Fundamenta Informaticae 132 (2014) i-iii
DOI 10.3233/FI-2014-1038
IOS Press

Warehousing and OLAPing Complex, Spatial and Spatio-Temporal Data

Preface

Complex, spatial and spatio-temporal data arise in a plethora of modern database and data mining applications and complex information systems. Complex, spatial and spatio-temporal data require more and more for effective and efficient models, algorithms and techniques for representing, managing, querying, indexing and discovering useful knowledge beyond such kind of data. A successful solution to issues above consists in applying well-consolidated methodologies coming from the *Data Warehousing and OLAP* research area. This allows us to take advantages from several nice amenities supported by Data Warehousing and OLAP, such as multidimensional and multi-resolution representation and analysis, multidimensional aggregations, hierarchy-based data representation and mining, complex query answering tools, and so forth. Application fields where Data Warehousing and OLAP over complex, spatial and spatio-temporal data have already demonstrated their success are many-fold.

As orthogonal to these emerging research contexts, a number of research challenges are capturing the attention of a large community of researchers, ranging from *models, algorithms and techniques for warehousing and OLAPing non-conventional data* to *ETL and data integration approaches* for supporting this task, from *storage issues* to *privacy-preserving issues of warehousing and OLAPing complex, spatial and spatio-temporal data*; from *models, algorithms and techniques for querying complex, spatial and spatio-temporal data in warehousing and OLAP environments* to *models, algorithms and techniques for mining complex, spatial and spatio-temporal data in warehousing and OLAP environments*.

In order to fulfill innovative research challenges and system-oriented requirements posed by the issue of warehousing and OLAPing such kind of data, this special issue on "*Warehousing and OLAPing Complex, Spatial and Spatio-Temporal Data*" of *Fundamenta Informaticae* presents a rigorous selection of best papers of the track *Warehousing and OLAPing Complex, Spatial and Spatio-Temporal Data* (WOCD 2010) of the 14^{th} *East-European Conference on Advances in Databases and Information Systems* (ADBIS 2010), held in Novi Sad, Serbia, during September $20 - 24$, 2010. WOCD 2010 has attracted a relevant number of submissions, and, after a rigorous selection two-round review process over the best papers, only 5 papers have been finally accepted for final publication in the special issue on "*Warehousing and OLAPing Complex, Spatial and Spatio-Temporal Data*".

The aim of the special issue is to offer an innovative, modern research perspective on the issue of warehousing and OLAPing complex, spatial and spatio-temporal data, by highlighting recent top-quality contributions and results in this scientific context, and, at the same, stimulating further investigation in the reference field. In the following, we provide a summary of papers contained in the special issue.

The first paper, titled "*Using a Time Granularity Table for Gradual Granular Data Aggregation*", by N. Iftikhar and T.B. Pedersen, starts from recognizing that the majority of today's systems increasingly require *sophisticated data management techniques* as they need to store and to query large amounts of data for analysis and reporting purposes. In order to keep more "detailed" data available for longer periods, "old" data has to be *reduced gradually* in order to save space and improve query performance, especially on resource-constrained systems with limited storage and query processing capabilities. A number of *data reduction solutions* have been developed recently, however an effective solution particularly based on *gradual data reduction* is still missing. Starting from this limitation, the paper presents *an effective solution for data reduction based on the so-called gradual granular data aggregation*. With the gradual granular data aggregation mechanism, older data can be made coarse-grained while keeping the newest data fine-grained. The proposed solution introduces a time granularity based data structure, namely a *relational time granularity table*, which enables (*i*) long term storage of old data by maintaining them at different levels of granularity and (*ii*) effective query processing due to reductions in data volume. In addition to this, the paper describes the implementation strategy derived from a farming case study that makes use of standard database technologies.

The second paper, titled "*Querying Cardinal Directions between Complex Objects in Data Warehouses*", by G. Viswanathan and M. Schneider, focuses on *cardinal directions* that have turned out to be very important *qualitative spatial relations* due to their numerous applications in spatial way-finding, GIS, qualitative spatial reasoning and in domains such as cognitive sciences, AI and robotics. They are frequently used as selection criteria in spatial queries. *Moving objects data warehouses* can help to analyze complex multidimensional data of a spatio-temporal nature and to provide decision support. However, currently there is no available methods to query for cardinal directions between spatio-temporal objects in data warehouses. Inspired by this main consideration, authors introduce the concept of a *moving objects data warehouse* (*MODW*) for storing and querying multidimensional spatio-temporal data. Furthermore, they also present a novel two-phase approach to model and query for cardinal directions between moving objects by using the MODW framework. This approach works as follows. First, a *tiling strategy* that determines the zone belonging to the nine cardinal directions of each spatial object at a particular time and then intersects them is applied. This leads to a *collection of grids* over time called the *Objects Interaction Graticule* (*OIG*). For each grid cell, the information about the spatial objects that intersect it is stored in the so-called *Objects Interaction Matrix (OIM)*. Then, an *interpretation method* is applied to these matrices to determine the cardinal direction between the moving objects. These results are integrated into MDX queries using spacialized *directional predicates*.

The third paper, titled "*Complex Object-Based Multidimensional Modeling and Cube Construction*", by D. Boukraa, O. Boussaid and F. Bentayeb, presents *a multidimensional model and a language for constructing OLAP cubes on so-called complex objects*, which model complex entities of the real world. Similarly to the real-word objects they point-out, even the multidimensional model and language proposed in this research are based on the concept of complex object/entity. In more details, the multidimensional model is presented at two layers: the *class diagram layer* and the *package layer*. Both layers are used by a *projection operation* that aims at extracting cubes: at the package diagram layer, the projection dynamically assigns the roles of fact and dimensions to complex objects of the multidimensional model whereas, at the class diagram layer, measures are designed. Formal operations constructing the final OLAP cubes are provided, along with operations that optimize the construction of new cubes from existing ones. The set of all operations supported by the proposed multidimensional model form a language. Finally, in order to show the feasibility of proposed multidimensional model and operators, authors provide implementation details of a real-world case study.

The fourth paper, titled "*A Decomposition Framework for Computing and Querying Multidimensional OLAP Data Cubes over Probabilistic Relational Data*", by A. Cuzzocrea and D. Gunopulos, introduces *a novel decomposition framework for efficiently computing and querying multidimensional OLAP data cubes over probabilistic data*. These data, which are characterized by *uncertainty* and *imprecision*, arise in novel database application scenarios, such as sensor network data management, logistic networks and telecommunication systems. Several models and algorithms supported in the proposed framework are formally presented and described in details, based on well-understood *theoretical statistical/probabilistic tools*, which converge to the definition of the so-called *probabilistic OLAP data cubes*, a relevant result of this paper. Finally, authors complete their analytical contribution by introducing an innovative *Probability Distribution Function* (PDF)-based approach, which makes use of well-known *probabilistic estimators theory*, for efficiently querying probabilistic OLAP data cubes, along with a comprehensive experimental assessment and analysis over synthetic probabilistic databases.

Finally, the fifth paper, titled "*OLAPing Field Data: A Theoretical and Implementation Framework*", by S. Bimonte, M.-A. Kang, L. Paolino, M. Sebillo, M. Zaamoune and G. Vitiello, focuses the attention on the issue of *integrating spatial data into multidimensional models*, which leads to the concept of so-called *Spatial OLAP* (SOLAP). Usually, SOLAP models exploit discrete spatial data, while few initiatives propose to integrate *continuous field* data into dimensions and measures. Contrary to this main trend, authors provide a formal multidimensional model that supports measures and dimensions as continuous field data, without dependence on their proper implementation. In addition to this, authors also provide a proposal for a logical model allowing aggregation of field measures in a feasible ROLAP architecture.

The Guest Editors would like to express their sincere gratitude to the Editor-In-Chief of *Fundamenta Informaticae*, Prof. Damian Niwiński, for accepting their proposal of a special issue focused on OLAPing complex, spatial and spatio-temporal data, and for assisting them whenever required. The Guest Editors would also like to thank all the reviewers who have worked within a tight schedule and whose detailed and constructive feedbacks to authors have contributed to substantial improvement in the quality of final papers.

The Guest Editors

Alfredo Cuzzocrea
ICAR-CNR &
University of Calabria
e-mail: cuzzocrea@si.deis.unical.it

Zbigniew W. Ras
Department of Computer Science
University of North Carolina
e-mail: `ras@uncc.edu`