

Analyse du Comportement des utilisateurs exploitant une base de données vidéo

Sylvain Mongy*

*Univ. de Lille1, Bât. M3 59655 Villeneuve d'Ascq Cedex FRANCE
mongy@lil.fr,
<http://www-rech.enic.fr/MIIRE>

Résumé. Dans cet article, nous présentons un modèle de fouille des usages de la vidéo pour améliorer la qualité de l'indexation. Nous proposons une approche basée sur un modèle à deux niveaux représentant le comportement des utilisateurs exploitant un moteur de recherche vidéo. Le premier niveau consiste à modéliser le comportement lors de la lecture d'une vidéo unique (comportement intra vidéo), le second à modéliser le comportement sur l'ensemble d'une session (comportement inter vidéo). A partir de cette représentation, nous avons développé un algorithme de regroupement, adapté à la nature particulière de ces données. L'analyse des usages de la vidéo nous permet d'affiner l'indexation vidéo sur la base de l'intérêt des utilisateurs.

1 Introduction

De par le développement rapide des techniques de stockage et de diffusion, les vidéos, notamment digitalisées, sont de plus en plus nombreuses et accessibles. En particulier, les agences de presse, les diffuseurs TV, les agences de publicité travaillent sur des ressources vidéo grandissantes. Pour être à même de travailler sur de tels volumes, des technologies adaptées doivent être mises en oeuvre. La « fouille des usages de la vidéo », qui cherche à analyser les comportements des utilisateurs sur des ensembles de vidéo est l'une des techniques clé émergentes pour optimiser les accès aux vidéos.

Dans cet article, nous proposons d'analyser les comportements des utilisateurs d'un moteur de recherche vidéo pour améliorer la qualité de l'indexation textuelle. Notre objectif est de comprendre pourquoi et comment chacune des séquences vidéo est visionnée. Par exemple, les utilisateurs recherchant des vidéos concernant le mot-clé « montagne » visionnent successivement les vidéos (18, 73, 29) qui sont retournées dans cet ordre par le moteur de recherche. Si l'on note que dans la majeure partie des cas, la vidéo 29 est visionnée totalement alors que les vidéos 18 et 73 ne le sont que partiellement, on en déduit que, selon l'utilisateur, le concept de « montagne » est mieux exprimé par la vidéo 29 que par les vidéos 18 et 73. En conclusion, la vidéo 29 doit être proposée en premier aux utilisateurs lors des futures recherches sur le concept « montagne ». Son poids dans la vidéo 29 s'en trouve augmenté et celui des vidéos 18 et 73 réduit.

Dans ce papier nous présentons une approche qui combine usage intra-vidéo et usage inter-vidéo pour générer des profils de visite sur un moteur de recherche vidéo dans le contexte de