

# Identification de communautés d'usage du web depuis un graphe issu des fichiers d'accès

Yacine Slimani<sup>\*,\*\*</sup>, Abdelouahab Moussaoui<sup>\*</sup>,  
Yves Lechevallier<sup>\*\*</sup>, Ahlem Drif<sup>\*\*\*</sup>

<sup>\*</sup>Laboratoire de Recherche en Informatique Appliquée (L.R.I.A)  
Université Ferhat Abbas Sétif 19000, Algérie  
slimani\_y09@univ-setif.dz, moussaoui.abdel@gmail.com

<sup>\*\*</sup>Institut National de Recherche en Informatique et en Automatique (I.N.R.I.A)  
Domaine de Voluceau Rocquencourt B.P. 105, 78153 Le Chesnay Cedex, France  
yves.lechevallier@inria.fr

<sup>\*\*\*</sup>Laboratoire de Réseau et Systèmes Distribuées (L.R.S.D)  
Université Ferhat Abbas Sétif 19000, Algérie  
adrif.univsetif@gmail.com

## Résumé

L'analyse du comportement des utilisateurs dans le processus de fouille de données d'usage du Web (Web Usage Mining) a suscité un intérêt croissant et fait l'objet de plusieurs travaux de recherches qui ont été proposés pour analyser les fichiers d'accès au web (Logs) et offrir de nouvelles connaissances utiles afin de restructurer et personnaliser le design des sites web. Dans ce papier, nous présentons une approche d'extraction des données basées sur la fonction de modularité des graphes. La découverte des communautés existantes dans le graphe issu de la structuration des navigations des usagers en sessions a permis de discriminer ces communautés par leur sujet d'intérêt et donc d'extraire des connaissances pertinentes.

## Summary

The extraction of user behavior in the process of Web Usage Mining has attracted a growing interest and is the subject of several researches that have been proposed to analyze log files and provide a useful knowledge to restructure and customize the design of a web site. In this paper, we present an approach for extracting data based on the modularity function and we discuss the obtained results. This approach discovers the existing communities by modeling the data obtained in the pre-processing operation as a weighted graph. The method discriminates the communities through their subject of interest and extract a relevant knowledge.