

Recommandation et personnalisation de contenu sur Facebook

Proposition de stage M2
Collaboration : LARIS-EISTI & Graphinium
www.eisti.fr, www.graphinium.com

Contexte de l'étude

L'analyse des réseaux sociaux est définie comme étant l'étude des entités sociales (les personnes dans les organisations qu'on appelle acteurs) ainsi que leurs interactions et leurs relations. Nous pouvons étudier les propriétés [2, 1] de la structure et son rôle ainsi que la position et le prestige de chaque acteur social [5]. Nous pouvons rechercher aussi les différents types de sous-graphes comme par exemple les communautés formées par des groupes d'acteurs ayant des intérêts communs. Le réseau social peut être aussi une source permettant l'élaboration de recommandations : trouver un expert dans un domaine donné, suggérer des produits à vendre, proposer un ami, etc.

Une première version de système de recommandation est déjà conçue au sein du laboratoire LARIS [3]. La recommandation dépend d'un côté, de la similarité entre les préférences des utilisateurs et la requête soumise et d'un autre côté de l'intermédiarité des nœuds-utilisateurs se trouvant sur les chemins de la solution. Pour effectuer une recherche dans le graphe, l'arbre couvrant le plus représentatif est extrait et ensuite il est exploré à l'aide d'un heuristique. D'un autre côté, un service de recherche social de concerts exploitant le protocole Open Graph de Facebook (POC PureConcerts) a été développé dans le cadre du projet SEPIA par Graphinium.

Description du stage

L'objectif de stage est d'améliorer le système de recommandation existant en intégrant une dimension sémantique avancée apportant plus de pertinence et puissance d'analyse au service de recommandation et personnalisation de

contenu sur Facebook. Cette étude sera menée dans le cadre de la collaboration LARIS/Graphinium (projet SEPIA).

Nous souhaitons intégrer dans le système de recommandation une librairie de mesures de similarité ou de pertinence dans le but d'améliorer la qualité des recommandations proposées [4]. Ces mesures seront de plusieurs natures : certaines mesures sont basiques et ressemblent à des mesures de similarités classiques, d'autres intégreront une dimension sémantique ou structurelle qui prend en compte la connaissance du domaine représentée sous forme d'ontologie, ou les ressemblances de comportements entre acteurs. Il s'agit d'étudier cet ensemble de mesures et d'élaborer un plan de test, afin de dresser un bilan permettant de mesurer la qualité des recommandations proposées aux acteurs en fonction des différentes mesures.

Une des applications possible sera le développement d'un service de partage sociale d'intérêts sur les sports, regroupant les sportifs, les événements, les compétitions, les stades, etc. Le but étant d'identifier les intérêts d'un utilisateur pour le sport afin d'en tirer partie des «posts» d'avis sur les événements sportifs. Les recommandations permettront d'identifier des sports/amis/compétitions et amis/communautés associés.

Tâches à réaliser

Nous nous intéressons dans un premier temps à l'identification des domaines d'intérêt et niveaux d'expertises d'un utilisateur ainsi qu'à la détection des sous groupes de personnes connectées sous Facebook. Ces communautés doivent être caractérisées par leurs natures de lien (fort, faible), leurs typologies, etc.

Une étude des différentes mesures de similarités est à élaborer dans un deuxième temps, ainsi que leur intégration dans le processus de la recommandation.

Un ensemble d'exigences est à respecter :

- Les services seront embarqués dans des outils déployés sur des sites web online pour accompagner la navigation des internautes.
- Une exigence de performance/temps de réponses en plus de la pertinence est nécessaire.
- Les différentes technologies et outils devront privilégier quand c'est possible des ressources open source.

Mots Clés

Recommandation sur Facebook, préférences sémantiques de l'utilisateur, analyse des réseaux sociaux.

Pré-requis et contraintes particulières

- Bon niveau et autonomie en programmation.
- Connaissances de base en théorie des graphes.
- Des connaissances sur les ontologies seront appréciées.

Environnement technique

Java, neo4j, mongodb, Solr, api Facebook/OpenGraph.

Informations complémentaires

Cadre de collaboration : Le laboratoire LARIS/EISTI et Graphinium (projet SEPIA).

Rémunération : 600 euros net mensuel.

Durée et début de stage : 6 mois à partir de mars.

Contacts : maria.malek@eisti.fr, jimi.fontaine@gmail.com, benmira@gmail.com.

Références

- [1] M. G. Everett and S. P. Borgatti. The centrality of groups and classes. *Journal of Mathematical Sociology*, 23(3) :181–201, 1999.
- [2] Linton C. Freeman. Centrality in social networks : Conceptual clarification. *Social Networks*, 1(3) :215–239, 1979.
- [3] Maria Malek and Dalia Sulieman. Exhaustive and guided algorithms for recommendation in a professional social network. In *7th conference on Application of Social Network Analysis (ASNA), Zurich*, September, 2010.

-
- [4] M.Malek, D. Sulieman, and H. Kadima. Integration of semantic user profile within a social recommendation system : semantic relevance measures characterization. *International Journal of complex systems in sciences, (Acte de IV International Conference Net-Works 2011 Complex Networks : Structure, Applications and Related Topics, El Escorial, Spain, 26-28 octobre.)*, 1, 2011.
- [5] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45 :167–256, Mar 2003.