

# Problèmes d'additivité dus à la présence de hiérarchies complexes dans les modèles multidimensionnels : définitions, solutions et travaux futurs

Marouane Hachicha, Jérôme Darmont

Université de Lyon (ERIC Lyon 2)  
5 Avenue Pierre Mendès-France  
69676 Bron Cedex  
France

Courriel:prénom.nom@univ-lyon2.fr

**Résumé.** De nos jours, les entrepôts de données et les outils d'analyse OLAP sont très utilisés dans les entreprises qui ont besoin de systèmes décisionnels qui s'adaptent à toutes les situations particulières du monde réel, pour éviter les erreurs d'analyse (plus connues dans la littérature sous le nom de problèmes d'additivité ou *summarizability issues* en Anglais). Dans cet article, nous présentons un état de l'art des travaux sur les problèmes d'additivité dus à la présence de hiérarchies complexes dans les modèles multidimensionnels et discutons des travaux restant à mener dans ce domaine.

## 1 Introduction

De nos jours, les systèmes d'information décisionnels, basés sur les entrepôts de données et les analyses OLAP (*On-Line Analytical Processing*), sont très utilisés en entreprise dans un but d'aide à la décision. Afin de les préparer au stockage et à l'analyse, les données décisionnelles sont modélisées, au niveau conceptuel, d'une façon dite multidimensionnelle. La *modélisation multidimensionnelle* consiste à représenter les données décisionnelles selon leurs niveaux d'agrégation (faits et dimensions qui peuvent être hiérarchisées). Par exemple, le modèle conceptuel UML de la figure 1 représente des faits *ventes* décrits selon deux dimensions : (1) la dimension *géographique* organisée en hiérarchie : *ville-région-pays* (du plus détaillé au plus général) et (2) la dimension *date* organisée en hiérarchie : *jour-semaine-mois-année* (du plus détaillé au plus général). D'après Mazón et al. (2009), dans un modèle multidimensionnel, le concepteur doit garantir (1) la représentation adéquate des associations fait-dimension et (2) la représentation adéquate des différents niveaux d'agrégation d'une dimension exprimée en hiérarchie. Généralement, les situations du monde réel sont modélisées selon des hiérarchies simples. Les associations entre les différents niveaux d'une *hiérarchie simple* sont de type "un à plusieurs". Mais, certaines situations du monde réel sont modélisées selon des hiérarchies dites complexes. Par définition, une *hiérarchie complexe* est une hiérarchie qui est à la fois non-stricte et non-couvrante (Hachicha et al., 2012).

Une hiérarchie est dite *non-stricte* lorsque l'association entre deux niveaux hiérarchiques est de type "plusieurs à plusieurs" et non de type "un à plusieurs" (Malinowski et Zimányi,