

Résumé de cubes de données multidimensionnelles à l'aide de règles floues¹

Yeow Wei Choong*/**, Anne Laurent**
Dominique Laurent***, Pierre Maussion*

*HELP Institute - BZ-2 Pusat Bandar Damansara - 50490 Kuala Lumpur -
MALAYSIA

choongyw@help.edu.my, maussp@help.edu.my

**LIRMM - Université Montpellier II - 161 rue Ada - 34392 Montpellier Cedex 5
anne.laurent@lirmm.fr

***LICP - Université de Cergy-Pontoise - 2, av Chauvin - 95302 Cergy-Pontoise
dominique.laurent@dept-info.u-cergy.fr

Résumé. Dans le contexte des entrepôts de données, et des magasins de données multidimensionnelles, les outils OLAP fournissent des moyens aux utilisateurs de naviguer dans leurs données afin d'y découvrir des informations pertinentes. Cependant, les données à traiter sont souvent très volumineuses et ne permettent pas une exploration systématique et exhaustive. Il s'agit donc de développer des traitements automatisés facilitant la visualisation et la navigation dans les données. Dans cet article, nous étudions une méthode originale permettant de construire et d'identifier de manière automatique et efficace des blocs de données similaires présents dans les cubes de données pouvant être exprimés sous la forme de règles. Cette méthode est fondée sur l'utilisation combinée d'un algorithme par niveaux (de type Apriori) et de la théorie des sous-ensembles flous. Cette théorie nous permet en effet de pallier les problèmes posés par le fait que les blocs de données calculés par notre algorithme peuvent se recouvrir.

MOTS CLES : Bases de données multidimensionnelles, algorithmes par niveaux, règles floues, visualisation de données.

1 Introduction

Les bases de données multidimensionnelles sont étudiées et utilisées depuis maintenant un peu plus de dix ans [Codd *et al.*, 1993]. Les bases de données relationnelles, définies pour des traitements de type OLTP (*On-Line Transactional Processing*) pour la production de données, se sont en effet révélées insatisfaisantes pour analyser et exploiter les gros volumes de données produits dans le cadre des traitements OLAP (*On-Line Analytical Processing*). Les bases de données multidimensionnelles ont alors été proposées à cette fin.

Les données, regroupées dans des entrepôts de données (ou *Data Warehouses*), sont sélectionnées pour construire des magasins de données (ou *Data Marts*) qui regroupent des données dédiées à une tâche d'analyse particulière. Ces magasins de données sont

¹Ces travaux ont été partiellement financés par l'Ambassade de France en Malaisie.