

A UML profile for multidimensional modeling in data warehouses

Sergio Luján-Mora, Juan Trujillo

Dept. de Lenguajes y Sistemas Informáticos

Universidad de Alicante

Alicante, España

slujan@dlsi.ua.es, jtrujillo@dlsi.ua.es

Il-Yeol Song

Dept. of Information Systems

Drexel University

Philadelphia, USA

songiy@drexel.edu

Resumen

Here, we briefly summarize our paper entitled “A UML profile for multidimensional modeling in data warehouses”, published in the Data and Knowledge Engineering (DKE) journal, vol. 59, December 2006, pp. 725-769. The full paper can be downloaded from the following Digital Object Identifier (DOI): <http://dx.doi.org/10.1016/j.datak.2005.11.004>. The DKE journal is highly cited in the JCR (Journal Citation Report) with an impact factor of 1.085 in 2005 (no data about 2006 is still available).

Moreover, let us point out that previous research works related to this paper were published in ER 2002 (21st International Conference on Conceptual Modeling) and UML 2002 (5th International Conference on the Unified Modeling Language). The acceptance rate of these two conferences were 23.7% and 30% respectively.

1. Summary of the publication

The multidimensional (MD) modeling, which is the foundation of data warehouses (DWs), MD databases, and On-Line Analytical Processing (OLAP) applications, is based on several properties different from those in traditional database modeling. In the past few years, there have been some proposals, provid-

ing their own formal and graphical notations, for representing the main MD properties at the conceptual level. However, unfortunately none of them has been accepted as a standard for conceptual MD modeling.

In this paper, we present an extension of the Unified Modeling Language (UML) using a UML profile. This profile is defined by a set of stereotypes, constraints and tagged values to elegantly represent main MD properties at the conceptual level. We make use of the Object Constraint Language (OCL) to specify the constraints attached to the defined stereotypes, thereby avoiding an arbitrary use of these stereotypes. We have based our proposal in UML for two main reasons: (i) UML is a well known standard modeling language known by most database designers, thereby designers can avoid learning a new notation, and (ii) UML can be easily extended so that it can be tailored for a specific domain with concrete peculiarities such as the multidimensional modeling for data warehouses. Moreover, our proposal is Model Driven Architecture (MDA) compliant and we use the Query View Transformation (QVT) approach for an automatic generation of the implementation in a target platform. Throughout the paper, we will describe how to easily accomplish the MD modeling of DWs at the conceptual level. Finally, we show how to use our extension in Rational Rose for MD modeling.