

Outil de représentation des évolutions de communautés d'intérêts

Anne Lavallard, Luigi Lancieri

France Telecom R&D
42 Rue des coutures 14000 Caen
anne.lavallard@francetelecom.com
luigi.lancieri@francetelecom.com
<http://www.ensicaen.ismra.fr/~lancieri>

Résumé. Cet article présente un système de visualisation permettant l'observation des comportements collectifs implicites. Il s'agit de reconnaître et de représenter des communautés à partir des connexions Internet des utilisateurs : les utilisateurs sont répartis en communautés en fonction des similarités entre des listes de termes établies sur l'analyse des documents consultés par chacun d'eux. L'étude est rendue dynamique par la comparaison des communautés reconnues sur des périodes de temps connexes. L'outil décrit ci après offre deux représentations différentes de ces communautés : une vision des liaisons thématiques entre les utilisateurs sur chaque période étudiée et une vue comparative des communautés reconnues sur toute la durée de l'étude.

1 Introduction

La problématique de cette étude est de proposer des outils de visualisation de la dynamique des communautés d'intérêts. Même si des algorithmes permettent de calculer des communautés thématiques en analysant les documents consultés par les utilisateurs, la lisibilité des résultats est parfois faible. En effet, d'une part, il est difficile de prendre en compte certains paramètres comme l'aspect dynamique du comportement. D'autre part, les résultats sont parfois difficilement exploitables directement, faute de représentation visuelle adaptée pour guider l'interprétation des données.

Pour répondre à cette problématique, nous avons mis au point différents outils afin de pouvoir identifier et visualiser différents paramètres liés à l'évolution des groupes d'utilisateurs. L'accent a été porté sur l'aspect temporel des données observées. Le but final étant de permettre une observation optimisée des comportements collectifs par la recherche des propriétés dynamiques liant des groupes d'utilisateurs (communautés implicites). L'identification et la visualisation de ces propriétés sur des périodes de temps connexes permet par confrontation d'évaluer les évolutions dans le temps de ces communautés.

Cette étude a été réalisée sur la base de données d'activités réelles des utilisateurs sur une période de 17 mois (traces de consultations et analyse des documents associés). Ces données ont été rendues anonymes pour concilier notre souci de préservation de la vie privée avec notre approche non intrusive de l'observation de l'activité.

Outre l'apport du point de vue de l'ergonomie de la représentation de l'activité collective, ces premiers travaux ont permis d'ouvrir des nouvelles pistes de réflexions permettant d'affiner l'analyse des résultats. Cet objectif est particulièrement important dans un contexte d'optimisation de l'exploitation des phénomènes d'intelligence collective.