CHAINING DISTRIBUTED GIS SERVICES

Trias Aditya *) and Rob Lemmens **) 

Abstract

The fast development of internet technology has motivated the improvement of geographic information sharing and now, the advent of web services enable GIS functionality to be shared and consumed in a distributed computing environment. Web services are loosely-coupled functions that can be executed remotely by users on the internet regardless the platforms implemented. This concept supports the technical requirement in developing the Spatial Data Infrastructure; in which distributed information and functions supplied by the providers can be integrated and distributed to the users as services. Of these services, a new service can be potentially gained by means of combining several services to achieve a specific task. This paper will present the scenario in developing the SDI as service architecture and solving a geographical problem-oriented with a service chaining approach. Here, the scenarios of chaining for national disaster management and for incorporating commercial services in the daily activities will be discussed. To improve the possibility of requester in investigating of the chaining, the metadata of services should deal with semantic aspect of the service definition. The semantic aspect is the highest level of interoperability in resolving information integration. It deals with the meaning of resources and their metadata. The metadata of a service contains information about input and output of the service. Ontology then is used as a reference for the input and the output. In such setting, distributed components within the infrastructure will have the same interpretation about the input and the output. Thereby the chaining, which is pipelining the output of preceding service to the input of following service, can be better examined.

Keywords: GIS Services, SDI, service chaining, ontology, semantic interoperability.

*) Author is lecturer at Dept. of Geodetic Engineering, UGM, Indonesia. 
**) Co-author is assistant professor at Dept. of Geo-Information Processing, ITC, the Netherlands.

Abstrak

1. INTRODUCTION

Presently, with the advance of information technology, geographic data and geographic functionalities (known as components) can be shared and consumed over the web. One of the major backings is the maturity of infrastructure for information exchanging on the web. The infrastructures are geared towards interoperable platforms. This can be seen from the development of W3C’s standards on the areas of: networking, middle-tier, and interchange formats (i.e. HTTP GET/POST/SOAP, Web Services, and XML respectively). Based on XML, the paradigm of web services now is emerging as the recent deployment for consuming components. In business applications, the web services are implemented from bookseller up to café finder services on the web. In the field of GIS, the concept of web service is implemented in the OpenGIS Service Architecture by the OGC. The draft specification of this architecture has been implemented to support the Spatial Data Infrastructure (SDI), such as in WestPhalia Germany (Bernard, 2002). The concept of web service is also adopted in the framework of the Canadian Geospatial Data Infrastructure (Geoconection, 2001).

This paper will describe the scenario of web services and services chaining deployment to support the SDI. Main tasks that need to be solved here are to model the service chaining and examine its implementation with considering the aspect of semantic interoperability. In this introduction the concepts of XML web services, service chaining, and SDI are discussed briefly. The second section will provide the conceptual model of service chaining in the SDI framework. Subsequently, the results of the research will be presented. The discussions and the suggestions for further research will conclude this paper.

1.1. Spatial Data Infrastructure (SDI)

The (Geo) Spatial Data Infrastructure is containing networked spatial databases and its technologies of data sharing, access and data uses across institutions with their complexities through a clearinghouse function (Groot & McLaughlin, 2000). The clearinghouse is known mechanism for the user to search and find spatial data which are distributed on the data provider’s servers. Commonly this mechanism is based on metadata management for each thematic spatial data of the providers. Hence, the basic foundations for a SDI are networking of the distributed spatial databases and the metadata. The obvious advantage of employing this architecture is the cost-benefit of reusing the available spatial data. The institutions, which need particular spatial data, can find & purchase the related spatial data via the SDI’s clearinghouse. This approach is better rather than conducting a new data acquisition from the scratch. To build the networked spatial databases for interchanging information, there is a barrier that must be handled: the infrastructure must capable of overcoming interoperability among the involved computers. This requires suitable protocols that can easily move in and out the firewalls to carry out the requests and the responses. Departing from this need, the web service concept is becoming appropriate solution to build the SDI (Geoconections, 2001). In this direction, with using the OGC’s framework, there is a possibility to extend the perception of the SDI. The SDI can also be a place to find and use the services (functionality & data components), not just about finding and using the spatial data (see the OGC 2002, Bernard 2002). Figure 1 shows the overview of the SDI’s components.

Figure 1. The components of SDI (Nebert & Lance, 2002).

1.2. GIS Services

Interoperability situation can be achieved only if and when distributed software components are able to deal with the invocation tasks reciprocally in the same way, and they can handle the obstacles caused by the variety of processing and data (ISO19119, 2001). The web services promise seamless interoperability between applications, in spite of platforms used (Schmelzer 2002).

Web services are loosely-coupled functions that can be executed remotely by users on the internet regardless the platforms implemented. This technology can be seen as an improvement of middle-tier technology (i.e. CORBA, DCOM) in the distributed computing
environment (Schmelzer, 2002). The web services communicate via XML-based protocols. XML (eXtensible Markup Language) is a W3C specification which is human readable, text-based format, self-describing, and firewall friendly. Those benefits make the XML is superior to the other middle-tier protocols mentioned above.

The concept underlying of the web services's protocols can be seen in Figure 2. This so called SOA (Service Oriented Architecture), comprises three components (provider, requester, broker) and three operations (publish, find, bind). SOAP (Simple Object Access Protocol) is a standard protocol to remotely bind component in the service provider regardless the provider's platform. The service requester needs WSDL (Web Services Description Language) to specify how to consume the web services. This include specifications of which message should be sent and to which port it should go. Next protocol is UDDI (Universal Description, Discovery, Integration). This is used to publish the service in order to make the requester (user) can find the service. All these standards are XML-based protocol.

![Figure 2. The SOA concept (Flurry, 2000).](image)

The benefits of the web services have motivated a lot of GIS functionalities to be provided as the web services. Terminology of GIS Services in this paper refers to GIS or geo web services. It is the web services with geographic interests that have capabilities to perform GIS functionalities (ESRI, 2002). The domains of this application now are ranging from commercial sector up to institutional level. In commercial sector for example, now can be seen many applications make use of ESRI Arc Web Services or developed using Microsoft MapPoint .NET. A lot of geocoding services and map services have been implemented over the internet (see URL: services brokerage3 or visit Microsoft’s or IBM’s UDDI4).

The OGC, an international consortium which endorses the interoperability program for GIS, formulates the GIS service concept in their OpenGIS Services Architecture. In this framework, GIS services can be of: data, function, or catalog services. Albeit it heavily adopts the W3C standards, the OGC has its own version of standards in support of the GIS web services (see Table 1).

<table>
<thead>
<tr>
<th>operations in web services</th>
<th>Industrial standards (W3C / UDDI org.)</th>
<th>GIS community (OGC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>binding</td>
<td>HTTP GET / POST SOAP</td>
<td>HTTP GET / POST SOAP-RPC</td>
</tr>
<tr>
<td>describing</td>
<td>WSDL</td>
<td>GetCapabilities doc.</td>
</tr>
<tr>
<td>brokering</td>
<td>UDDI</td>
<td>Registry Services (based on ebXML)</td>
</tr>
</tbody>
</table>

Clearly, the concept of web services promotes the need of interchanging and interacting openly across the platforms. These benefits support the technical requirements in developing of the SDI.

1.3. Service Chaining

It is likely that users have possibility to make a combination of one or more the GIS services into their applications. Without having to buy a complete license of GIS software, it is viable to bring GIS technologies more straightforward for the users.

A service chaining can be interpreted as pipelining services in a dependent series to attain a customized request. There are three types of the service chaining i.e.: user-defined, aggregate, and workflow-managed service chaining (ISO 19119, 2001). The first one insists the users to have prior knowledge of interfaces, inputs, and outputs of the services used. With the second one, the services appear as a single aggregate service which handles the coordination of the individual services in a chain. They perform static chain of services and present the result to the user as one. This approach has some drawbacks, in which the user loses some of flexibility and control over parameters of the individual services, and it can give false results (Alameh, 2002). The last one uses mediating services as gateways to coordinate several services without necessarily storing any data of their own. The user needs to invoke pre-specified workflow parameters which are provided by mediating service in order to have the requested result (ISO19119, Microsoft UDDI registry at http://uddi.microsoft.com/).

3 See at [http://www.salcentral.com/](http://www.salcentral.com/) as an example
4 IBM UDDI Registry at [https://uddi.ibm.com/ubr/find](https://uddi.ibm.com/ubr/find)
There is a need to have a discovery mechanism that can manage and present expected metadata about the service. Service mediators intend to provide the users with solutions that are explicitly met to their requirements (Alameh, 2002).

In the field of the web services, to mediate the services is a task that belongs to broker (registry service), at least for the syntactic case, not yet for the semantic case. There are several types of broker that can be used to locate the web service, such as: UDDI, ebRIM (ebXML Registry Information Model), and brokerage. The broker is a service intended to provide a common mechanism to classify, register, search and access descriptions of the services (the OGC, 2002).

2. SERVICE CHAINING IN THE FRAMEWORK OF SDI

The SDI we mean here is an infrastructure of geographic information services (see Bernard 2002 & 2003). Then, the content of the infrastructure is not merely about the spatial data. Referring to the OGC paradigm, it can be a complete framework of data, functionality, and registry services (the OGC, 2002). In such setting, intuitively the SDI is encompassing the providers of the components (i.e. GIS data and GIS functions) and a broker as a subset of the clearinghouse functionality. On the whole, the term of service chaining itself is loosely applied on demand. That means it is used generally in the context of business execution on the internet. Here, the chaining will be introduced for pipelining the services in the context of SDI.

2.1. The use of service chaining

The implementation of four-tier concept in the web services is benefitting for distributed computing environment. Since it is not a centralized server-client, the components on the web (i.e. the services) can be reused and then chained with the other service(s) by any computer that is granted for the access. With the chaining, regardless of the type used, the possible solution that can be achieved is depending on the availability of services and their granularities. There is a possibility that a single service actually is implicitly composed by several services\(^5\). In another case, it can happen that the run-time of a process will be longer when the result of the preceding two services then is chained to subsequent service and another subsequent service. Hence, the built chaining between proximity and geocoding services (with different granularity in their implementation), as shown in figure 3., can be made longer by sending the result of chaining to the subsequent service (e.g. Route Finder).

Figure 3 describes the flow process for the scenario of finding the nearest restaurant for the traveler using the clearinghouse of SDI. In this respect the SDI is considered as the infrastructure for interacting with the services which are provided by either governmental agencies or private companies. The scenario is given as follows: a traveler who visits Jakarta can access the proximity service on the web to find the nearest restaurant from his/her standing point (input used is street address, since this is only the geographic element he/she knows about regarding to his/her position). Then the suitable result of the nearest restaurant is chained to the following geocoding service. Assuming the geocoding service on the web is providing latitude and longitude, then these values can be used further for another chaining. Just for example: the result is combined with a route finder to determine the best route to get the restaurant or to be combined with the proximity service again to determine the nearest parking parcel near the restaurant.

\(^5\) See Storelocator of MapPoint .NET at http://demo.mappoint.net/storelocator/

![Figure 3. One possible solution of chaining proximity and geocoding within SDI's framework.](image-url)
resources and their metadata in a distributed environment.

Now there is another example for the use of chaining. Suppose that during a major flooding of Jakarta, a disaster management team wants to get a clear picture of the damaged and endangered areas. Up-to-date information has to be combined and this information may come from several disparate sources. The flood extent has to be compared with the actual locations of population concentrations. The first is derived from scattered observations in the field and most of them are based on GPS coordinates in latitude, longitude. The population data is stored in a database with local coordinate reference. For the geographic comparison we need a coordinate transformation service that converts lat,lon into the local system. Subsequently an overlay operation has to be able to combine the data.

It is possible to simply build a static application that is dedicated to solve this demand which means the disaster management team no need to worry about the chaining (with aggregated services approach). Even, this built application can be seen as a new service that can be combined with another suitable service for a specific task. However, the metadata for each single service is an important element in that chaining. Therefore description of inputs, outputs and necessary information we need to know should be described clearly.

For example purpose, coordinate transformation and overlay operation are not aggregated in one application, rather loosely used for another possibility of chaining. This application is bound in a framework of SDI as seen in Figure 4.

2.2. Chaining for interoperability

For building an aggregated and static application, a developer can use any services development framework such as .NET or Java (see Chang, 2003). While for a larger task and to allow many services involved in a workflow model, there are some languages available. The recent one is BPEL4WS. This language is a merger between WSFL of IBM and XLANG of Microsoft. In workflow chaining, WSDL definitions on messages & ports are used as the basis for collaborating services (Leymann, 2001). Figure 5 shows the possibility in using WSFL to make a chaining between proximity and geocoding. These sample codes of chaining based on two WSDL documents of proximity & geocoding respectively.

```
<activity name="ProximityAnalysis">
  <input message="mp:FindNearby"/>
  <output message="mp:FindNearbyResponse"/>
  <performedBy serviceProvider="StoreLocator"/>
  <implement>
    <export>
      <target portType="mp:FindServiceSoap" operation="mp:FindNearby"/>
      <export>
        <implement>
        <activity>
```

```
<activity name="GeocodingNameOfPlace">
  <input message="pf:FindPlace1n"/>
  <output message="pf:FindPlace1nOut"/>
  <performedBy serviceProvider="ESRIPlaceFinder"/>
  <implement>
    <export>
      <target portType="pf:PlaceFinderSample" operation="pf:findPlace1"/>
      <export>
        <implement>
        <activity>
```

Figure 5. Chaining definition with WSFL with based on messages and portTypes of WSDL.

In the open environment, such as in the GDI environment, a lot of users and providers are participating. It will be difficult for the user to know all about output or input of services.

To solve questions and requirements indicated in section 2.1. we describe each service with a structure as shown in table 2. It is based on the implementation of UDDI. The service basically is described based on its identity and its functionality. Identity elements are based on ISO19119 specification. Functionality elements are based on DAML-S specification (Aditya, 2003).

<table>
<thead>
<tr>
<th>Type</th>
<th>Identification (ISO19119 based)</th>
<th>Functionality (DAML-S based)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main elements</td>
<td>ServiceType, SpatialReference, OperationCategory, DataCategory, AccessProperties.</td>
<td>Provided By, Geographicradius, ServiceCategory, Input, Output.</td>
</tr>
</tbody>
</table>

Table 2. Structure of metadata about service

Figure 4. Modeling interaction sequences of service chaining with UML (Unified Modeling Language).
For collaborating services, input and output elements for each service are vital information. Unfortunately, the elements of input and output of services can have different tag names for referring the same object. This is where we propose the use of ontology for the services. An ontology associates to a taxonomic hierarchies of classes and it is combined with inference rules (Gruber, 1993). Classes and attributes within ontology in this research are defined with the ontology language (this work uses DAML+OIL to build the ontologies of service, operation, and data). The input and output of service are referenced to particular classes in the data ontology defined. DAML+OIL is an ontology language which is based on RDF and RDF Schema (DAML Org, 2002). Figure 6 shows an example how to define a proximity service as a class (concept) with the use of DAML+OIL.

\[
\text{Figure 6. Axiom in Ontology: definition for a proximity service that uses SQL selection operation for point of interest and is also a database application.}
\]

Since the research done here is to develop a solution that fulfills the requirements of chaining with regard to semantic interoperability, then building a static chaining application (which is relatively straightforward with software development) is out of the question here. The next section will focus on the results of metadata development for the semantic chaining purpose.

### 3. CHAINING DISTRIBUTED GIS SERVICES

Based on the structure shown in table 2, the metadata for each service is built as RDF-XML document. RDF-XML refers to the use of RDF (Resources Description Framework) model through XML serialization. Further, the resources of elements (e.g. the input and the output) are enhanced with the use of ontology (Lemmens, et al, 2003).

### 3.1. The scenario of finding the nearest restaurant: chaining proximity and geocoding

In this scenario, the user-defined chaining will be used. Metadata for each service is created using the structure in table 2. Then the providers publish the metadata to the broker (in this case SDI’s clearinghouse). In such, subsequently the users will have possibilities to examine the chaining since the input and the output are referenced to the data ontology.

For clarity, figure 7 will show the elements definition of input for the geocoding service.

```
<rdf:parseType="Resource">
  <input_value>Place</input_value>
  <input_description>the input of the service can be the street address, name of city and state, or the value of ZIP code</input_description>
</rdf:parseType>
```

**Figure 7. The input definition in geocoding.**

The definition for the “Name of Place” that is referred by the `<input_value>` is shown in figure 8.

```
<owl:Restriction>
  <onProperty rdf:resource="http://www.itc.nl/ontology/geodata#NameOfPlace"/>
  <minCardinality>"1"</minCardinality>
</owl:Restriction>
```

**Figure 8. The ontology definition which is used for the “Name of Place”.**

The same way of input and output definitions is implemented to the other services. Then the possible step is given below (regarding to the scenario of finding nearest restaurant within SDI framework):

- The traveler visits the clearinghouse and types keywords to search for a service that provides functions to determine the nearest point of interest. Next, s/he searches the geocoding service that accepts the output of preceding service as the input.
Figure 9. The result of searching for “geocoding”.

- The clearinghouse will provide the available services and their elements in which it is possible to click on them to determine their compatibility (with the link to ontology definition). (See Figure 10 in appendix).

As an example here, from the result it can be concluded that the StoreLocator can be chained with the PlaceFinder.

3.2. Flood disaster scenario: Chaining coordinate transformation and overlay

The coordinate transformation and overlay are registered into the broker (SDI’s clearinghouse). This is done with using publication function. The metadata of service in the broker side will have the same structure as shown in the table 2.

The linking entity between coordinate transformation and overlay operations is the common coordinate system of output and input respectively. The coordinate system is referenced to a common ontology, e.g. based on the EPSG coding system. The output of the coordinate transformation is described accordingly as can be seen in figure 11.

Figure 11. The output element for a coordinate transformation. #4238 is the EPSG code for the Indonesian datum 1974.

The possible overlay that can be chained with that service must be capable of handling input data that carries the same reference.

Using the same means of interacting with the broker as discussed in 3.1., the chaining for locating the damaged and endangered area then can be done.

4. CONCLUSIONS

This paper presented the principle of building the user-defined chaining in the distributed environments. The metadata of services are residing in the servers where the GIS functionalities are available. These metadata (commonly using WSDL) mainly describe syntactic functionalities of the services. The extensive metadata is needed to provide a clear description about the capability and the functionality of the services for the users. In this paper, the main goal of the metadata extension is aiming at defining the input and the output of services clearly for the chaining purpose. Additionally, an ontology is used to provide a reference for specifying the resource elements. In such, the broker can provide a comprehensive way to examine the possibility to make a chaining. This kind of broker functions can be embedded on the SDI’s clearinghouse. Both scenarios explained, use single services to be combined for solving geographic questions. Those two cases showed the benefits of combining services for commercial and disaster management within the SDI’s framework.

A prototype has been developed according to the idea presented in this paper. However, the chaining with using this prototype is only allowing user-defined chaining through broker assistance with fictional services. This research needs to be extended in providing an intelligent chaining with combining workflow language and ontology language.

REFERENCES


http://www.epsg.org/


Vitae

Trias Aditya was born in Yogyakarta in 1975. He holds an ST degree in Geodesy from UGM and an MSc degree in Geoinformatics from ITC, The Netherlands. He is lecturer at the Dept. of Geodetic Engineering UGM, Yogyakarta. (contact: aditya@itc.nl)

Rob Lemmens holds an MSc degree in Geodesy from Delft University of Technology. He is assistant professor at the Department of Geo-Information Processing at ITC (International Institute for Geo-Information Science and Earth Observation), The Netherlands. (contact: lemmens@itc.nl)
Figure 10. The results of searching and the use of data ontology as the reference for the input and the output elements.