

Proposition d'un thème à la journée du 20 mai 2005

**GDR I3 - Fouille de données**

**Groupe de travail 3.4**

Prof. Rokia Missaoui

Laboratoire LARIM

Université du Québec en Outaouais, Canada

**Titre** : Prospection de données dans les cubes de données

**Résumé**

Les techniques OLAP (*Online Analytical Processing*) utilisées dans les entrepôts de données permettent d'explorer les cubes de données selon divers axes d'analyse (dimensions), divers degrés d'abstraction et différentes valeurs des dimensions. Cependant, ces techniques ne sont pas appropriées pour la découverte automatique et efficace de groupements, d'associations ou d'exceptions dans des représentations multidimensionnelles.

Dans le cadre de cette présentation, nous allons proposer quelques alternatives de prospection de données dans les cubes (tableaux multidimensionnels). Il s'agit en particulier de techniques d'extraction de règles d'association et de groupements (*clusters*) à partir des tables de faits, et de mécanismes de détection des valeurs aberrantes. Dans le premier cas, nous exploitons nos récents travaux sur la génération des motifs fermés fréquents pour une extraction efficace des règles d'association et des groupements. Dans le deuxième cas, nous présentons le principe ainsi que les avantages et les inconvénients d'une solution fondée sur la modélisation log-linéaire qui permet de sélectionner un modèle parcimonieux de représentation des données (avec un nombre réduit de paramètres) et d'identifier les valeurs extrêmes dans les cellules du cube.