

Analyse en ligne d'objets complexes avec l'analyse factorielle

Loïc Mabit*, Sabine Loudcher*, Omar Boussaid*

*Laboratoire ERIC, Université Lumière Lyon 2
Université de Lyon
5 avenue Pierre Mendès-France
69676 Bron Cedex
loic.mabit@gmail.com
{sabine.loudcher | omar.boussaid}@univ-lyon2.fr

Résumé. Les entrepôts de données et l'analyse en ligne OLAP (*On-line Analysis Processing*) présentent des solutions reconnues et efficaces pour le processus d'aide à la décision. Notamment l'analyse en ligne, grâce aux opérateurs OLAP, permet de naviguer et de visualiser des données représentées dans un cube multidimensionnel. Mais lorsque les données ou les objets à analyser sont complexes, il est nécessaire de redéfinir et d'enrichir ces opérateurs OLAP. Dans cet article, nous proposons de combiner l'analyse OLAP et la fouille de données (*data mining*) afin de créer un nouvel opérateur de visualisation d'objets complexes. Cet opérateur utilise l'analyse factorielle des correspondances.

1 Introduction

Les entrepôts de données et l'analyse en ligne OLAP (*On-line Analysis Processing*) présentent des solutions reconnues et efficaces pour le processus d'aide à la décision. Notamment l'analyse en ligne, grâce aux opérateurs OLAP, permet de naviguer et de visualiser des données représentées dans un cube multidimensionnel. Cette technologie est bien rodée quant il s'agit de données simples où les faits sont analysés à travers des indicateurs numériques, souvent additifs, selon des descripteurs en général qualitatifs. Cependant, l'avènement des données complexes a remis en cause ce processus d'entreposage et d'analyse en ligne.

Souvent les données complexes sont plutôt représentées par un ensemble de descripteurs de bas niveau et/ou sémantiques. Lorsqu'il s'agit d'analyser par exemple une image, une vidéo ou tout autre objet de l'univers, il est alors plus efficace de considérer chacun de ces éléments comme une entité, à part entière, à observer. Celle-ci est alors perçue comme un objet complexe. Ce dernier peut être considéré comme un agrégat hétérogène de données qui, une fois réunies, forment une unité sémantique.

La vocation de l'analyse en ligne (OLAP) est de permettre d'agrèger des données pour résumer l'information qu'elles contiennent et de représenter celle-ci sous différents angles. L'utilisateur peut alors naviguer dans les données afin de les explorer. Les opérateurs OLAP sont définis pour des données classiques. Ils sont par conséquent inadaptés quand il s'agit d'objets complexes. Le recours à d'autres techniques, par exemple de fouille de données, peut s'avérer