Extending the Multidimensional Model for Linking Cubes

Alberto Sabaini*, Esteban Zimányi**, Carlo Combi*

*Department of Computer Science,
University of Verona, Italy
{alberto.sabaini, carlo.combi}@univr.it
**Department of Computer and Decision Engineering (CoDE),
Université Libre de Bruxelles, Belgium
ezimanyi@ulb.ac.be

Abstract. Data warehouses structure data in multidimensional cubes, where dimensions specify different ways in which measures in facts may be viewed, aggregated, and sorted. It is essential for data analysts to combine data from heterogeneous multidimensional cubes to enhance their analysis capabilities. For this, users are restricted in using only shared dimensions for navigating related multidimensional cubes. In this paper, we show that this limits the analysis possibilities and introduce an explicit link that relates two multidimensional cubes, indicating that they represent different aspects of the same reality, and hence they may be connected. We argue that the standard drill-across operator is not suited to perform such operation, and we extend it by proposing a new operator called drill-across-link.

1 Introduction

Data warehouses structure data according to a multidimensional space, where dimensions specify different ways the data can be viewed, aggregated, and sorted. Events of interest for an analyst are represented as facts which are associated with cells or points in this multidimensional space and described in terms of a set of measures. Thus, every fact is based on a set of dimensions that determine the granularity adopted for representing fact measures. Dimensions are organized as hierarchies of levels that allow analysts to aggregate data at different degrees of detail. One of the major challenges that must be faced by designers of multidimensional models is to adequately represent the interactions between dimensions and facts (Song et al., 2001). Another issue is to represent the connections between different facts in the same schema. Indeed, data analysts often need to combine data from heterogeneous multidimensional cubes. A key requirement in navigating multidimensional cubes from several sources, according to Kimball and Ross (2013), is that they must have *shared* dimensions. Their widely adopted view requires shared dimensions to be either from the same instance or identical in terms of schema and data. The non-conformity problem arises when there is the need of combining multidimensional cubes and using non-shared dimension. The inclusion of non-shared dimensions in the navigation and visualization of multidimensional data from multiple sources provides the analyst with the ability to view and analyze data that would be otherwise not