Current data warehouse and OLAP models include a time dimension that, like other dimensions, is used for grouping purposes (using the roll-up operation) or in a predicate role (using the slice-and-dice operation). The time dimension also indicates the time frame for measures (for example, in order to know how many units of a product were sold in March 2007). However, the time dimension cannot be used to keep track of changes in other dimensions, for example, when a product changes its ingredients or its packaging. Consequently, the “nonvolatile” and “time-varying” features included in the definition of a data warehouse (Sect. 2.5) apply only to measures, and this situation leaves to applications the responsibility of representing changes in dimensions. Kimball et al. [147] proposed several solutions for this problem in the context of relational databases, the slowly changing dimensions. Nevertheless, these solutions are not satisfactory, since they either do not preserve the entire history of the data or are difficult to implement. Further, they do not take account of all research that has been done in the field of temporal databases.

Temporal databases are databases that provide structures and mechanisms for representing and managing information that varies over time. Much research has been done in the field of temporal databases over the last few decades (e.g., [55, 72, 276, 277]). Therefore, combination of the research achievements in temporal databases and data warehouses has led to the new field of temporal data warehouses. Temporal data warehouses raise many issues, including consistent aggregation in the presence of time-varying data, temporal queries, storage methods, and temporal view materialization. Nevertheless, very little attention has been given by the research community to conceptual and logical modeling for temporal data warehouses, or to the analysis of what temporal support should be included in temporal data warehouses.

In this chapter, we propose a temporal extension of the MultiDim model. The chapter starts by giving in Sect. 5.1 a general overview of slowly changing dimensions, which is the mostly used approach in the data warehouse community for managing changes in dimension data. Section 5.2 briefly introduces
Temporal Data Warehouses

some concepts related to temporal databases. In Sect. 5.3, we give a general overview of the proposed temporal extension. Then, we refer to the inclusion of temporal support in the various elements of the model, i.e., levels, hierarchies, fact relationships, and measures. In Sect. 5.4, we discuss temporal support for attributes and for a level as a whole, and in Sect. 5.5 we refer to temporal hierarchies and present different cases considering whether it is important to store temporal changes to levels, to the links between them, or to both levels and links. Temporal fact relationships are discussed in Sect. 5.6, and temporal measures in Sect. 5.7. Section 5.8 examines the issue of different temporal granularities: for example, the source data may be introduced on a daily basis but the data in the data warehouse may be aggregated by month. The temporal support in the MultiDim model is summarized in Sect. 5.9, where the metamodel is presented.

This chapter also includes a mapping of our temporal multidimensional model into the classical (i.e., nontemporal) entity-relationship and object-relational models. After describing in Sect. 5.10 the rationale for mapping to these two models, Sect. 5.11 presents the logical representation of temporality types, temporal levels, temporal hierarchies, temporal fact relationships, and temporal measures. This section refers also to the inadequacy of relational databases for representing temporal data. After summarizing the mapping rules in Sect. 5.12, we present in Sect. 5.13 various implementation considerations that should be taken into account in performing aggregation of measures. Finally, Sect. 5.14 surveys work related to temporal data warehouses and temporal databases, and Sect. 5.15 concludes the chapter.

5.1 Slowly Changing Dimensions

The problem of managing changes to dimension data represented as a star or snowflake schema was addressed by Kimball et al. [147]. These authors defined slowly changing dimensions as dimensions where attributes do not change over time very frequently and users need to store these changes in a data warehouse. They proposed three solutions, called type 1, type 2, and type 3, to which we refer next.

In the type 1 or overwrite model, when an attribute changes, the new value overwrites the previous value. This is the current situation in traditional data warehouses, where attributes always reflect the most recent value without maintaining the history of changes. Suppose that in Fig. 5.1a the first product changes its size from 375 to 500. In this case, when the record of the product is updated with the new value, 500, we lose track of the previous value, 375. Note that this solution may give incorrect analysis results, since measures may be affected by changes to dimension data and these cannot be observed.

In the type 2 or conserving-history model, every time an attribute changes, a new record is inserted containing the new value of the attribute that has changed and the same values as before for the remaining attributes.