

Impact de l'évolution de la nomenclature des membres de dimension sur l'entrepôt de données

Inès ZOUARI TURKI*, Faïza GHOZZI JEDIDI**, Rafik BOUAZIZ***

Laboratoire MIRACL,
Faculté des Sciences Économiques et de Gestion de Sfax,
Route de l'Aérodrome km 4.5, B.P. 1088 – 3018 Sfax, Tunisie.

*Ines.Zouari@isimsf.rnu.tn

**Jedidi_Faiza@yahoo.fr

***Raf.Bouaziz@fsegs.rnu.tn

Résumé. Les systèmes courants de gestion d'entrepôts de données (**Data Warehouses** : DW) permettent de traiter les évolutions des faits, mais pas les évolutions affectant les schémas des DW et les instances de dimension. Plusieurs solutions, basées essentiellement sur l'historisation et/ou le versionnement des éléments de DW, ont été proposées pour ces dernières. Néanmoins, peu sont les travaux qui ont traité des évolutions des nomenclatures des instances de dimension et de leurs effets sur les analyses. Dans cet article, nous proposons une classification des différents types d'évolution de nomenclature, et étudions les effets de ces évolutions sur la cohérence des analyses. Les solutions que nous envisageons à chacun de ces aspects d'évolution se situent dans le cadre d'un système de gestion de DW temporel multiversions.

1 Introduction

Un entrepôt de données, mieux connu sous le vocable anglais *Data Warehouse* (DW), est un outil précieux pour la préparation de l'information de synthèse, nécessaire au processus de prise de décision. La structuration d'un DW conformément au modèle multidimensionnel de données est constituée par un fait et un certain nombre de dimensions. Un fait comporte des mesures dont les valeurs dépendent du contexte établi par les dimensions, qui regroupent les paramètres pertinents des différentes entités d'analyse. Une dimension est souvent organisée selon une ou plusieurs hiérarchies dont chacune comporte un certain nombre de niveaux. Une hiérarchie spécifie la façon selon laquelle les mesures peuvent être agrégées. Les instances des niveaux sont appelées **membres de dimensions** (MemD). La modélisation d'un DW dépend du nombre de niveaux par dimension. Elle donne soit un modèle en étoile lorsque chaque dimension présente un seul niveau, soit un modèle en flocons de neige, lorsque certaines dimensions présentent des niveaux hiérarchisés. Par ailleurs, le modèle en constellation est d'un grand intérêt pour structurer un DW comportant plusieurs tables de fait. Il permet de mettre en relief le partage des dimensions par ces tables.

Sachant que les systèmes classiques de gestion de DW sont conçus pour traiter l'évolution des données de transaction, ils s'approprient mal à prendre en considération les changements qui peuvent affecter les dimensions, aussi bien au niveau de leurs données qu'au niveau de leurs structures. Ceci est dû à l'hypothèse de l'orthogonalité de la dimension