

# Indicateurs de diversité dans les graphes de recommandation

Lionel Tabourier, Raphaël Fournier-S'niehotta

stages@complexnetworks.fr, fournier@cnam.fr

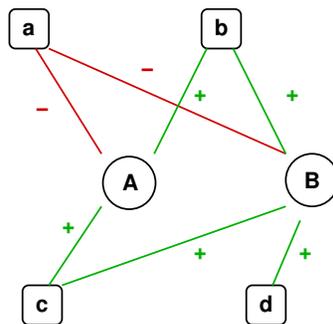
<http://complexnetworks.fr>

LIP6 – CNRS et UPMC – Paris

Au cours de la dernière décennie, les systèmes de personnalisation de recherche sont devenus omniprésents sur le web, qu'il s'agisse des moteurs de recherche mêmes ou des pratiques de recommandations (vidéos, sites, contacts, achats de produits etc). La recherche de contenus en ligne est désormais assistée par des algorithmes qui conditionnent l'expérience de navigation des utilisateurs. Cette évolution soulève des questions capitales aussi bien du point de vue informatique que sociologique. En particulier, un effet désormais bien identifié est le phénomène de **bulle de filtres** (ou *filter bubble*), correspondant à la restriction de la diversité des sources d'information imposée par les systèmes de recommandation. Comprendre et mesurer ce phénomène nécessite dans un premier temps de définir des outils pour **mesurer la diversité** des contenus recommandés.

L'objectif de ce stage est de proposer des métriques qui s'appuient sur la **structure du graphe** entre agents et objets recommandés pour définir la diversité dans ce contexte. La représentation des données sous forme de **graphes bipartis** (agent/objet) et **valués** permet en effet de rendre compte des évaluations faites par les agents des différents objets. Par ailleurs, l'utilisation d'une représentation en graphe permet d'introduire naturellement, par l'analyse de la structure, des concepts tels que la proximité de nœuds ou la ressemblance des environnements.

Un principe fondamental de la recommandation est de proposer à un utilisateur des objets bien évalués par d'autres utilisateurs ayant des goûts similaires aux siens. Par exemple, l'objet **d** serait recommandé à l'utilisateur **A** dans le graphe ci-dessous.



Une telle démarche peut sembler pertinente car elle exploite la connaissance du réseau dans le but d'offrir une recommandation plus sûre. En revanche, cela se fait au détriment de l'exploration : on limite l'accès de l'agent à de nouvelles communautés et de potentiels nouveaux intérêts.

Les mesures définies seront appliquées sur des données publiques correspondant à des domaines variés (recommandation de films, de titres musicaux, de marchandises etc). Selon les contextes, nous chercherons à comprendre si l'analyse de la structure peut révéler des choix favorables à la recherche d'un bon compromis exploration-exploitation.

Le stage sera co-encadré par Lionel Tabourier (LIP6 - UPMC) et Raphaël Fournier-S'niehotta (CEDRIC - CNAM). Il s'adresse à des étudiants issus de formations variées (réseaux, fouille de données, apprentissage ...). Les principales compétences recherchées sont une certaine connaissance de l'algorithmique de graphes, la capacité d'adaptation à des méthodes nouvelles et de l'intérêt pour les thématiques interdisciplinaires. Du point de vue des langages et outils utilisés, le/la stagiaire bénéficiera d'une grande liberté.