

Des graphes de documents aux réseaux sociaux

Michel Plantié*, Michel Crampes*

*EMA-LGI2P, Parc Scientifique Georges Besse, 30035 Nîmes Cedex, France
michel.plantie,michel.crampes@mines-ales.fr, <http://www.lgi2p.ema.fr>

1 Introduction

Dans ce papier, nous cherchons à constituer des réseaux sociaux à partir de documents partagés par des personnes dans des plateformes collaboratives. Puis nous cherchons à constituer des listes de diffusions personnalisées.

2 Construction du réseau social

Organiser socialement des documents et élaborer des réseaux sociaux à partir du contenu des documents est depuis un certain temps un domaine actif de recherche. Beaucoup de travaux comme par exemple (Velardi et al., 2008; Mika, 2007) cherchent à utiliser les relations entre personnes dans un collectif afin de découvrir des connaissances communes et les exploiter pour différents objectifs.

Le concept de tribu : Nous appelons Tribu un sous-ensemble de personnes x_i concernées par un partage de document personnalisé. Tous les individus qui appartiennent à la même tribu possèdent un réseau social commun qui est mis en évidence par l'apparition de ces individus sur différents documents sous des formes différentes. Nous résonnerons ici sur un couple d'individus x, y , plus précisément sur le lien entre x et y .

2.1 Modèles de Forces de couples et tribus

Ci-dessous nous définissons différentes lois qui constitueront des tribus différentes et nous comparerons ensuite ces différents découpages. Nous définissons d'abord trois forces brutes, et nous définirons ensuite une mesure de pondération de ces forces.

La force simple d'un couple : Ici nous définissons la "fréquence" d'un couple (deux individus dans une certaine forme de relation) comme le nombre d'occurrences de ce couple parmi les concepts divisé par le nombre de concepts :

La proximité ou force pondérée d'un couple : La proximité d'une tribu (respectivement d'un couple) est la somme des inverses des nombres d'individus présents dans les concepts qui contiennent la tribu, divisée par le nombre total de concepts. Donc, plus le couple est "dilué" dans un grand nombre d'individus pour un concept, moins la contribution de ce concept à la proximité est forte. Quand une tribu (un couple) n'apparaît dans aucun concept, la force est