matches the query attributes are ranked on the basis of lowest number of# unmatched attributes. The overall match of a query service to repository service# $M_{sem,\ i}$# is given by#

$$M_i^{sem} = \min\left\{\frac{m_i}{n_i}, \frac{m_i}{|Q|}\right\}$$

#

where m_i# (1≤ i ≤ k)# is the number of attributes of services S1, S2,..., Sk that have# been matched out of |Q| query attributes using the semantic match score.# Also n_i# (1≤ i ≤ k) are the number of attributes present in services S1, S2,...,Sk in the# repository. Further, the values of# M_i^{sem} can be sorted to get a ranked list of# matching services. The service with# $\max\{M_i^{sem}\}$ is chosen as the candidate service# set.# If a service has similarity score less than some predefined threshold score# $M_{threshold,\ }$# are pruned from the possible list of services.# Further refinement is done# at the application ontology level.#

#

The application ontological# similarity is derived by an inference on the semantic# annotations associated with web service descriptions.# This matching is directly# influenced by application ontology description (DAML-S) of the services itself.# The# ranking procedure is applied on the candidate service set generated previous# process. Similar ranking is generated using# previous method# and may be written as#

$$M_i^{ont} = \min\left\{\frac{m_i}{n_i}, \frac{m_i}{|Q|}\right\}$$

#

The match between X and Y, where X and Y has the same meaning as in domain# ontology matching, is derived as Match (A, B) = C_{ij}# Where C_{ij}# = 1# # dist (i, j) and dist# (i,j) is the ontological distance between the jth term in service S_i# and a# corresponding query term.#

#

For each potential matching service S_i, let# M_i^{sem} be the matching score using# semantic matching and# M_i^{ont} be the matching score using ontological matching.# Then a new score can be computed for each potential matching service Si as#

#

$$M_i = \max\left\{M_i^{sem}, M_i^{ont}\right\}$$

Similarity Matching and Ranking#

When we say matching for services, matches generically stands for# **satisfies,**# **meets,**# or# **is equivalent to**. It is rarely the case that we would want one# component to match the other# **exactly**. In retrieval, we want a close match; as in# other information retrieval contexts [MMJL94]. Some services will match all the# requested parameters, while others will not. To distinguish between them, the# categorization can be done based on [RKPS02] as# **Exact, Plug-in, Subsumption**, and#

#

fail. A service matches a request when all the outputs the request match all the output of the service and all the inputs of the services match all the inputs of the request. Matchmaking here is defined as a process that requires to take a service query as input, and to return all the published advertisements which may potentially satisfy the QoS requirements specified in the input inquiry. For Matchmaking purpose, similarity matching of the services in the previous subsection is utilized. Formally the matchmaking can be specified as follows:

For a given inquiry P, the matchmaking algorithm should return a subset of the published advertisements which are compatible. Two service descriptions, say C1 and C2, are compatible iff their intersection is satisfiable by:

$$compatible\left(C1, C2\right) = \Leftrightarrow \neg\left(C1 \cap C2 \subseteq \perp\right)$$

The matchmaking procedure is based on function subtypes, which are the basis for the notion of safe substitution of software components [LB87, SG89]: A software component c that is a subtype of another component d can safely substitute this component, i.e. c can replace d without changing d's behavior. When applied to web service discovery, this means that an operation offered by a web service is a suitable match if it can be safely substituted for (i.e. is a subtype of) the operation described by the requester in her query. A service S can be categorized at different degree of match [RKPS02] for a query Q if the service's input concepts (C_{in},S) and its output concepts (C_{out},S) satisfy the following conditions:

Exact: When all the outputs of the request are same as those of the advertisements and all the inputs of the request are same as those of the advertisements. This is a more restrictive definition. We can also say matching is exact if the request outputs are *subClassOf* the advertisement outputs in the ontological concept hierarchy. Thus if request parameter for search is say, *RainFall* of any place, then a service *WeatherReport* is a super class of the request service. Thus we can say this is an exact matching. Formally we can say,

$$Match_{exact}\left(S,Q\right) \Rightarrow \left(C_{in},Q\right) = \left(C_{in},S\right) \wedge \left(C_{out},S\right) = \left(C_{out},Q\right) \text{ OR } \left(C_{out},Q\right) subClassOf \left(C_{out},S\right) \wedge \left(C_{in},Q\right) subClassOf \left(C_{in},S\right)$$

PlugIn: If the outputs of the request are a subset of the outputs of the advertisements, then we can say that the service advertisement can be *Plugged In* for the service request. This type of matching is said to be *Plug In* matching and comes after the *Exact* matching in the degree of matching hierarchy. For example, the service *NaturalHazards* may provide the information of the request that is looking for *FloodArea* information. Although we expect it to provide the expected information, it is not guaranteed. The formal definition is as follows

$$Match_{plug-in}\left(S,Q\right) \Rightarrow \left(C_{in},Q\right) \subseteq \left(C_{in},S\right) \wedge \left(C_{out},S\right) \subseteq \left(C_{out},Q\right)$$

Subsumption: If we just reverse the condition for the outputs of request and advertisements, a situation will arise where we will get those service

advertisements those partially fulfills request. This sort of matching comes next in the degree of matching hierarchy.

$$Match_{subsump}(S,Q) \Rightarrow (C_{in},S) \subseteq (C_{in},Q) \wedge (C_{out},Q) \subseteq (C_{out},S)$$

Fail: A matching between service request and advertisements is said to be *Fail* if none of the matching types gives any positive output. Thus

$$Match_{fail} \Rightarrow (C_{in},S) \cap (C_{in},Q) = \Phi \wedge (C_{out},Q) \cap (C_{out},S) = \Phi$$

MatchMaking Algorithm: In this section we discuss the algorithm for *MatchMaking*. First, we need to generate the candidate service set on the basis of the analysis of request inputs and advertisement inputs as well as request outputs to those of the advertisements. For this purpose, service request is compared against all the service advertisements from the registry of services.

```
Let S_cand = Set of all candidate advertisement
        services that can satisfy request
  A = Set of all service advertisements
  Q = Service request
  S_cand = ϕ
  AddService = TRUE

For each service advertisement a ∈ A do
 For each output of request q_out ∈ Q_out do
        If Match (a_out, q_out) then
        AddService = FALSE
        EndIf
 EndFor

 If AddService = TRUE
        S_cand = S_cand U a
        Group S_cand in the order of matching degree
 EndIf

EndFor

 Return S_cand
```

Scoring Algorithm: Scores are given to each matching services on a discrete scale where the **Exact** match, of course, gets highest score. **PlugIn** is the next candidate in the scoring scale. **Subsumes** comes next as it can fulfill user requirement partially. **Fail** is given a score of zero, as it cannot satisfy user request at all. The matching process makes use of the similarity score obtained using domain and application ontology. The ordering among the services is maintained in the following order

Exact > PlugIn > Subsumption > Fail

The candidate services S_{cand} are grouped on the discrete scale of four types. They are again ordered on the basis of the similarity score obtained by application of ontology based similarity matching process discussed in the previous section. The services with very low similarity score are pruned out. The algorithm for scoring the services S_{cand} in each groups are as follows:

For each candidate $s \in S_{cand}$ *do*

Calculate domain ontology matching score M_i^{sem}

Calculate application ontology matching score M_i^{ont}

Set $M_i = \max\{M_i^{sem}, M_i^{ont}\}$

If $M_i < M_{threshold}$ *do*

$S' = S_{cand} - s$ */*services being pruned which are less than threshold score*/*

EndIf

EndFor

Sort S' *on similarity score*

Return S'

With the definition of match degrees, we can use a DL reasoning engine to match a request. The ranking for the services are necessary for preventing the exploitation of of the system form the following cases

- A service may advertise in more generic way such that it appears that many of the requests can be served by those services. These sorts of services should have lower degree of matching as they do not serve any specific purpose.
- On the same manner, a requester can make a general request rather than exactly specifying what he expects.

These sorts of efficacy can be removed by measuring the degree of matching between the service advertisements and service request.

Precision-Recall Analysis

To evaluate the effectiveness of the propose algorithm for generating matched services, precision/recall studies have been carried out.

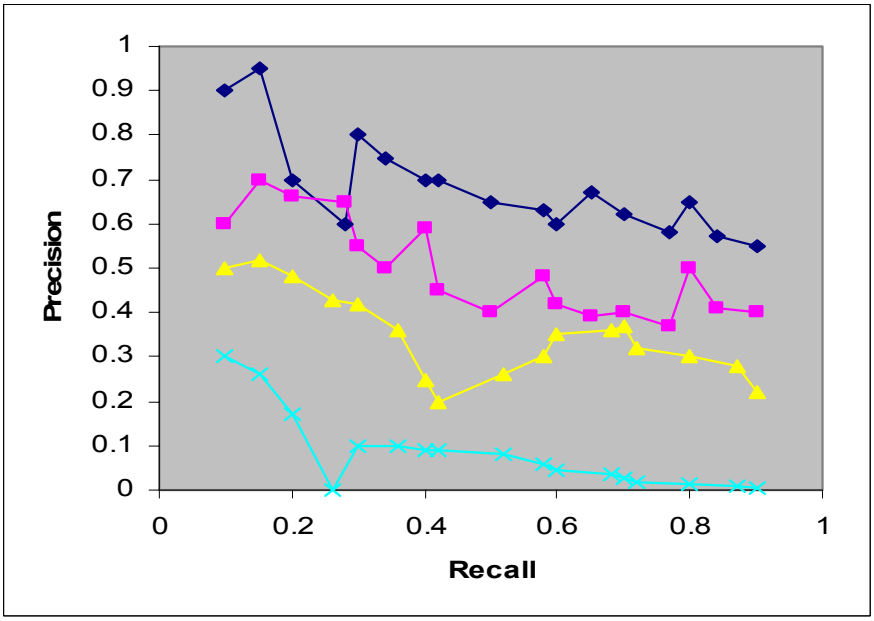
Precision is the percentage of relevant matches retrieved with respect to the number of retrieved services. Thus

$$\textbf{Precision} = \frac{\textbf{Matched Service}}{\textbf{Matched Service + False Match}}$$

Recall#that is, the percentage of relevant#matches#retrieved with respect to the# total#number of relevant#services#in the database.#

$$\text{Recall} = \frac{\text{Matched Service}}{\text{Matched Service + Missed Service}}$$

Due to the large size of the available GI services, it is practically impossible to# compare every query with each service. To compute recall, for each query, the# candidate services are merged and this set is considered to contain the total# number of correct answers. This is a valid sampling method known as pooling# method [VH98]. This method allows for relative judgments#(e.g., method A# retrieves 10% more relevant answers than method B) but does not allow for# absolute judgments (e.g., method A retrieved 10% of the total relevant answers).# The precision-recall studies have been carried out for different values of threshold# score and a graph as shown in figure 12 has been obtained. It has been observed# that threshold value#$M_{threshold}$# = 0.6#gives better precision-recall values. Decreasing# the threshold will give more false positives. On the other hand, increasing it above# the#threshold will lower the recall and also the possibility of getting semantically# similar services. A closer look into the results reveals that the efficiency of the# proposed similarity matching approach is mostly due to the contribution of non-identical#but semantically similar terms.#



Figure#12: Precission-Recall Analysis Diagram of the proposed method with VSM method#

QoS Considerations#

The Quality of Service (QoS) specifications of a web services characterize the# performance and other qualitative/quantitative aspects of services. We have# already discussed the detailed procedure for service similarity matching, which# also attributes to the QoS of service. Besides that we consider other QoS#

parameters that influence the services#-# Performance, Reliability, Security. Ref.# [WG99] divided the services QoS into three categories: system centric, process# centric and information centric. The performance, reliability, security, etc. are# located in the system centric category. These general QoS metrics are normally# needed in QoS description. The basic profile contains response time, cost,# reliability and throughput.##

- Response Time is defined as the total time needed by the service requester# to invoke the service. It is measured from the time the requester initiates# the invocation to the time the requester receives the last byte of the# response.#
- Cost represents the cost associated with the execution of the service. It is# necessary to estimate the guarantee that financial plans are followed. The# cost can be broken into major components, which include the service# execution cost and network transportation cost.#
- Reliability corresponds to the likelihood that the service will perform when# the user demands it and it is a function of the failure rate. Each service has# two#distinct terminating states: one indicates that a web service has failed# or aborted; the other indicates that it is successful or committed. By# appropriately designed redundancy, one can build highly reliable systems# from less reliable components.#
- Throughput represents the number of Web service requests served at a given# time period. It is the rate at which a service can process requests.#

#
Although DAML-S aims to make Web Services computer-interpretable and to enable# automated Web service discovery, invocation, composition and monitoring, the# specification has not provided a detailed set of classes, properties and constraints# to represent QoS descriptions. Because, GI services are distributed as well as# autonomous by their very nature, and can be invoked dynamically by third parties# over the Internet, their QoS can vary greatly. With the proliferation of GI services# and service providers, it is inevitable that there will be services offered by multiple# providers with the same functionality. In such scenarios,#the users should be able# choose the services with best QoS support out of the discovered services based on# some criteria. Thus, it is vital to have an infrastructure which takes into account# the QoS provided by the service provider and the QoS desired by#the service# requester, and ultimately find the (best possible) match between the two during# service discovery. In the proposed methodology, the services discovered by the# similarity based#*Match Making*#procedure, are further analyzed on the basis of the# QoS. For each discovered service, these QoS parameters will serve as a choice for# the client for selecting services on the basis of their QoS.#

QoS based Service Filtering#

We have already discussed about filtering and ranking of services based on# semantic similarity matching of attributes and operation.#Let#$\hat{S}$ = {S_1,• • • , S_n}# denote the set of services which are available (or registered with our system). Then# S'#⊆ $\hat{S}$#is the set of services#discovered in the semantic matchmaking procedure.# We perform the#QoS analysis on this set of services#S', rather than on the whole set#

#

of available services. This gives an improvement in the processing time for service discovery.

An ontology based on DAML-S can be constructed for QoS matchmaking purpose, similar to the ontology for semantic similarity matching of request and advertisements. Ontology can be designed describing the **QoSProfile** for the services. The development of QoS ontology allows customized QoS properties, and suitable QoS metrics for the QoS properties. The available property types, their domain constraints and range constraints are also defined as a part of the ontology for the **QOSProfile** design process. This consists of the **ProviderProfile** - the description of QoS ontology of the provider, **QueryProfile** - description of QoS ontology of the service requester, **QoSInput** - the input properties of the QoSProfile, **QoSOutput** - Output properties of the QoSProfile, **QoSPrecondition** - external conditions to be satisfied to ensure that service provider can provide the promised QoS level, **QoSPostCondition** - The expected effect that may occur due to the execution of the service. Thus we can say

$$QoSProfile \subseteq T$$
$$ProviderProfile \subseteq QoSProfile$$
$$QueryProfile \subseteq QoSProfile$$
$$QoSInput \subseteq QoSProfile$$
$$QoSOutput \subseteq QoSProfile$$
$$QoSPrecondition \subseteq QoSProfile$$
$$QoSPostCondition \subseteq QoSProfile$$

The above definition, however, has not provided any constraints over **QoSProfile** so that the **ProviderProfile** contains all the possible QoS combinations for the published service. No constraint means no useful information for the QoS matchmaking process and this makes no sense. The possible solution is to use property definition and cardinality to define the QoS constraints. After the definition of the QoS properties, cardinality constraints are ready to be added on these defined properties in **QoSProfile** for matchmaking purpose. We have created an ontology to represent the generic metrics, as well as domain specific QoS metrics. QoS of an operation j of a Web service i can be defined with the parameters on which the QoS is analyzed. This can be represented as follows where the interpretation of different terms are explianed in table 2:

$$QoS(si, oj) = \langle T(si,oj), C(si,oj), R(si,oj), A(si,oj), DS1(si,oj), DS2(si,oj), \ldots, DSN(si,oj) \rangle$$

T(s,o)	Execution time of Web service ‘s’ when operation ‘o’ is invoked
C(s,o)	Cost of Web service ‘s’ when operation ‘o’ is invoked
R(s,o)	Reliability of Web service ‘s’ when oper ‘o’ is invoked
A(s, o)	Availability of Web service ‘s’ when oper ‘o’ is invoked
DSi(s, o)	Service/operation level domain specific QoS metrics

Table 2: QoS aspects of service description

A service can then be defined as describing the semantics of the service operation (with the semantic observation on services as described in the previous sections)

and the QoS of those operations. Thus, the overall semantics of an operation are defined as:

OP(si,oj) = <F(si,oj), QoS(si,oj)>

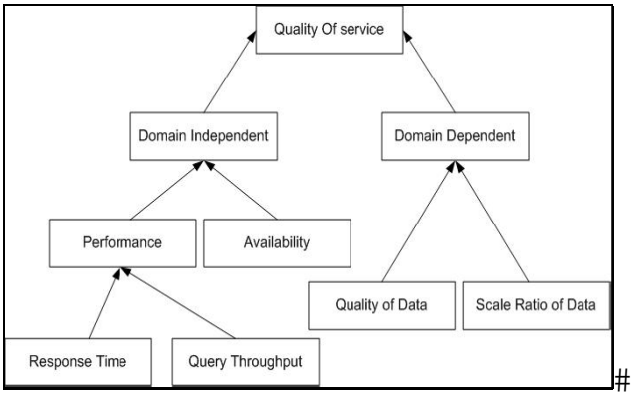
where F(si,oj) and QoS(si,oj) are functional semantics and QoS specifications of the required operation, respectively. The former tells what the service operation does while the latter tells how well it does it. Quality of Service can also be represented as a set of non-functional attributes that may impact the service quality offered by a Web service [JMS88]. In DL syntax, this advertisement can be written as

$$\begin{aligned} \textit{Advert} &\sqsubseteq \textit{QoSProfile} \\ &\sqcap (\leq 100\,\textit{cost:CostUSCentMetric}) \\ &\sqcap (\leq 20000\,\textit{responseTime:RespMSMetric}) \end{aligned}$$

A service request can also be specified with similar sort of DL syntax

$$\begin{aligned} \textit{Inquiry} &\sqsubseteq \textit{QoSProfile} \\ &\sqcap (\leq 500\,\textit{cost:CostUSCentMetric}) \\ &\sqcap (\leq 4000\,\textit{responseTime:RespMSMetric}) \end{aligned}$$

Towards improving the selection of services, we refine the candidate service set S′ based on several non-functional attributes to provide better quality service to the user. Some QoS property can be domain specific while others are not. Therefore, it essential to categorize them into: domain dependent and domain independent QoS. As an example, figure 13 shows the hierarchical categorization of QoS parameters. A detailed list and explanation about such attributes can be found in [SKF01]. When the service provider has built up a common QoS Profile for its specific service domain, this common QoS Profile can be defined as a template, which is a strong assistant for extensibility. Based on the template, service provider can expand the QoS Profile Layer, add more QoS properties and set stricter constraints over properties. Using the template to inherit the original QoS Profile avoids building the whole profile block from scratches. A similar algorithm to that of the semantic matching approach of service descriptions can be used for QoS matchmaking purpose.



Figure#13. QoS Ranking Hierarchy#

A quality matrix,#**Q = {V (Q_{ij}); 1#≤ i ≤ m; 1 ≤ j ≤ n}**, refers to a collection of# quality attribute-values for a set of candidate services, such that, each row of the# matrix corresponds to the value of a particular QoS attribute (in which the user is# interested) and each column refers#to a particular candidate service. In other# words, V (Q_{ij}), represents the value of the ith QoS attribute for the jth candidate# service. These values are obtained from the profile of the candidate service# providers and mapped to a scale between 0#to#1 by applying standard mathematical# maximization and minimization formulas based on whether the attribute is positive# or negative.# Let the Matchmaking algorithm discovered the the service# advertisements A1, A2 and A3 on the top score in similarity order.#The resepective# QoS is annotated with the service description as shown in table 3. Finally, it is to# the dicretion of the requester to select service based on the QoS provided.# #

#	A1	A2#	A3#
Performance#	#0.85#	0.30#	0.55#
Availability#	0.75#	0.95#	0.15#
Quality of Data#	0.45#	0.50#	0.90#

Table#3: Matrix for QoS based Ranking#

In a real world scenario, given a service request, it is conceivable that there exist# scores of service providers, which not only satisfy the functional requirements of# the requester, but also the#non-functional requirements. We use degree of match# to categorize (and implicitly order) the set of candidate service providers based on# the functional requirements of the user. We further refine each category and# select only those candidate service providers which satisfy the non-functional# requirements of the user.#

Conclusion#

With the increasing availability of GI services and the growing need of its# integration at enterprise level, the issue has been put in important concern. In a# real world scenario, given a service request, it is conceivable that there exist# scores of service providers, which not only satisfy the functional requirements of# the requester, but also other QoS requirements. As a result, it is also necessary to# consider the service quality#criteria for discovering a service.#

In this paper we investigate an efficient design for Enterprise-GIS (E-GIS) for# enhancing business process decision support. We have addressed how Semantic and# web Services technologies can be used to support service advertisement and# discovery. Although GI service based computing adheres much to the general# service based technology, some issue needs to be considered specific for the# purpose of GI services. In order to provide methodologies for discovering services# based#on both general and domain-specific search criteria, UDDI Specifications# seems promising, yet limited for GI domain. In particular, we describe the E-GIS# design and implementation of a service matchmaking prototype which uses a DAML-

S based ontology and a#Description Logic reasoner to compare ontology based# service descriptions. The proposed design of GSB and GSP has considered all these# issues to be efficient enough for using in the GI service domain. The paper# identifies that semantic search for services#based on similarity matching. The# domain ontology and application ontology has been used for ranking the services in# terms of degree of match. This along with the consideration of QoS further# improves the set of services for the serving request.#

To summarize, the paper mainly focuses on (1) designed the basic framework that# accommodates the OGC specifications in order to provide the integrated geo-spatial data in the standardized manner; (2) extended the basic framework for# semantic and QoS based searching of GI services.#

References#

[ABH02]# Ankolekar,#A.; Burstein, M.; Hobbs, J. et al: DAML-S: Web Service Description for# the Semantic Web.#*In Proceedings of the 1st International Semantic Web# Conference (ISWC)*, 2002#

[AMJM97]# ZAREMSKI A. M.; WING J. M.: Specification Matching of Software Components,#*ACM# Transactions on Software Engineering and Methodology*, Vol. 6, No. 4, October# 1997, Pages 333-369.#

[BB99]# Buttenfield, B.: Sharing Vector Geospatial Data on the Internet,#*Conference of the# international Cartographic Association*, 1999, PP.35-44#

[BE01]# Bertolotto, M.; and Egenhofer, M.: Progressive Transmission of Vector Map Data over# the World Wide Web, Proc. GeoInformatica, vol. 5, PP.345-373, 2001#

[BG03]# Brickley, D.; Guha, R.V.: RDF vocabulary description language 1.0: RDF schema.# *W3C Proposed Recommentation*.#Available at http://www.w3.org/ TR/rdf-schema/,# accessed on Dec. 2003.#

[BHL01]# Berners-Lee, J.; Hendler, J.; Lassila, O.: The Semantic Web,#*Scientific American,# vol 184, no. 5, pp. 34-43*, 2001.#

[BK02]# Bernstein, A.; and Klein, M.: Towards high-precision service retrieval.#*In First# International Semantic Web Conference (ISWC 2002)*, pages 84–101, Sardinia, Italy,# 2002.#

[BLV87]# B. Liskov. Data abstraction and hierarchy. ACM Sigplan Notices, 23(5):17–34, 1987.#

[BN03]# Baader,#F.; and Nutt, W.: Basic description logics. In F. Baader, D. Calvanese, D.# McGuinness, D. Nardi, and P. Patel-Schneider, editors,#*The Description Logic# Handbook*. Theory, Implementation and Applications, pages 43–95. Cambridge# University Press, Cambridge, 2003.#

[CGK02]# Curbera, F.; Goland, Y.; Klein, J. et al : Bpel4ws white paper . http://www-3.ibm.com/software/ solutions/ webservices, 2002#

[CRL95]# CRUZ, R. L. 1995. Quality of service guarantees in virtual circuit switched networks.# IEEE J.#Select. Areas Commun. 13(6): 1048-1056.#

[DAML01]# DAML Joint Commitee. DAML+OIL Language, http://www.daml.org/2001/03/daml+# oil-index.html, March 2001.#

[DFK02]# Dan,#A.; Franck,#A. R.; Keller,#A.; King,#R.; Ludwig,#H.: Web Service Level# Agreement (WSLA) Language Specification",# *http://dwdemos.alphaworks.ibm.com/wstk/common/wstkdoc/services/utilities# /wslaauthoring/ WebServiceLevelAgreementLanguage.html*, 2002.#

[ebXML05]# *EbXML specifications*#at#http://www.ebxml.org/, 2005#

[EJ02]# Egenhofer, J.: Toward the Semantic Geospatial Web.#*10th ACM International# Symposium on Advances in Geographic Information Systems*. McLean (USA). 2002.#

[ET04]# Erl, T.: Service-Oriented Architecture: A Field Guide to Integrating XML and Web# Services.#*Prentice Hall, New Jersey*, 2004.#

[GM99]# M. Gahegan. Characterizing the Semantic Content of Geographic Data, Models and# Systems. Interoperating Geographic Information Systems (1999), pages 71–84.#

[HA02]# Horrocks, I.: DAML+OIL: a reason-able web ontology‡language. In Jensen, C.S.,# Jeffery, K.G., Pokorn´y, J., Saltenis, S., Bertino, E., B¨ohm, K. and Jarke, M. eds.# Proc. of EDBT 2002, number 2287 in Lecture Notes in Computer Science, pages 2–13.# Springer, Mar. 2002.#

[HK03]# Hess, A.; and Kushmerick, N.: Learning to attach metadata to Web services,‡*In‡the# Proceedings of‡International Semantic web conference(ISWC)*, 2003.#

[JLC03]# J. Riecken, L. Bernard, C. Portele, and A. Remke, North-rhine westphalia: Building a# regional sdi in a cross-border environment / ad-hoc integration of sdis: Lessons# learnt. In 9th EC-GI and GIS Workshop‡‡ESDI: Serving the User, A Coru˜na, Spain,# 2003.#

[JMS88]‡ Radatz, J.; Sloman, M. S.: A Standard Dictionary for Computer Terminology: Project# 610,‡*IEEE Computer*, 21(2), 1988.#

[KRH01]# Kreger, H.: Conceptual Architecture (WSCA 1.0). IBM Technical White Paper.# http://www-‡3.ibm.com/software/solutions/webservices/pdf/WSCA.pdf, 2001#

[KS98]# V. Kashyap and A. Sheth. Semantic Heterogeneity in Global Information System: The# Role of Metadata, Context and Ontologies. In M. Papazaglou and G. Schlageter,# editors, Cooperative Information Systems: Current Trends and Directions (1998),# pages 139–178. Academic Press.#

[LT92]# Laurini, R.; and Thompson,‡D.: Fundamentals of Spatial Information Systems,# London, Academic Press, 1992#

[MDL98]# McGuinness, D.L.: Ontological issues for knowledge-enhanced search. In Proceedings# of Formal Ontology in Information Systems, 1998. Also published in Frontiers in# Artificial Intelligence and Applications, IOS-Press, 1998.#

[MGA83]# Miller, G.A.: WordNet: A lexical database for the English language,‡*In Comm. ACM*# 1983.#

[MMLJ94]# Mauldin, M.; Lesvitt, J.: Web-agent related research at the CMT, ‡*In ACM Special# Interest Group on Networked Information Discovery and Retrieval (SIGNIDR-94)*# (Aug.). ACM, New York, 1994.#

[MRDB03]# Mart´ın-D´ıaz, O.; Ruiz-Cort´es, A.; Dur´an, A.; Benavides, D.; and Toro, M.: # Automating the Procurement of Web Services.‡*In‡the‡Proceedings‡of the First# International‡Conference‡on Service Oriented Computing (ICSOC 2003)*, Trento,# Italy, December‡15-18 2003.#

[MS04]# Maximilien E.M.; Singh M.P.: A Framework and Ontology for Dynamic Web Services# Selection,‡*IEEE Internet Computing*, vol. 8, no. 5, 2004, pp. 84–93.#

[MSA05]# Mahmood, T. S.; Shah G.; Akkiraju, R.; Ivan, A. A.; Goodwin, R.: Searching Service# Repositories by Combining Semantic and Ontological Matching,‡*In the Proceedings of# the IEEE International Conference on Web Services (ICWS'05)*, 2005#

[MSZ01]# McIlraith, S.; Son, T.; and Zeng, H.: Semantic Web Services.‡*IEEE Intelligent# Systems*, 16(2):46–53, 2001.#

[MW95]# Zaremski, A. Moormann; Wing, J. M.: Signature Matching: a Tool for Using Software# Libraries,‡*ACM Transactions on Software Engineering and Methodology*, 1995#

[OCS01]# Open Geospatial Consortium Inc, Catalog Services Specification, version 1.1,‡28-March 2001#

[OGC03]# Open Geospatial Consortium Inc, Open Geospatial Reference Model, version 0.1.2,# 04-March 2003#

[OGM02]# Open Geospatial Consortium Inc, Geography Markup Language Implementation# Specification, version 2.1.2, 17-September 2002#

[OSL02]# Open Geospatial Consortium Inc: Styled Layer Descriptor Implementation# Specification, version 1.0.0, 19-September 2002#

[OWC02]# Open Geospatial Consortium Inc, Web Coverage Service Implementation# Specification, version 0.9, 18-December 2002#

[OWF02]# Open Geospatial Consortium Inc, Web Feature Service Implementation# Specification, version 1.0.0, 19-September 2002#

[OWM02]# Open Geospatial Consortium Inc, Web Map Service Implementation Specification,# version 1.1.1, 16-January 2002#

[OWR03]# Open Geospatial Consortium‡Inc, Web Registry Service Implementation# Specification, version 0.7, 18-January 2003#

[PGa06]# Paul,#M.;#Ghosh,#S. K.:#A Service Oriented Architecture for Enterprise-GIS.#*In the# Proceedings of 9th Annual International Conference and Exhibition on Geographic# Information Science, Technology and Application (Map India 2006)*, New Delhi, India,# 2006.#

[PGb06]# Paul,#M.;#Ghosh,#S. K.: Enterprise Geographic Information System (E-GIS): A Service-based Architecture for Geo-spatial Data Interoperability.#*In the Proceedings IEEE# International Geoscience And Remote Sensing Symposium (IGARSS 2006)*, Denver,# Colorado, USA, July-August 2006 (To Appear)#

[PGc06]# Paul,#M.;#Ghosh,#S. K.: An Approach for Service Oriented Discovery and Retrieval of# Spatial Data.#*In the Proceedings of International Workshop on Service Oriented# Software Engineering (IW-SOSE '06), In conjunction with the 28th International# Conference on Software Engineering (ICSE 2006), ACM Press*, Shanghai, China, May# 2006#

[PG05]# Paul,#M.;#Ghosh,#S. K.:#A Framework#for Web-enabled Distributed GIS.#*In the# Proceedings of 8th Annual International Conference GIS, GPS and Remote Sensing# (Map India 2005)*, New Delhi, India, 2005#

[PKPS02]# Paolucci, M.; Kawamura, T.; Payne, T.; Sycara, K.: Semantic Matching of Web# Services Capabilities,#*In the Proceedings of 1st International Semantic Web# Conference*, 2002.#

[PT06]# Peristeras, V.; Tarabanis, K.: The Connection, Communication, Consolidation,# Collaboration Interoperability Framework (C4IF) For Information Systems # Interoperability.#*International Journal of Interoperability in Business Information# Systems*## Issue 1 (1), 2006#

[RBPA03]# Riecken, J.; Bernard, L.; Portele, C.; and Remke,#A.: North-rhine westphalia:# Building a regional sdi in a cross-border environment / ad-hoc integration of sdis:# Lessons learnt.#*In 9th EC-GI and GIS Workshop*## *ESDI: Serving the User*, A Coru˜na,# Spain, 2003.#

[RGM00]# R. Groot and J. McLaughlin, editors. Geospatial data infrastructure## Concepts,# cases, and good practice. Oxford University Press, Oxford, 2000.#

[RLBO03]# Rocha,#A.; Lopes, J. C.; Bartolo, L.; and Oliveria, M. A.: An Interoperable GIS# Solution for the Public Administration.#*In the Proceedings of. EGOV#2003*, LNCS# 2739, PP.345-350, 2003#

[RP97]# Roth, M.T.;#and Schwarz,#P.:#Don't scrap it, wrap it! A wrapper architecture for# legacy data sources.#*In Proceedings of the International Conference on Very Large# DataBases*,#1997.#

[RS05]# Richardson, R.; and Smeaton, A.F.: Using WordNet in a knowledge-based approach# to information retrieval.#*Dublin City University School of Computer Applications# Working Paper*#CA-0395, 2005.#

[SBF98]# Studer,#R.; Benjamins, V.; and Fensel, D.: Knowledge engineering: Principles and# methods.#*Data and Knowledge Engineering*, 25(1-2):161–197, 1998.#

[SG89]# Salton, G.: Automatic Text Processing: The Transformation Analysis and Retrieval of# Information by Computer.#*Addison-Wesley*, 1989.#

[SKF01]# Salmatian, K.; and Fdida, S.: Measurement based modeling of quality of service in# the Internet: a methodological approach, IWDC 2001:158-174., 2001#

[SMSJ02]# Sahai, A.; Machiraju, V.; Sayal, M.; Jin, L. Jie.; Casati, F.: Automated SLA# Monitoring for Web Services,#*http://www.hpl.hp.com/techreports/2002/HPL-2002-# 191.pdf*, 2002.#

[SOAP01] ## Simple Object Access Protocol (SOAP)#http://www.w3.org/TR/2003/REC-soap12-# # # part0-20030624/#

[TBP02]## Trastour, D.; Bartolini, C.; and Preist, C.: Semantic web support for the business-to-# # business e-commerce lifecycle. In Proceedings of the Eleventh International## # # conference on World Wide Web, pages 89–98. ACM, 2002.#

[TD05]## Tidwell, D.: Web services: the Web's next revolution.#*IBM Developer#Works*,## # # www.ibm.com, 2005.#

[TG01]## Taldoire,#G.:#Geospatial Data Integration and Visualisation Using#Open Standards.# # # *7th EC-GI & GIS WORKSHOP (2001)*, Potsdam, Germany, June 13-15#

[UDDI02]# UDDI technical paper,#http://www.uddi.org/pubs/Iru_UDDI_Technical_White_# # # Paper.pdf#

[VH98]## Voorhees, E.; Harmann, D.: Overview of the Seventh Text REtrieval Conference# (TREC-7).#*In NIST Special Publication 500-242: The Seventh Text REtrieval* ## *Conference (TREC-7)*,#pages 1-23, 1998.#

[W3C03]# W3C Consortium:#Web Services Description Language version#1.2, 11-June 2003#

[WG92]# Wiederhold, G.:#Mediators in the architecture of future information systems.#*IEEE*# *Computer*, pages 38-49, 1992#

[WG99]# Weikum, G.: Towards guaranteed quality and dependability of information service.# *In 8th GI Fachtagung: Datenbanksysteme in Buero*, Technik und Wissenschaft, 1999.#

[WSDL02]# Web Services Description Language (WSDL) (WSDL) http://www.w3.org/TR/wsdl#

[YRNB99]## Yates, R.B.; Neto, B.R.: Modern Information Retrieval.#*Addison Wesley Longman*,# 1999.#

[ZAWJ95]# Zaremski, A. M.;##and#Wing, J. M.:#Signature Matching: a Tool for Using Software# Libraries.#*ACM Transactions on Software Engineering and Methodology*, Vol. 4 No. 2,# 146-170, Apr. 1995.#

Manoj Paul# has received a Post Graduate Diploma in# Information Technology form Indian Institute of Technology,# Kharagpur, India. He is currently pursuing MS in Information# Technology from the same institute.#He is working in the# area fo spatial data analysis and interoperability for the last# three years.#His research interests include Spatial Data# Analysis,#Spatial#Web Services, Semantic Geospatial Web.# He is a student member of IEEE.#

S. K. Ghosh#is presently working as Assistant Professor in# the School of Information Technology, Indian Institute of# Technology, Kharagpur, India. Prior to IIT Kharagpur, he# worked for Indian Space Research Organization (ISRO),# Department of Space, Government of India, for about eight# years in the field of Satellite Remote Sensing and GIS. His# research interest includes Image Processing, Computer# Networking, Education Technology, Remote Sensing & GIS.#