

RecoOLAP : Un système de recommandation de requêtes OLAP

Elsa Negre

Université Paris-Dauphine, LAMSADE
elsa.negre@dauphine.fr

Résumé. Une session d'analyse OLAP peut être définie comme une session interactive durant laquelle un utilisateur lance des requêtes pour naviguer dans un cube. Très souvent, choisir quelle partie du cube va être naviguée par la suite, et, de ce fait, concevoir la prochaine requête, est une tâche difficile. Dans cette démonstration, nous proposons un système de recommandation qui utilise ce que tous les utilisateurs du système OLAP ont fait pendant leurs précédentes explorations du cube afin de recommander des requêtes MDX à l'utilisateur.

1 Introduction

La recommandation de requêtes dans les bases de données a fait l'objet de nombreuses recherches, (Khoussainova et al. (2009), Chatzopoulou et al. (2009)), en particulier dans les systèmes OLAP (OnLine Analytical Processing) où l'utilisateur navigue interactivement dans un cube en lançant une séquence de requêtes (et plus particulièrement des requêtes MDX) sur un entrepôt de données, ce que nous appelons une session d'analyse (ou session) dans la suite de l'article. Ce processus est souvent pénible puisque l'utilisateur peut ne pas avoir d'idée sur ce que pourrait être la prochaine requête (Sarawagi (2000)).

Les systèmes de recommandation existants sont généralement classés selon deux catégories : les méthodes basées sur le contenu et celles basées sur le filtrage collaboratif (Adomavicius et Tuzhilin (2005)). Les méthodes basées sur le contenu recommandent à l'utilisateur des objets similaires à ceux qui l'ont intéressé dans le passé, tandis que les méthodes basées sur le filtrage collaboratif recommandent des objets qui ont intéressé des utilisateurs similaires. Au vu de la nature multi-utilisateur des entrepôts de données, dans nos travaux précédents (cf. Negre (2009) pour une vue d'ensemble), nous avons proposé des techniques issues du filtrage collaboratif pour recommander des requêtes OLAP à l'utilisateur. L'idée principale est de calculer une similarité entre la séquence de requêtes de l'utilisateur courant et les séquences de requêtes précédentes qui ont été enregistrées par le serveur.

Nous présentons, dans cette démonstration, notre cadre générique de génération de recommandations élaboré lors de nos travaux, appliqué sur des données synthétiques, qui permet de recommander des requêtes MDX, en utilisant le log du serveur, c'est-à-dire, l'ensemble des sessions précédentes sur le cube, et la séquence de requêtes de la session courante.

Cette démonstration est organisée comme suit : la section 2 présente notre cadre générique de recommandations de requêtes et la section 3 détaille l'architecture de notre système. Enfin, nous concluons section 4.