Editorial

Jaroslav Pokorný

Faculty of Mathematics and Physics
Charles University, Czech Republic
E-mail: jaroslav.pokorny@mff.cuni.cz

Robert Wrembel*

Institute of Computing Science
Poznań University of Technology, Poland
E-mail: robert.wrembel@cs.put.poznan.pl
*Corresponding author

Biographical notes: Jaroslav Pokorný received his PhD degree in Theoretical Cybernetics in 1984. Currently he is a Professor of Computer Science at the Faculty of Mathematics and Physics at Charles University and at the Czech Technical University in Prague. Pokorný has published more than 250 papers and books on data modelling, relational databases, query languages, file organisation and XML. His research interests also include database design, information retrieval and semantic web. Pokorný is a member of the ACM and IEEE. In 2004 he became the representative of the Czech Republic to the IFIP.

Robert Wrembel works as an Assistant Professor at the Poznań University of Technology, Poland. In 2001 he received a PhD in Computer Science (Databases). In 1996–2006 he took part in four research projects on databases and four industrial projects in the field of information technologies. He has visited a number of research and education centres, including the INRIA Paris-Rocquencourt (France), Paris Dauphine University (France), Klagenfurt University (Austria) and Loyola University (USA). His main research interests encompass data warehouse technologies (temporal, multiversion, object-relational) and object-oriented systems (views, data access optimisation).

1 Introduction

The modern way of managing enterprises, institutions and organisations is based on knowledge, which in turn is gained from data analysis. Business decisions are taken based on the analysis of past and current data, continuously collected during the lifetime of an enterprise. In practice, a data analysis technology, widely accepted by research and industry, is based on a data warehouse architecture. In this architecture, data coming from multiple distributed and heterogeneous storage systems are integrated in a central repository, called a Data Warehouse (DW). Such integrated data are analysed by Online Analytical Processing (OLAP) applications for the purpose of discovering trends, discovering patterns of behaviour and anomalies, finding hidden dependencies between data and predicting business trends.

From a technical point of view, a data warehouse is a huge database (often exceeding the size of dozens of terabytes (TB)). It typically offers dedicated data structures for efficient OLAP processing, advanced OLAP query optimisation methods and analytical extensions to SQL.

The continuous development in data warehousing technology for more than 20 years has focused mainly on warehousing data of simple formats, *i.e.*, texts, dates and numbers. The development has resulted in widely accepted technologies and standards used by commercial products, *i.e.*, schema design methodologies and implementation models; data loading and DW refreshing techniques; efficient query processing with the support of advanced indexes, materialised views and query optimisers; parallel processing; data partitioning techniques; metadata management; advanced prediction and data mining algorithms; advanced data visualisation techniques and user interfaces; DW development tools and OLAP application development tools.

Nowadays, information systems store and process data other than just simple records composed of strings, dates and numbers. Typically, various web systems store huge amounts of data in the form of XML documents, images and other complex data structures. Advanced image-processing technologies make the use of images and maps easier and more common, *e.g.*, Google Maps, NASA Earth Observing System Data and Information System. Various commonly used mobile devices (*e.g.*, PDAs, different types of sensors) challenge current DW technologies to support the analysis of data that are 'mobile' and are being continuously supplied into a system.

These various new types of data are as important for business as traditional text, dates and numbers, and therefore there is a need to analyse them in a similar way to traditional DWs. This requirement has resulted in several attempts to build data warehouses from web data sources, and to provide OLAP functionality for XML documents, multimedia data and object data. Moreover, knowledge acquisition from images requires combining the functionality of Geographical Information Systems with the functionality of data warehouses and OLAP. These facts lead to the conclusion that warehousing data stored in web systems as well as in mobile and wireless environments is of high importance.

The aim of this special issue of the *International Journal of Web Engineering and Technology* is to present problems, challenges and solutions for developing DW and OLAP technologies for analysing complex data that come from web, mobile and wireless environments.

The first paper, an invited one, entitled 'Warehousing complex data from the web' by Boussaïd *et al.*, describes an approach to warehousing and analysing complex data that originate from the web. The authors present a complex data warehousing methodology that exploits XML as a pivot language. Their approach includes the integration of complex data in operational data storage, in the form of XML documents; their dimensional modelling and storing in an XML data warehouse; and their analysis with combined OLAP and data mining techniques.

The second paper, entitled 'Web services oriented architectures for mobile SOLAP applications' by Badard *et al.*, deals with geospatial (web) Services Oriented Architectures (SOA) for mobile processing of geo-decisional information, including the requirement analysis for deploying Spatial Online Analytical Processing (SOLAP) applications on mobile devices using such architectures. The authors also propose different web services based on an extended geospatial SOA in order to tackle several issues related to mobile SOLAP. One of the proposed web services is currently under development and is detailed in the paper.

The third paper, entitled 'A web-based architecture for temporal OLAP' by Vaisman *et al.*, studies temporal OLAP operating over the web. After introducing a temporal data model supporting historical dimensions and fact table versioning, the authors introduce a three-tier architecture based on web services, SOAP and XML, allowing efficient querying over the web. They also present TOLAP-QL, an SQL-like temporal query language and study query optimisation in a temporal OLAP. The applied technique is based on materialisation of historical views.

The fourth paper, entitled 'Integrating XML data in the TARGIT OLAP system' by Pedersen *et al.*, presents the results of industrial work on the logical integration of OLAP and XML data sources in a framework of the TARGIT prototype. It is an OLAP client developed by a Danish vendor. The developed prototype allows XML data stored outside the OLAP system to be used as dimensions and measures in this system. A part of the paper is devoted to a novel multigranular data model and query language that formalises and extends the TARGIT data model and query language.

Finally, the fifth paper, entitled 'Quantity comparison of concurrency control methods for XML database systems based on DOM API' by Jankiewicz *et al.*, concerns concurrency control in accessing and modifying XML documents, which are stored in an XML database system. The aim of this paper is the analysis and quantity comparison of concurrency control methods for XML database systems based on DOM API.

2 Reviewers of this special issue

The editors would like to thank the following researchers for their valuable assistance in the development of the issue.

- Yvan Bédard, Centre for Research in Geomatics and Laval University, Canada
- Mauricio Minuto Espil, Pontificia Universidad Católica, Argentina
- Pedro Furtado, University of Coimbra, Portugal
- Matteo Golfarelli, University of Bologna, Italy
- Hyoil Han, Drexel University, USA
- Mikolaj Morzy, Poznán University of Technology, Poland
- Torben Bach Pedersen, Alborg University, Denmark
- Juan Manuel Perez Martinez, Castellon University, Spain
- David Taniar, Monash University, Australia
- Zahir Tari, RMIT University, Australia
- Alejandro Vaisman, University of Buenos Aires, Argentina.