

Modèles d'arbre pour XOLAP

Marouane Hachicha, Jérôme Darmont

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

Courriel : prénom.nom@univ-lyon2.fr

Résumé. Avec l'avènement de XML comme standard de représentation de données décisionnelles, les entrepôts de données XML trouvent leur place dans le développement de solutions décisionnelles. Dans ce contexte, il devient nécessaire de permettre des analyses OLAP sur des cubes de données XML. Afin de contribuer à ces recherches, de définir un cadre formel et de permettre l'optimisation indispensable des requêtes décisionnelles exprimées en XQuery, nous travaillons à définir une algèbre XML-OLAP (ou XOLAP). En premier lieu, nous avons exprimé avec l'algèbre XML TAX (*Tree Algebra for XML*) les opérateurs OLAP usuels. Il s'agit maintenant de prendre en compte les structures complexes permises par XML. Comme premier pas, nous proposons dans cet article un opérateur *rollup* basé sur un modèle d'arbre qui prend en compte des données XML multidimensionnelles organisées en hiérarchies complexes.

Mots clés : XML, Entrepôts de données, XOLAP, Modèles d'arbre, Hiérarchies complexes, Algèbre, Opérateur *rollup*

1 Introduction

La complexité des technologies décisionnelles telles que les entrepôts de données et l'analyse en ligne (OLAP, *On-Line Analytical Processing*) les rend peu attractives pour de nombreux utilisateurs potentiels. Dans ce contexte, une nette tendance à l'entreposage de données en ligne se dégage, avec notamment l'entreposage XML (Park et al., 2005; Zhang et al., 2005; Boussaïd et al., 2006). Le langage XML (*eXtensible Markup Language*) est en effet de plus en plus utilisé pour représenter des données décisionnelles (Beyer et al., 2005) et se montre particulièrement adapté pour modéliser des données dites complexes (Darmont et al., 2005) issues de sources hétérogènes et notamment du Web. Ainsi, plusieurs travaux visent à étendre le langage XQuery pour supporter des requêtes de type OLAP (Kay, 2006). Ces extensions doivent non seulement permettre d'effectuer des analyses OLAP classiques, mais aussi de prendre en compte dans l'analyse en ligne les spécificités des données XML, comme par exemple les structures hiérarchiques complexes telles que les *ragged hierarchies* (Beyer et al., 2005), qui seraient très difficiles à gérer dans un environnement relationnel.

Nous travaillons dans ce contexte à concevoir une algèbre XML-OLAP ou XOLAP (Wang et al., 2005) permettant d'exécuter des requêtes OLAP sur des données natives XML. Notre but général sera de définir un cadre formel actuellement inexistant dans le contexte XOLAP.