

电子商务辅导：数据仓库的建模方法电子商务师考试 PDF 转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/645/2021_2022__E7_94_B5_E5_AD_90_E5_95_86_E5_c40_645463.htm

数据仓库的建模方法
逻辑建模是数据仓库实施中的重要一环,因为它能直接反映出业务部门的需求,同时对系统的物理实施有着重要的指导作用。目前较常用的两种建模方法是所谓的第三范式 (3NF,即 Third Normal Form)和星型模式 (Star-Schema),我们将重点讨论两种方法的特点和它们在数据仓库系统中的适用场合。

一、什么是第三范式 范式是数据库逻辑模型设计的基本理论,一个关系模型可以从第一范式到第五范式进行无损分解,这个过程也称为规范化 (Normalize)。在数据仓库的模型设计中目前一般采用第三范式,它有非常严格的数学定义。如果从其表达的含义来看,一个符合第三范式的关系必须具有以下三个条件: 1. 每个属性的值唯一,不具有多义性. 2. 每个非主属性必须完全依赖于整个主键,而非主键的一部分. 3. 每个非主属性不能依赖于其他关系中的属性,因为这样的话,这种属性应该归到其他关系中去。我们可以看到,第三范式的定义基本上是围绕主键与非主属性之间的关系而作出的。如果只满足第一个条件,则称为第一范式.如果满足前面两个条件,则称为第二范式,依此类推。因此,各级范式是向下兼容的。

二、什么是星型模式 星型模式是一种多维的数据关系,它由一个事实表(Fact Table)和一组维表(Dimension Table)组成。每个维表都有一个维作为主键,所有这些维则组合成事实表的主键,换言之,事实表主键的每个元素都是维表的外键。事实表的非主属性称为事实 (Fact),它们一般都是数值或其他可以进行计算的数据.而维大

都是文字、时间等类型的数据。三、第三范式和星型模式在数据仓库中的应用 一个数据仓库的基本结构可以分成如图所示的四层:也有一些企业由于这样那样的原因,没有建立全企业范围的数据仓库,而是建立基于部门应用的独立数据集市(有关数据集市与数据仓库的比较,请参阅本报今年第 27 期上笔者编译自 Bill Inmon 的文章)。大多数人在设计中央数据仓库的逻辑模型时,都按照第三范式来设计.而在进行物理实施时,则由于数据库引擎的限制,不得不对逻辑模型进行不规范处理 (De-Normalize), 以提高系统的响应速度,这当然是以增加系统的复杂度、维护工作量、磁盘使用比率 (指原始数据与磁盘大小的比率)并降低系统执行动态查询能力为代价的。100Test 下载频道开通, 各类考试题目直接下载。详细请访问 www.100test.com