

Chapter I

GML as Database: Present and Future

Jose E. Córcoles

Castilla La-Mancha University, Spain

Pascual González

Castilla La-Mancha University, Spain

ABSTRACT

An interesting feature of GML is to consider it as a database, but only in the strictest sense of the term. That is, as a collection of data. As a database format, it can be queried. In order to do this, we need a query language with spatial operators. In addition, in order to use any query language over GML, it is necessary to find an implementation that allows to exploit all its features, i.e., an efficient storage of GML documents is necessary. The general aim of this chapter is to discuss different approaches for storing and querying GML documents. In order to achieve our aim we discuss well-known approaches to the storage of XML documents (with only alphanumeric data) and their application to GML documents. Although there are many approaches to storing and retrieving XML documents with only alphanumeric features, few approaches are applicable to query GML documents.

INTRODUCTION

An interesting feature of eXtensible Markup Language (**XML**) (W3C, 2005) is to consider it as a database, but only in the strictest sense of the term. That is, as a collection of data. In many ways, this makes it no different from any other file.

As a *database* format, XML has several advantages. For example, it is self-describing (the markup describes the structure and type names of the data, although not the semantics), it is portable (Unicode), and it can describe data in tree or graph structures.

This point of view makes it possible to open a new set of XML applications, all of them

involving storage and retrieval of information represented by XML.

One example of XML as a database is a restaurant catalog. It could be defined with alphanumeric features, for example name, phone number, address, and capacity. An advantage of XML is that the data is portable, and it can easily be manipulated for inserting, updating, and deleting information.

Another example could be an extension of the previous example. Besides representing alphanumeric information, the authors can include spatial features. Thus, they include a polygon (spatial coordinates) to represent the parcel where the restaurant is located. In this case, our set of data with alphanumeric features (name, number, etc.) and spatial feature (parcel) would be a Geography Markup Language (**GML**) document (Open Geospatial Consortium, 2003), instead of an XML document.

Since XML (GML by extension) is a database, it can be queried. In order to do this, we need a query language (of general use) to retrieve information from an XML document. Nevertheless, it is necessary to enrich the query language over XML features with spatial operators if we wish to apply it over spatial data encoded with GML. Otherwise, the query language could only be used to query alphanumeric features of an XML document and not, for example, the topological relationship between two spatial regions.

Today, there is a large set of **query languages** over XML. These query languages are different with respect to syntax, available operators and environment of applicability. However, they share the same features, that is, features of query languages over **semi-structured data**. This is because XML is not structured data, but instead has a structure that is flexible (Abiteboul et al., 1997).

More specifically, in order to use any query language over GML (Córcoles and González, 2001), it is necessary to find an implementation that allows to exploit all its features, that is, an efficient storage of GML documents is necessary.

Although there are many approaches to storing and retrieving XML documents with only alphanumeric features (McHugh et al., 1997; Yoshikawa and Amagasa, 2001; Bohannon et al., 2002), few approaches are applicable to query GML documents (Córcoles and González, 2002; Huang et al., 2006).

BACKGROUND

The general aim of this chapter is to discuss different approaches for storing and querying GML documents. In order to achieve our aim we discuss well-known approaches to the storage of XML documents (with only alphanumeric data) and their application to GML documents.

In the following sections we show that this is not a trivial problem because, due to the resources required to query and store spatial elements, appropriate XML-based approaches with alphanumeric operators do not obtain good results when combined with spatial operators. Furthermore, some XML-based approaches are not applicable to GML documents.

Many approaches to storing and retrieving XML documents have been implemented to date, and several database management systems for storing XML documents have been developed (e.g., McHugh et al., 1997). There are also several approaches based on the relational model or object-oriented model. When XML documents are stored in off-the-shelf database management systems, the problem of storage model design for storing XML data becomes a database schema design problem. Yoshikawa and Amagasa (2001) divide such database schemas into two approaches: structure-mapping and model-mapping. In the former, the design of the database schema is based on the understanding of Document Type Descriptor (DTD) or XML Schema that describes the structure of XML documents (e.g., Bohannon et al., 2002; Kappel et al., 2000; Lee and Chu 2000; Huang et al., 2006; Klettke and Meyer, 2000;

8 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/gml-database-present-future/20380

Related Content

Urban Sprawl and the Quantification of Spatial Dispersion

Federico Martellozzo and Keith C. Clarke (2013). *Geographic Information Analysis for Sustainable Development and Economic Planning: New Technologies* (pp. 129-142).

www.irma-international.org/chapter/urban-sprawl-quantification-spatial-dispersion/69053

Effects of Human Culture Among AEC Professionals Towards Adaptation of Collaborative Technology in Industrialized Project Delivery

Maszura Abdul Ghafar and Rahinah Ibrahim (2020). *International Journal of Digital Innovation in the Built Environment* (pp. 36-48).

www.irma-international.org/article/effects-of-human-culture-among-aec-professionals-towards-adaptation-of-collaborative-technology-in-industrialized-project-delivery/255178

Framework for GeoSpatial Query Processing by Integrating Cassandra With Hadoop

S. Vasavi, Mallela Padma Priya and Anu A. Gokhale (2019). *Geospatial Intelligence: Concepts, Methodologies, Tools, and Applications* (pp. 353-388).

www.irma-international.org/chapter/framework-for-geospatial-query-processing-by-integrating-cassandra-with-hadoop/222907

Using Semantic Search and Knowledge Reasoning to Improve the Discovery of Earth Science Records: An Example with the ESIP Semantic Testbed

Kai Liu, Chaowei Yang, Wenwen Li, Zhipeng Gui, Chen Xu and Jizhe Xia (2014). *International Journal of Applied Geospatial Research* (pp. 44-58).

www.irma-international.org/article/using-semantic-search-and-knowledge-reasoning-to-improve-the-discovery-of-earth-science-records/111100

BIM Macro Adoption Study: Establishing Ireland's BIM Maturity and Managing Complex Change

Barry McAuley, Alan V. Hore and Roger P. West (2018). *International Journal of 3-D Information Modeling* (pp. 1-14).

www.irma-international.org/article/bim-macro-adoption-study/216885