Chapter 3.26
A Survey of Spatio–Temporal Data Warehousing

Leticia Gómez
Instituto Tecnológico de Buenos Aires, Argentina

Bart Kuijpers
Hasselt University and Transnational University of Limburg, Belgium

Bart Moelans
Hasselt University and Transnational University of Limburg, Belgium

Alejandro Vaisman
Hasselt University and Transnational University of Limburg, Belgium

ABSTRACT

Geographic Information Systems (GIS) have been extensively used in various application domains, ranging from economical, ecological and demographic analysis, to city and route planning. Nowadays, organizations need sophisticated GIS-based Decision Support System (DSS) to analyze their data with respect to geographic information, represented not only as attribute data, but also in maps. Thus, vendors are increasingly integrating their products, leading to the concept of SOLAP (Spatial OLAP). Also, in the last years, and motivated by the explosive growth in the use of PDA devices, the field of moving object data has been receiving attention from the GIS community. However, not much has been done in providing moving object databases with OLAP functionality. In the first part of this article we survey the SOLAP literature. We then move to Spatio-Temporal OLAP, in particular addressing the problem of trajectory analysis. We finally provide an in-depth comparative analysis between two proposals introduced in the context of the GeoPKDD EU project: the Hermes-MDC system, and Piet, a proposal for SOLAP and moving objects, developed at the University of Buenos Aires, Argentina.

INTRODUCTION

Geographic Information Systems (GIS) have been extensively used in various application domains, ranging from economical, ecological and demographic analysis, to city and route plan-

ning (Rigaux, Scholl, & Voisard, 2001; Worboys, 1995). Spatial information in a GIS is typically stored in different so-called *thematic layers* (also called *themes*). Information in themes can be stored in data structures according to different data models, the most usual ones being the *raster model* and the *vector model*. In a thematic layer, spatial data is annotated with classical relational attribute information, of (in general) numeric or string type. While spatial data is stored in data structures suitable for these kinds of data, associated attributes are usually stored in conventional relational databases. Spatial data in the different thematic layers of a GIS system can be mapped univocally to each other using a common frame of reference, like a coordinate system. These layers can be overlapped or overlayed to obtain an integrated spatial view.

On the other hand, OLAP (On Line Analytical Processing) (Kimball, 1996; Kimball & Ross, 2002) comprises a set of tools and algorithms that allow efficiently querying multidimensional databases, containing large amounts of data, usually called Data Warehouses. In OLAP, data is organized as a set of *dimensions* and *fact tables*. In the multidimensional model, data can be perceived as a *data cube*, where each cell contains a measure or set of (probably aggregated) measures of interest. As we discuss later, OLAP dimensions are further organized in hierarchies that favor the data aggregation process (Cabibbo & Torlone, 1997). Several techniques and algorithms have been developed for query processing, most of them involving some kind of aggregate precomputation (Harinarayan, Rajaraman, & Ullman, 1996).

The Need for OLAP in GIS

Different data models have been proposed for representing objects in a GIS. ESRI (http://www.esri.com) first introduced the *Coverage* data model to bind geometric objects to non-spatial attributes that describe them. Later, they extended this model with object-oriented support, in a way that behavior can be defined for geographic features (Zeiler, 1999). The idea of the Coverage data model is also supported by the Reference Model proposed by the Open Geospatial Consortium (http://www.opengeospatial.org). Thus, in spite of the model of choice, there is always the underlying idea of binding geometric objects to objects or attributes stored in (mostly) object-relational databases (Stonebraker & Moore, 1996). In addition, query tools in commercial GIS allow users to overlap several thematic layers in order to locate objects of interest within an area, like schools or fire stations. For this, they use indexing structures based on R-trees (Gutman, 1984). GIS query support sometimes includes aggregation of geographic measures, for example, distances or areas (e.g., representing different geological zones). However, these aggregations are not the only ones that are required, as we discuss below.

Nowadays, organizations need sophisticated GIS-based Decision Support System (DSS) to analyze their data with respect to geographic information, represented not only as attribute data, but also in maps, probably in different thematic layers. In this sense, OLAP and GIS vendors are increasingly integrating their products (see, for instance, Microstrategy and MapInfo integration in http://www.microstragey.com/, and http://www.mapinfo.com/). In this sense, aggregate queries are central to DSSs. Classical aggregate OLAP queries (like "total sales of cars in California"), and aggregation combined with complex queries involving geometric components ("total sales in all villages crossed by the Mississippi river and within a radius of 100 km around New Orleans") must be efficiently supported. Moreover, navigation of the results using typical OLAP operations like roll-up or drill-down is also required. These operations are not supported by commercial GIS in a straightforward way. One of the reasons is that the GIS data models discussed above were developed with "transactional" queries in mind. Thus, the databases storing non-spatial attributes or objects are designed to support those (non-aggregate)

Related Content

Knowledge Business Examples

Petter Gottschalk (2007). *Business Dynamics in Information Technology (pp. 230-351).*

www.irma-international.org/chapter/knowledge-business-examples/6060

State of the Art Solutions in Enterprise Interoperability

Silke Balzert, Thomas Burkhart, Dirk Werth, Michal Laclavík, Martin Šeleng, Nikolay Mehandjiev, Martin Carpenterand Iain Duncan Stalker (2010). *Business Information Systems: Concepts, Methodologies, Tools and Applications (pp. 2134-2162).*

www.irma-international.org/chapter/state-art-solutions-enterprise-interoperability/44189

Ontology-Driven Method Engineering for Information Systems Development

Roland Holten, Alexander Dreilingand Jörg Becker (2005). *Business Systems Analysis with Ontologies (pp. 174-217).*

www.irma-international.org/chapter/ontology-driven-method-engineering-information/6123

Evaluating E-Learning from an End User Perspective

Robert Costello (2014). *Frameworks of IT Prosumption for Business Development (pp. 259-283).*

www.irma-international.org/chapter/evaluating-e-learning-from-an-end-user-perspective/78781

Highway Alignment Optimization Using Cost-Benefit Analysis Under User Equilibrium

Avijit Majiand Manoj K. Jha (2013). *Optimizing, Innovating, and Capitalizing on Information Systems for Operations (pp. 313-327).*

www.irma-international.org/chapter/highway-alignment-optimization-using-cost/74024