

Approche de modélisation multidimensionnelle des données complexes : Application aux données médicales

Sid Ahmed Djallal Midouni*, Jérôme Darmont**, Fadila Bentayeb**

* Département d'informatique
Université Abou Bekr Belkaid - Tlemcen
B.P.230- Tlemcen 13000, Algérie
djmidouni@hotmail.com

** Université de Lyon (ERIC Lyon 2)
5 avenue Pierre Mendès-France
69676 Bron Cedex - France
jerome.darmont@univ-lyon2.fr, bentayeb@eric.univ-lyon2.fr

Résumé. La vocation d'un entrepôt de données est l'analyse de données pour l'aide à la décision dans les entreprises. La modélisation multidimensionnelle est la base des entrepôts de données et de l'analyse en ligne (OLAP). Ces techniques sont efficaces pour traiter les données simples numériques, mais elles ne sont pas adaptées aux données variées et hétérogènes provenant de différentes sources, appelés communément données complexes. Dans cet article, nous abordons le problème de la modélisation multidimensionnelle des données complexes à travers le cas des données médicales du projet MAP (Médecine d'Anticipation Personnalisée). Nous proposons un métamodèle multidimensionnel étendu pour les données médicales en généralisant le modèle cardiovasculaire du projet MAP. Enfin, nous avons spécifié et réalisé un outil d'aide à la conception d'entrepôt de données médicales.

Mots Clés : Modélisation des entrepôts de données, modèle multidimensionnel, analyse en ligne, données complexes.

1 Introduction

L'intérêt pour l'analyse de données s'est développé énormément ces dernières années. Les entreprises se sont rendues compte de l'efficacité de la technologie OLAP (*On-line Analytical Processing*) dans l'analyse et l'exploration des données. Cette technologie est utilisée dans les systèmes d'aide à la décision. Le plus souvent, ces systèmes sont basés sur des techniques d'entreposage de données pour exploiter la grande masse d'informations disponibles dans les entreprises à des fins d'analyse et d'aide à la décision.

La manière la plus appropriée pour faciliter cette analyse OLAP est la modélisation multidimensionnelle des données. Cette dernière représente les données comme des points dans un espace multidimensionnel, Kimball (1996), Vassiliadis et Sellis (1999).

La modélisation multidimensionnelle est donc une technique qui vise à organiser les données de telle sorte que les applications OLAP soient performantes et efficaces. Cependant, cette technique n'est pas adaptée à un certain type de données, dites complexes.