XML and Inter-Operability in Distributed GIS

KIM Do-Hyun and KIM Min-Soo, Republic of Korea

Key words: GML(Geography Markup Language), Interoperability, GIS.

SUMMARY

Web GIS (Geographic Information Systems) service systems provide the various GIS services of analyzing and displaying the spatial data with friendly user-interface.
But, these systems don't support data inter-operability. It is difficult to access diverse data sources because of various spatial data format and data access methods.
In this paper, we design and implement data provider components based on the inter-operability of OGC(OpenGIS Consortium) in web distributed environment. Data provider components provide unique accessing method to distributed data sources based on OLE-DB technology of Microsoft. In addition, each data source process the analyzing the spatial data of its domain.

XML and Inter-Operability in Distributed GIS

KIM Do-Hyun and KIM Min-Soo, Republic of Korea

1. INTRODUCTION

Web GIS(Geographic Information Systems) is expanding the class of user with the advantage of the friendly user interface environment and various geographical operations, topological analysis. But, because of closing-style system architecture these systems don't support the interoperability. That is, there is a lack of interoperability between them because most of them have their own unique data format according to their application fields. This brings about the duplication of data construction. The practical use of Constructed already GIS data drop being been subordinate in GIS package and caused by the GIS data format and the type of DBMS.

OGC(Open GIS Consortium) proposes the open service architecture of web GIS to support data-interoperability. It suggests the GML (Geographic Markup Language) based on XML(eXtended Markup Language) to exchange the data between the web client and the web GIS [2].

And, OLE DB technology of Microsoft provides unique accessing method to distributed data sources. OLE DB uses the Component Object Model (COM) infrastructure, which reduces unnecessary duplication of services and provides a higher degree of interoperability, not only among diverse information sources [1][3].

In this paper, we describe the GML and OLE DB technologies for interoperability in GIS domain. And we present our developed components based these technologies. The rest of the paper is organized as follows. Section 2 discusses two technologies for interoperability. Section 3 describes system architectures of our web GIS environment, Finally, Section 4 consider the advanced technologies for interoperability in GIS domain.

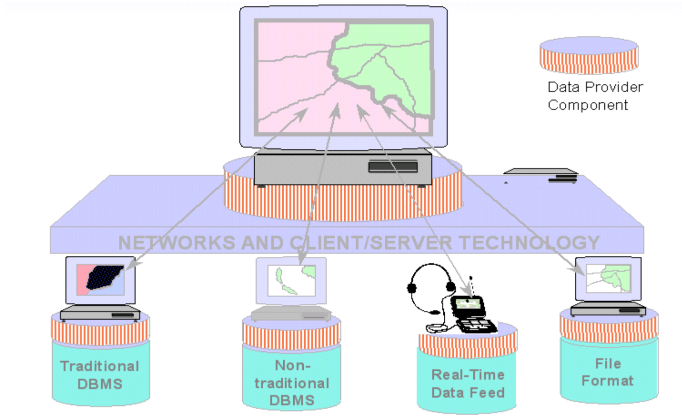
2. RELATED WORK

This paper is described using two opened studies. One is OLE DB and the other is GML based on XML .

OLE DB is set of Component Object Model (COM) interfaces that provide applications with uniform access to data stored in diverse information sources and that also provide the ability to implement additional database services [3].

There are three fundamental categorizations of software – Data Providers, Service Providers, Consumers.

The Data Provider category is the most fundamental set of components that must be implemented in order to allow geographic data to be shared among different applications. These applications may be as diverse as data collection, analysis or simple viewing. With OLE DB interfaces to relevant geographic data, customers and other software vendors will be able to view and analyze heterogeneous collections of data from a wide range of data sources without first trying to convert them all to a compatible data format[1]. Service Providers include spatial query processors, buffer zone services, geocoding services, or network analysis services. And, Consumer is application or tools. [Picture 1] shows the concept of data provider component.



Picture 1 Data Provider Component

OGC announced OLE DB provider implementation specification in order to support interfaces that could be freely accessible to GIS data on distributed computing environment. This implementation specification says that we firstly must implement the Microsoft OLE DB interfaces [3] and then we must extend these interfaces to OGC requirements usually composed of spatial engine functions.

- OGIS Data Provider Registry Entry: OGIS data providers must register support for the "OGISDataProvider" component category so that consumers can distinguish them from other OLE DB provider.
- GIS Metadata Supporting: The manner on retrieving some GIS metadata must be provided. Table 1 shows metadata information that we have to support in an OGC OLE DB provider [6].

GUID	Meta Info.	Description
DBSCHEMA_COLUMNS	OLE DB Info.	Column Information of Table
DBSCHEMA_PROVIDER_TYPES	OLE DB Info.	Data Types of Data Provider
DBSCHEMA_TABLES	OLE DB Info.	All Table Information of Data Provider
DBSCHEMA_OGIS_FEATURE_TABLES	OLE DB Info.	GIS Feature Table Information of Data Provider
DBSCHEMA_OGIS_GEOMETRY_COLUMNS	OLE DB Info.	Geometry Column Information of GIS Feature table
DBSCHEMA_OGIS_SPATIALREF_SYSTEMS	OLE DB Info.	Spatial Reference Information of GIS Feature Table
DBSCHEMA_OGIS_SPATIAL_OPS	OLE DB Info.	Spatial Operator Property Set of Data Provider

Table 1 Metadata of OGIS Data Provider

- IColumnRowset Interface Updates: IColumnRowset interface is available on accessing a meta information for column data. OGC requires that OGC provider must give data consumer data information for geometry and spatial reference column data. Therefore, there is some need to modify the IcolumnRowset interfaces to satisfy the OGC's requirement.
- Geometry Information Supporting: OGC requires that the methods are described for acquiring geometry from columns as a WKB(Well-Known Binary) type.
- Spatial Reference Information Supporting: OGC requires that the methods are described for acquiring spatial reference information as a WKT (Well-Known Text) type.
- Spatial Filter Supporting: OGC requires that spatial filtering method have to work through parameterized manners. These spatial filter parameters must be defined as Spatial Filter, Spatial Operator and Geometry Column Name.

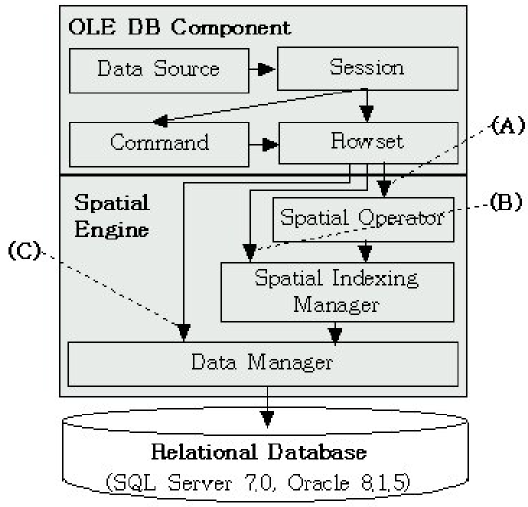
GML is an XML encoding for the transport and storage of geographic information, including both the spatial and non-spatial properties of geographic features [2]. GML uses the W3C XML Schema Definition Language to define and constrain the contents of its XML documents. The GML v2.0 Specification defines some basic conformance requirements for users to develop their own application schemas. Software applications attempting to process any arbitrary GML user application schema must understand GML and all of the technologies upon which GML depends, including the W3C XML Schema [2][4].
This specification defines the XML Schema syntax, mechanism, and conventions that provide an open, vendor-neutral framework for the definition of geospatial application schemas and object. And, it allows profiles that support proper subsets of GML Framework descriptive capabilities. [Picture 2] shows the example of GML feature schema.

```
<my:Road
        xmlns:my="http://www.ned.dem.csiro.au/XMML"   xmlns:gml="http://www.opengis.net/gml"
        xmlns:xlink="http://www.w3.org/1999/xlink"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
        xsi:schemaLocation="http://www.ned.dem.csiro.au/XMML ./road.xsd"
        gml:id="H23">
 <gml:description>An example of a simple Road feature</gml:description>
 <gml:name>Oliver Highway</gml:name>
 <my:nLanes>8</my:nLanes>
 <my:number>M99</my:number>
 <my:surfaceTreatment>gravel</my:surfaceTreatment>
 <my:centreLine>
  <gml:LineString srsName="epsg:4256">
   <gml:coordinates>23.0,45.9 23.1,46.6 23.5,48.2</gml:coordinates>
  </gml:LineString>
 </my:centreLine>
 <my:destination>Leederville, W.A.</my:destination>
 <my:destination xlink:href="urn:au:gov:geoscience:places:Canberra"/>
</my:Road>
```

Picture 2 GML Feature Schema : Road

3. THE IMPLEMENTATION OF OGC OLE DB COMPONENT

[Picture 3] shows the overall system configuration of OGC OLE DB component software using only RDBMS.



Picture 3 System Configuration of
OGC OLE DB Component

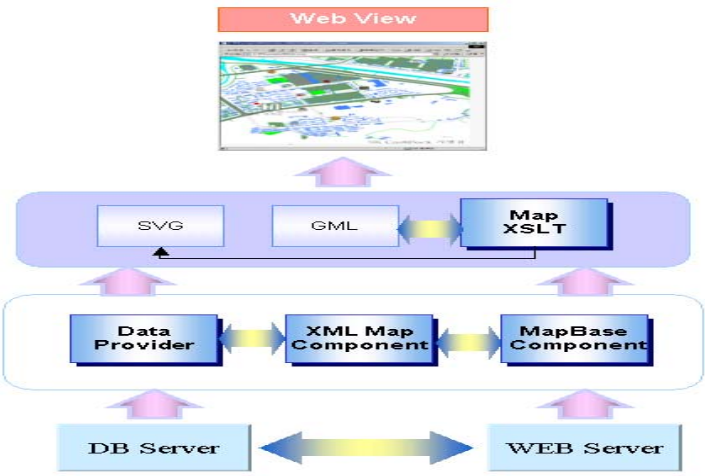
Firstly, OLE DB component specification announced by Microsoft[3] exposes the standard interfaces for accessing to global data server, so public users are guaranteed to access a variety of relational databases or file systems in the same manner through the standard interfaces. In addition to this standard OLE DB component specification, OGC's OLE DB component requires modified and extended interfaces for satisfying OGC's GIS concerned functionality. Extended OLE DB component specification leaves OLE DB component interfaces as it is, additionally must provide user with GIS concerned functions. For this GIS concerned functions, we additionally implemented a spatial engine in this study.
As shown the [Picture 3], OLE DB component consists of data source, session, command and rowset objects. Data source object initializes database and sets up the connection to database. Session object obtains the database concerned meta information and creates a command object for running the SQL statements. Command object executes SQL statements, processes SQL parameters and creates rowset objects as a result of SQL statements. Rowset object plays a role of management of result features that transferred from database as a result of command object. And let's look at the detail process and roles of RDBMS, spatial engine and OLE DB component.

- OGC requires that data provider component should provide various meta information about what GIS features exist in the GIS server, what is geometry column's name of feature table, what spatial reference systems are supported for each feature and what is each feature table's column schema. As these meta information could be obtained through *IDBSchemaRowset* interface of session object, session object of pure OLE DB component have to be extended so as to accommodate the meta information. In [Picture 3], meta information could be easily obtained by calling**(C)** data manager of spatial engine.

- OGC requires that data provider component should provide 9 spatial operators using spatial filter. This spatial filter function demands three SQL parameters of ‘spatial filter object’, ‘spatial operator’ and ‘feature table’s geometry column name’. Actually *ICommandWithParameters* interface of command object processes these three parameters and asks spatial indexing manager and spatial operator of spatial engine for processing these spatial filtering functions. Therefore, such spatial indexing manager for processing spatial filter is indispensable to spatial engine and R*-tree indexing methodology is implemented in this study. In this [Picture 3], this functions are processed through the flow **(A)** or **(B)**.
- OGC also requires that WKB(Well-Known Binary) typed geometry information and WKT(Well-Known Text) typed spatial reference system information are provided by data provider component. Data manager of spatial engine also has to process these functions through the flow **(C)**.

4. THE IMPLEMENTATION OF DATA CONVERSION COMPONENT

[Picture 4] describes geometry conversion components of our web GIS. There are four layers which process the spatial data – data source layer, data provider layer, data conversion layer, and data display layer .

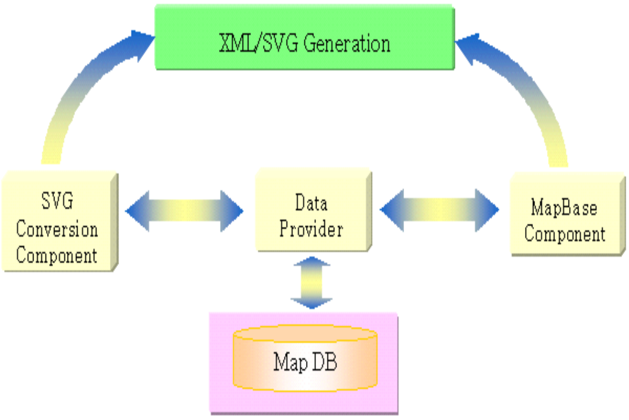


Picture 4 Geometry Conversion Components

Data source layer consists of DB Server and WEB Server. DB Server is composed of the distributed data sources have the spatial data and non-spatial data. That is, our web GIS was developed based on the distributed environment. And Web Server provides various these data of DB Server via Internet.

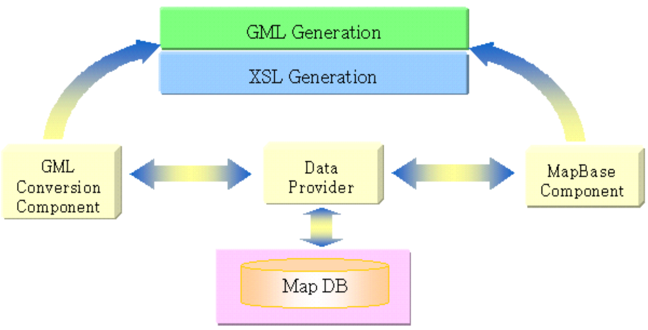
Data provider layer is the set of the components implemented based on specifications of Microsoft’s OLE DB. Data provider components have metadata about the spatial data. These metadata are what are feature tables, what is column of spatial in feature table, what is the

spatial reference system of feature table , which operation is supported by data source layer, and what is extents of feature table.
Data conversion layer have two components. One is the conversion between spatial data of data source and GML. The other is between spatial data of data source and SVG.
SVG is graphic standard of displaying XML announced by W3C [5]. The role of SVG is displaying the GML and graphic contents based on XML in the web application. This SVG component converts the spatial data from data providers to SVG format. [Picture 5] presents SVG conversion component.



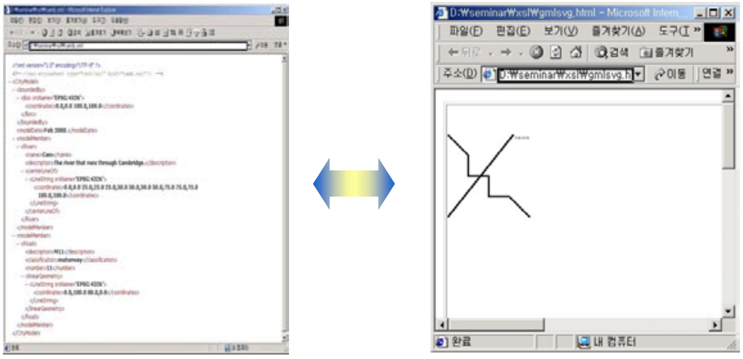
Picture 5 SVG Conversion Component

GML component convert the spatial data from data providers to GML format . But, GML doesn't have the styling information. GML Conversion Component receives WKB data form Data Provider and transfers GML document based on OGC's GML Schema[2]. [Picture 6] presents GML conversion component.



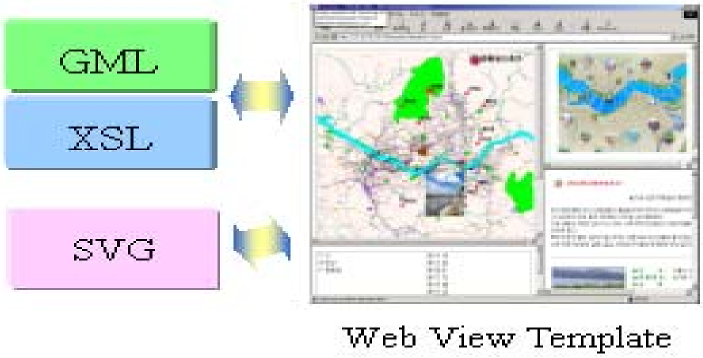
Picture 6 GML Conversion Component

GML data is converted SVG datat with styling information. [Picture 7] is the displaying the GML data with styling information. we define styling information as specific characters of thematic map.



Picture 7 GML-SVG XSLT

Data display layer is web view template component displaying GML and SVG contents . This template has the fundamental mapping functions and SVG services. In addition, it is possible to convert automatically GML to SVG without XSLT. [Picture 8] presents web view template component.



Picture 8 Web View Template Component

5. CONCLUSION

We designed and developed the distributed web GIS services based on the GML and interoperability.
OLE DB technology provides system architecture model related to interoperability. This model provider unique interfaces accessing the spatial and non-spatial data in distributed data sources. Especially, clients can access and retrieve GIS data source support the different GIS data format using data provider component based on OLE-DB technology.
OGC publishes the GML based on XML as vector data format transmitted from server to client. We developed the conversion components between web data contents. This component support the description of spatial application schemas for specialized domains and information communities And increase the ability of organizations to share geographic application schema and the information they describe.
We modeled multiple components using the UML. The components were implemented with ATL/COM provides the environment of the language-independency. These components can

be appropriately composed to the web GIS application, and they have advantages of the reusability and interoperability. The reusability guarantees the low cost of the software development and the interoperability shares the data between diverse GIS data server.

REFERENCES

[1] OpenGIS Consortium Inc, OpenGIS Simple Features Specification For OLE/COM, 1999

[2] Simon Cox, Adrian Cuthbert, Ron Lake, Richard Martell. Eds, Geography Markup Language (GML) 2.0, Open GIS Consortium Inc, 2001

[3] Microsoft Press, Microsoft OLE DB 2.0 Programmer's Reference and Data Access SDK, 1998

[4] Henry S. Thompson, David Beech, Murray Maloney, and Noah Mendelsohn. Eds, XML Schema Part 1: Structures. W3C Recommendation, 2001

[5] W3C Candidate Recommendation, Scalable Vector Graphics (SVG) 1.0 Specification, 2000

[6] Kin, T.J., Metadata for Geo-spatial Data Sharing: A Comparative Analysis, Open GIS Research Reviews, 1999

CONTACTS

XML and Inter-Operability in Distributed GIS
Do-Hyun Kim, Min-Soo Kim
4S Application Research Team, Electronics and Telecommunications Research Institute(ETRI)
161 Gajeong-Dong, Yuseong-Gu
Daejon
KOREA
Tel. + 82 042 860 5529
Fax + 82 042 860 4844
E-mail: dohyun@etri.re.kr, minsoo@etri.re.kr
Web site: www.sitc.re.kr