

An UML Profile and SOLAP Datacubes Multidimensional Schemas Transformation Process for Datacubes Risk-Aware Design

Elodie Edoh-Alove, National Research Institute for Science and Technology for the Environment and Agriculture (IRSTEA), Aubiere, France

Sandro Bimonte, National Research Institute for Science and Technology for the Environment and Agriculture (IRSTEA), Aubiere, France

François Pinet, National Research Institute for Science and Technology for the Environment and Agriculture (IRSTEA), Aubiere, France

ABSTRACT

Spatial Data Warehouses (SDWs) and Spatial On-Line Analytical Processing (SOLAP) systems are new technologies for the integration and the analysis of huge volume of data with spatial reference. Spatial vagueness is often neglected in these types of systems and the data and analysis results are considered reliable. In a previous work, the authors provided a new design method for SOLAP datacubes that allows the handling of vague spatial data analysis issues. The method consists of tailoring SOLAP datacubes schemas to end-users tolerance levels to identified potential risks of misinterpretation they encounter when exploiting datacubes containing vague spatial data. In this paper, the authors further their previous proposal by presenting different formal tools to support their method: it is an UML profile providing stereotypes needed to add vague, risks and tolerance levels information on datacubes schemas plus the formal definition of SOLAP datacubes schemas transformation process and functions.

Keywords: *Spatial Data Warehouse, Spatial OLAP, Spatial OLAP Datacubes Schemas, Spatial Vagueness, UML Profile*

DOI: 10.4018/ijdwm.2015100104

1. INTRODUCTION

In Bédard and Han (2009), Spatial On-Line Analytical Processing (SOLAP) is defined as a “Visual platform built especially to support rapid and easy spatiotemporal analysis and exploration of data following a multidimensional approach comprised of aggregation levels”. In this type of systems, cartographic, tabular and chart displays are used to interactively visualize and explore the data. The data are stored in a Spatial Data Warehouse (SDW) as a datacube that is the multidimensional model implementation (Salehi, Bédard, & Rivest, 2010). The decision-makers who use SOLAP tools are rarely fully aware of the spatial data uncertainty. Spatial data are usually represented by crisp objects in information systems, while, in fact it is often impossible to exactly know the spatial object boundaries. Thus, a clear gap is created between the majority of real world phenomena and their formal representation in spatial data warehouses (Cheng, Molenaar, & Lin, 2001). The SOLAP end-users are then exposed to erroneous analysis due to the uncertainty issues on data sources.

As indicated in Edoh-Alove, Bimonte, and Bédard (2014), only some recent research introduces spatial vagueness in the multidimensional model (Jadidi, Mostafavi, Bédard, & Long, 2012; Siqueira, Ciferri, Times, & Ciferri, 2014) and SOLAP operators to reduce the analysis errors. While Jadidi et al. (2012) have proposed an approach based on Fuzzy Set Theory and Tessellation to address spatial vagueness in SOLAP datacubes, and experimented it in coastal erosion management, Siqueira et al. (2014) have extended the multidimensional model to exploit vague objects exact models (Bejaoui, 2009; Pauly & Schneider, 2010) in SOLAP systems.

In Edoh-Alove et al. (2014), we proposed a design process where a risk-management approach was adopted to better deal with the spatial vagueness issues. This method was motivated by the need of sticking with existing SOLAP technologies which support vector-based

cartographic dimensions implementation. The method, called Risk-Aware Design of SOLAP datacube (RADSOLAP) method, considers the risks of misinterpretation induced by vague data in the SOLAP datacubes and identified by spatial data experts and end-users. It also considers the tolerance levels of end-users to those risks, and based on those parameters, proposes to end-users a set of SOLAP datacube prototypes iteratively tailored to their tolerance levels. This paper goes beyond the work presented in Edoh-Alove et al. (2014) by proposing formal tools to support the key steps of the process. More specifically, we extended an existing UML Profile intended for SOLAP datacube modeling to take into account vague, risk and tolerance elements and we described and formalized the SOLAP datacubes prototypes tailoring to end-users tolerance levels in the forms of multidimensional schemas transformation functions and a transformation process.

In Section 2 of this paper, we briefly describe our RADSOLAP method alongside with the ICSOLAP UML Profile we have extended; in Section 3, an agricultural case study used to illustrate our proposals is presented; then in section 4, we detailed our new RADSOLAP UML profile and in section 5 the SOLAP datacube transformation process and functions; section 6 present the related work and finally, Section 6 concludes the paper.

2. PRELIMINARY WORK

2.1. RADSOLAP Method

The RADSOLAP method is our risk-aware design method advocated in Edoh-Alove et al. (2014). Using this RADSOLAP method, SOLAP experts, with the participation of SOLAP datacube end-users, are able to create SOLAP datacubes PIMs where spatial vagueness on spatial data are considered and managed. The spatial vagueness management is done via the identification, assessment and reduction of risks of misinterpretation caused by the introduction of spatial vague data into the SOLAP datacubes.

18 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/article/an-uml-profile-and-solap-datacubes-multidimensional-schemas-transformation-process-for-datacubes-risk-aware-design/130667

Related Content

Boat Detection in Marina Using Time-Delay Analysis and Deep Learning

Romane Scherrer, Erwan Aulnette, Thomas Quiniou, Joël Kasarherou, Pierre Kolband Nazha Selmaoui-Folcher (2022). *International Journal of Data Warehousing and Mining* (pp. 1-16).

www.irma-international.org/article/boat-detection-in-marina-using-time-delay-analysis-and-deep-learning/298006

Collective Entity Disambiguation Based on Hierarchical Semantic Similarity

Bingjing Jia, Hu Yang, Bin Wuand Ying Xing (2020). *International Journal of Data Warehousing and Mining* (pp. 1-17).

www.irma-international.org/article/collective-entity-disambiguation-based-on-hierarchical-semantic-similarity/247917

Parallel Hybrid BBO Search Method for Twitter Sentiment Analysis of Large Scale Datasets Using MapReduce

Ashish Kumar Tripathi, Kapil Sharmaand Manju Bala (2022). *Research Anthology on Implementing Sentiment Analysis Across Multiple Disciplines* (pp. 1219-1236).

www.irma-international.org/chapter/parallel-hybrid-bbo-search-method-for-twitter-sentiment-analysis-of-large-scale-datasets-using-mapreduce/308541

Fuzzy Miner: Extracting Fuzzy Rules from Numerical Patterns

Nikos Pelekis, Babis Theodoulakis, Ioannis Kopanakisand Yannis Theodoridis (2005). *International Journal of Data Warehousing and Mining* (pp. 57-81).

www.irma-international.org/article/fuzzy-miner-extracting-fuzzy-rules/1748

Improving Healthcare with Data-Driven Track-and-Trace Systems

Eldar Sultanowand Alina M. Chircu (2016). *Big Data: Concepts, Methodologies, Tools, and Applications* (pp. 1229-1246).

www.irma-international.org/chapter/improving-healthcare-with-data-driven-track-and-trace-systems/150213