Contribution ID : 131

Combining Logical Information Systems and OpenGIS tools for geographical data exploration

Friday 15 Sep 2006 at 09:30 (00h30')

The amount of geographical data available around the world is ever increasing. This makes information retrieval a significant stake in geomatics. Traditionnaly, geographical data are organised as layers of information. However, the layer model is a rigid structure for navigation, as it does not permit to reorganize information beyond predefined thematics. Spatially extended SQL languages offer expressive querying capabilities that could balance layer navigation. However, building relevant queries often require external knowledge on the data. Recently, Logical Information Systems (LIS) introduced a new paradigm for information retrieval. In LIS, information is described by logical descriptors which are also used as arguments in data navigation and querying. These systems do not rely on a hierarchical data organisation and enable to tighly combine flexible navigation and expressive querying in the information search process. In this paper, we present GEOLIS, a prototype combining LIS and webmapping capabilities for geographical data exploration. The storage component of GEOLIS is LISFS, a generic implemention of LIS that takes the form of a genuine Linux file system. An important feature of LISFS is that it permits navigation and querying among a set of files as usual and inside files. The second service, called Part-of-files File System (PofFS) is used in GEOLIS. Geographical data are accessed through a graphical web interface designed for data navigation. This interface includes MapServer technology for map vizualisation and PHP based components for logical navigation. The Geographical Markup language (GML) and the OGR2OGR library offers a flexible data format, which is appropriate for interactions between MapServer and LISFS.

Data handling with LIS provides new functionnalities in data exploration. Indeed, logical descriptors used to describe data in LIS also play the role of navigation links and form the basis of the querying language. This enables to freely combine querying and navigation operations in the search process. For instance, a subset of data selected by a query could then be reduced using navigation links, which will refine the query. The resulting subset can be used as a new start point for querying

or navigation. Furthermore, LIS offer an assisted navigation. At each step of the search process, the system suggests the user navigation links, and guaranties that each proposed link reduces the current substet of data but does not lead to an empty set. GEOLIS combines geo visualization capabilities with LIS functionnalities in order to connect map representation and navigation with logical exploration of geographical data.

GEOLIS works with vectorial geographical data. It does not require to convert geographical data in a particular format. An existing geographical format can be used as soon as LISFS has been equiped with the corresponding transducer. A transducer is a plugin that extracts properties from parts of the geographical data. This mechanism makes GEOLIS open on several data sources. At the moment, GEOLIS prototype supports GML. GML is supposed to become a standard for geographical data sharing and has the avantage to gather all information in one file. As it is a subset of XML, writting a GML transducer based on XML-schema and XSLT is quite a simple task. Futhermore, each view in the PofFS is a syntaxically correct GML file, which can be accessed by MapServer (or other applications) through connectors like OGR2OGR. This is made in a transparent way, and views appear as traditionnal Geographical Information Systems (GIS) layers. As views are handled at the file system level, navigating in geographical data corresponds to browsing directories in LISFS. This is done in GEOLIS using a PHP based interface. GEOLIS has been tested on a real dataset about rodents distribution in the Sahelo Soudanian Africa. This database represents a large dataset : about 20,000 features with a mean of 39 attributes per feature. It is also composed of heterogeneous data : several attributes are not filled for many rodents; furthermore, half of the base has been collected in Senegal whereas the studied area goes far East to Chad.. Experiments have enabled to detect quickly anomalous entries in the dataset, e.g. individuals with incoherent values (numerical values to qualify the genus of some rodents) or uncertainty on values (values including or equals to '?'). These entries appear isolated in the logical navigation, which is useful for database cleansing.

To conclude, our prototype shows that LISFS cohabits well with existing mapping technologies. It provides directly, i.e. without implementing specific drivers, new services to mapping applications. First of all, it offers a new navigation paradigm that can be connected, or not, with traditionnal map navigation. Furthermore, it is appropriate for managing heterogeneous data and cleansing them. In the future, we plan to work on spatial logics to improve expressiveness and querying capabilities of GEOLIS with predicates including distance and topological relations. We also have in mind to connect LISFS with other GIS tools to take advantage of LIS view management, for inserting and updating features.

Primary authors : Mr. BEDEL, Olivier (IRISA Rennes / Université de Rennes 1)

Co-authors : Mr. RIDOUX, Olivier (IRISA Rennes / Université de Rennes 1) ; Mr. QUESSEVEUR, Erwan (Laboratoire Reso / Université de Rennes 2)

Presenter : Mr. BEDEL, Olivier (IRISA Rennes / Université de Rennes 1)

Session classification : Session 14 Development

Track classification : --not yet classified--

Type : Conference