

Chapter VI

Different Kinds of Hierarchies in Multidimensional Models

Elzbieta Malinowski

Universidad de Costa Rica, Costa Rica

INTRODUCTION

In the database design, the advantages of using conceptual models for representing users' requirements are well known. Nevertheless, even though data warehouses (DWs) are databases that store historical data for analytical purposes, they are usually represented at the logical level using the star and snowflake schemas. These schemas facilitate delivery of data for online analytical processing (OLAP) systems. In particular, hierarchies are important since traversing them, OLAP tools perform automatic aggregations of data using the roll-up and drill-down operations. The former operation transforms detailed data into

aggregated ones (e.g., daily into monthly sales) while the latter does the opposite.

In spite of the advantages of star and snowflake schemas, there are some inconveniences in using them. For example, since these schemas are based on the relational logical model, some implementation details (e.g., foreign keys) must be considered during the design process. Further, the star and snowflake schemas are not adequate for representing different kinds of hierarchies existing in real-world applications. Therefore, users are not able to express their analysis needs, and consequently, developers cannot implement them.

We advocate that it is necessary to represent DW data requirements at the conceptual level.

The conceptual model should clearly distinguish different kinds of hierarchies since they exist in real-world situations and are important for DW and OLAP applications. Further, developers should be able to implement these hierarchies. Therefore, considering that DWs and OLAP can use relational storage, we present how hierarchies can be mapped to a relational model.

BACKGROUND

The star and snowflake schemas include relational tables known as fact and dimension tables. The fact table represents the focus of analysis (e.g., analysis of sales). It usually contains numeric data called measures (e.g., quantity). Dimension tables contain attributes that allow users to see measures from different perspectives (e.g., analyze sales in different stores). Since users usually start from a general view of data and then, if required, the detail explorations follow, dimensions may contain attributes that form hierarchies. OLAP tools allow users to traverse hierarchies, aggregating measures automatically. For example, “moving” (i.e., using the roll-up operation) from store to city, the quantity of sold products in each store will be added according to the cities where the stores are located.

Depending on whether hierarchies are represented using flat (Figure 1(a)) or normalized tables (Figure 1(b)), the relational structure is called star or snowflake schemas, respectively. Nevertheless, both schemas are not adequate for representing different kinds of hierarchies existing in real-world situations. The star schema does not represent hierarchies clearly, and the hierarchy structure should be deduced based on the knowledge of the application domain. On the other hand, the snowflake schema only allows us to represent simple hierarchies such as Store, City, and State in Figure 1(b), even though there are different kinds of hierarchies in real-world applications.

There are several proposals of conceptual multidimensional models¹ that include hierarchies. Nevertheless, as we will see later, these models do not include all hierarchies as presented in this chapter. This lack of a general classification of hierarchies, including their characteristics at the schema and at the instance levels, leads to repeated research efforts in “rediscovering” hierarchies and providing solutions for managing them.

MAIN FOCUS

We first describe the MultiDim model, a conceptual multidimensional model used for representing requirements for DW and OLAP applications, including different kinds of hierarchies. Then, we present the hierarchy classification and refer in more detail to each hierarchy type. Last, we present mapping of these hierarchies to the relational model.

The MultiDim Model

To describe the MultiDim model (Malinowski & Zimányi, 2008), we use an example of a Sales DW shown in Figure 2 that contains different kinds of hierarchies; we refer to them in the next section.

A MultiDim schema is a finite set of dimensions and fact relationships. A dimension is an abstract concept for grouping data that share a common semantic meaning within the domain being modeled. A dimension is composed of a level or one or more hierarchies. The Store dimension in Figure 2 includes two hierarchies representing the administrative division and organizational structure.

Levels, such as a Product level in Figure 2, correspond to entity types in the ER model. Every instance of a level is called member. Levels contain one or several key attributes (underlined in Figure 2) identifying uniquely the members of

9 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/chapter/different-kinds-hierarchies-multidimensional-models/20687

Related Content

Adoption and Use of Open Source Infrastructure Software by Large Corporations: The Case of MySQL

RadhaKanta Mahapatra, Rashid Manzarand Vikram S. Bhadauria (2015). *Journal of Database Management* (pp. 1-17).

www.irma-international.org/article/adoption-and-use-of-open-source-infrastructure-software-by-large-corporations/153515

Co-creation and Collaboration in a Virtual World: A 3D Visualization Design Project in Second Life

Keng Siau, Fiona Fui-Hoon Nah, Brian E. Menneckeand Shu Z. Schiller (2012). *Cross-Disciplinary Models and Applications of Database Management: Advancing Approaches* (pp. 323-335).

www.irma-international.org/chapter/creation-collaboration-virtual-world/63672

Fault-Tolerant Quorum Consensus Scheme for Replication Control in Mobile Distributed Database Systems: FTQC

Siwoo Byunand Songchun Moon (1998). *Journal of Database Management* (pp. 16-24).

www.irma-international.org/article/fault-tolerant-quorum-consensus-scheme/51200

Business Data Warehouse: The Case of Wal-Mart

Indranil Bose (2009). *Selected Readings on Database Technologies and Applications* (pp. 112-133).

www.irma-international.org/chapter/business-data-warehouse/28549

Integrating Heterogeneous Data Sources in the Web

Angelo Brayner, Marcelo Meirellesand José de Aguiar Moraes Filho (2009). *Database Technologies: Concepts, Methodologies, Tools, and Applications* (pp. 2472-2488).

www.irma-international.org/chapter/integrating-heterogeneous-data-sources-web/8047