

# Résumés et interrogations de logs de requêtes OLAP

Julien Aligon\*, Patrick Marcel\*, Elsa Negre\*

\*Université François Rabelais Tours  
Laboratoire d'Informatique  
3 place Jean Jaurès  
41000 Blois  
prenom.nom@univ-tours.fr

**Résumé.** Une façon d'assister l'analyse d'entrepôt de données repose sur l'exploitation et la fouille de fichiers logs de requêtes OLAP. Mais, à notre connaissance, il n'existe pas de méthode permettant d'obtenir une représentation d'un tel log qui soit à la fois concise et exploitable. Dans ce papier, nous proposons une méthode pour résumer et interroger des logs de requêtes OLAP. L'idée de base est qu'une requête résume une autre requête et qu'un log, qui est une séquence de requêtes, résume un autre log. Notre cadre formel est composé d'une algèbre simple destinée à résumer des requêtes OLAP, et d'une mesure évaluant la qualité du résumé obtenu. Nous proposons également plusieurs stratégies pour calculer automatiquement des résumés de logs de bonne qualité, et nous montrons comment des propriétés simples sur les résumés peuvent être utilisées pour interroger un log efficacement. Des tests sur des logs de requêtes MDX ont montré l'intérêt de notre approche.

## 1 Introduction

Il est maintenant établi que l'exploitation des logs de requêtes peut aider l'utilisateur qui analyse de grandes bases de données ou des entrepôts de données (Khoussainova et al., 2009). C'est particulièrement le cas dans un contexte collaboratif, par exemple pour proposer des recommandations (Chatzopoulou et al., 2009; Giacometti et al., 2009a,b; Stefanidis et al., 2009). Mais, à notre connaissance, même le simple fait de fournir à l'utilisateur une représentation concise, i.e. un résumé, de ce qui apparaît dans un log de grande taille n'a que très rarement été abordé dans la littérature (Khoussainova et al., 2009). Or un résumé aurait de nombreux avantages, comme aider l'administrateur à paramétrer le serveur OLAP s'il indique les membres fréquemment accédés, permettre à un utilisateur d'avoir une première idée du genre de requêtes lancées par d'autres utilisateurs, ou aider l'utilisateur à lancer de nouvelles sessions d'analyse, par exemple en ne refaisant pas ce qui a déjà été fait, ou à écrire de nouvelles requêtes, en partant de, ou en prenant exemple sur, ce qui a déjà été fait.

Dans cet article, nous développons les travaux entrepris sur le résumé de log de requêtes OLAP (Aligon et al., 2010a), où nous proposons un langage de manipulation de requêtes (appelé *QSL*) pour résumer des requêtes OLAP, une mesure de la qualité d'un résumé et un algorithme glouton de construction automatique de résumés de log de requêtes utilisant *QSL*. Sur