

Recherche d'information personnalisée dans les bibliothèques numériques scientifiques

Thanh-Trung Van*, Michel Beigbeder *

* Centre G2I/Département RIM
Ecole Nationale Supérieure des Mines de Saint-Etienne
158 cours Fauriel, 42023 Saint-Etienne, FRANCE
{van,mbeig}@emse.fr

Résumé. Dans cet article nous présentons nos travaux sur la recherche d'information personnalisée dans les bibliothèques numériques. Nous utilisons des profils utilisateurs qui représentent des intérêts et des préférences des utilisateurs. Les résultats de recherche peuvent être re-triés en tenant compte des besoins d'informations spécifiques de différentes personnes, ce qui donne une meilleure précision. Nous étudions différentes méthodes basées sur les citations, sur le contenu textuel des documents et des approches hybrides. Les résultats des expérimentations montrent que nos approches sont efficaces et applicables dans le cadre des bibliothèques numériques.

1 Introduction

La recherche d'information dans les bibliothèques numériques est souvent une tâche ennuyeuse et fastidieuse. Les utilisateurs doivent répéter le processus d'envoyer les requêtes, regarder les résultats et modifier les requêtes jusqu'à ce qu'ils trouvent les informations pertinentes. Une des raisons principales est que les requêtes des utilisateurs sont souvent courtes et donc ambiguës. Par exemple, la même requête «java» peut être formulée par une personne qui s'intéresse au langage de programmation «java», et par une autre qui veut chercher des informations concernant une île en Indonésie. Cependant les moteurs de recherche renvoient le même résultat pour ces deux personnes. Même avec une plus longue requête comme «langage programmation java»; nous ne savons pas quels types de document cet utilisateur veut chercher. Si c'est un(e) programmeur(e), peut-être il/elle s'intéresse aux documents techniques sur le langage Java, si c'est un(e) enseignant(e), peut-être il/elle s'intéresse aux tutoriels de Java pour ses cours.

Le problème que nous avons mentionné peut être résolu en utilisant des techniques de personnalisation avec des *profils utilisateurs*. D'une manière générale, nous pouvons définir un profil d'utilisateur comme un ensemble structuré d'informations qui décrit les intérêts et/ou les préférences de cet utilisateur.

Le travail de Amato et Straccia (1999) est parmi les premiers travaux consacrés à définir un modèle de représentation de profil utilisateur dans les bibliothèques numériques, leur modèle est un modèle multidimensionnel dans lequel le profil utilisateur se compose de plusieurs