

# Approche semi-automatisée de conception de schémas multidimensionnels valides

Ahlem Soussi\*, Jamel Feki\*\*, Faïez Gargouri\*\*\*

\*Centre de Calcul El Khawarizmi,  
Campus Universitaire de la Manouba, 2010 ; Manouba, Tunisie  
ahlem.soussi@cck.rnu.tn

\*\*Faculté des Sciences Economiques et de Gestion de Sfax  
B.P. 1088 ; Sfax, Tunisie  
jamel.feki@fsegs.rnu.tn

\*\*\*Institut Supérieur d'Informatique et du Multimédia de Sfax  
Route Mharza km 1,5 ; Sfax, Tunisie  
faiez.gargouri@fsegs.rnu.tn

**Résumé.** Cet article s'inscrit dans le cadre d'une approche semi-automatisée de conception de schémas multidimensionnels. Cette approche accepte des besoins OLAP, exprimés indépendamment de toute source OLTP, sous forme de tableaux n-dimensionnels, puis construit des schémas en étoile *idéaux*. En vue de valider ces étoiles idéales par rapport à une source d'alimentation, nous proposons d'une part, une méthode systématique pour effectuer la correspondance de chaque concept multidimensionnel avec la source et d'autre part, des règles de raffinement des étoiles. Les schémas en étoile valides seront ensuite fusionnés pour générer des schémas en constellation.

## 1 Introduction

Les entreprises passent à l'ère de l'information : leur défi est de compléter leur système d'information transactionnel OLTP («*On-Line Transaction Processing*»), à vocation de production, par un système d'information décisionnel (SD), à vocation de pilotage. Ceci est facilité avec l'apparition des entrepôts de données (ED) qui permettent désormais à tous les utilisateurs d'accéder à l'information stratégique. Un ED (ou *Data Warehouse*) est le lieu de stockage centralisé d'un extrait des bases de production. Il intègre et «*historise*» les données utiles pour la prise de décision. Son organisation doit faciliter la gestion efficace des données et la conservation des évolutions [Ravat *et al.*, 2001]. Dans la majorité des SD, un ED est restructuré en un ensemble de magasins de données (MD) (ou *Data Marts*). Un MD est un extrait de l'entrepôt. Les données extraites sont adaptées à une classe de décideurs ou à un usage particulier ; il est modélisé de manière multidimensionnelle qui s'adapte bien aux analyses OLAP («*On Line Analytical Processing*») [Teste, 2000].

La phase de création d'un schéma d'ED ou de MD passe par les étapes de modélisation conceptuelle, logique et physique [Phipps et Davis, 2002]. Nous nous intéressons dans cet article à la modélisation conceptuelle des schémas de MD et à son automatisation. Dans notre approche, les besoins OLAP sont exprimés sous forme de tableaux n-dimensionnels. Nous partons de ces tableaux pour proposer une méthode systématique de génération et de validation de schémas multidimensionnels en *étoile* ou en *constellation*. Un schéma en étoile est constitué d'un fait central (sujet analysé) et de plusieurs dimensions (axes d'analyses) ; un