

Recherche de communautés Basée sur la complémentarité dans les Réseaux de personnes

Mouhamadou Niang *, Michel Plantié*

* Laboratoire LGI2P, Ecole des Mines, Parc Georges Besse, 30035 Nîmes
michel.plantie@mines-ales.fr,
niang20042002@gmail.com
<http://social-networks.mines-ales.fr>

Résumé. La recherche de communautés dans les graphes a fait l'objet de beaucoup de travaux de recherche. Les algorithmes existants se basent sur la notion de similarité, de distances ou de modularité. Dans cet article nous proposons une solution de recherche de communautés basée sur le partitionnement de graphes en **communautés de personnes complémentaires**.

1 Introduction

La recherche de communautés à partir de graphes a fait l'objet de plusieurs publications. Différentes méthodes d'algorithmes ont été proposées dans Fortunato (2009). La recherche de communautés est très importante pour résoudre de nombreux problèmes d'analyses des réseaux sociaux. La connaissance de ces communautés peut faciliter la recommandation d'amis, de groupes, d'événements, etc.

Dans nos travaux précédents nous avons étudié la complémentarité de personnes dans les réseaux représentés sous forme de graphes bipartis Crampes et Plantié (2015). La complémentarité permet d'étudier la cohésion d'un groupe, c'est à dire la capacité à associer des éléments d'un système pour un rendement meilleur. Par exemple composer une équipe de travail possible pour réaliser efficacement un projet, connaissant les compétences et les caractéristiques nécessaires pour la réalisation de projets ainsi que les compétences des candidats. Notre approche s'est basée sur la notion d'entropie et gain d'information.

2 Communautés de personnes complémentaires

Pour la conférence Marami2015, nous avons étudié trois approches de complémentarité de personnes dans les Réseaux. Les expérimentations qui ont été effectuées montrent que la complémentarité telle qu'elle est définie par Milgrom et Roberts (1990) donne les meilleurs résultats. Nous allons par la suite nous baser sur cette approche pour rechercher les communautés.

2.1 complémentarité de personne : définition et formalisation

La définition que nous avons retenue de la littérature sur la complémentarité est la suivante : deux éléments sont complémentaires, **si leur apport au système est plus grand quand ils sont présents ensemble que leur apport quand ils sont pris séparément.**

Communautés de personnes complémentaires

Elle a été formalisée par Milgrom et Roberts (1990) comme suit : Soit S un système et x et y deux éléments de S , x est complémentaire à y si et seulement si la condition de supermodularité suivante est vérifiée :

$$\prod(x \wedge y) + \prod(x \vee y) \geq \prod(x) + \prod(y) \quad (1)$$

où \prod est une fonction d'utilité. Une fonction d'utilité permet de mesurer la contribution d'un élément dans un ensemble. Dans notre approche nous utilisons comme fonction d'utilité : le gain d'information G . la formule 1 devient :

$$\text{complémentarité}(Pe_i, Pe_j) : G(Pe_i \wedge Pe_j) + G(Pe_i \vee Pe_j) \geq G(Pe_i) + G(Pe_j), \quad (2)$$

Pe_i (Pe_j) signifie la personne i (respectivement la personne j). L'algorithme de recherche de personnes complémentaires a été donné dans Crampes et Plantié (2015).

2.2 détection de communautés : définition et algorithme

Définition - communauté de personnes complémentaires :

Un sous-graphe dont les sommets sont des personnes ayant une forte complémentarité entre eux et une faible complémentarité avec les autres personnes du réseau. Nous utilisons ainsi la complémentarité comme moyen de partitionnement de nos communautés.

Algorithm 1 Algorithme de détection de communautés

donnée initiale :

- liste complète des candidats : *listecandidats*
- tableau des communautés M =vide

Calcul entropie système $H(S)$

Trouver Pe_j | $Gain(Pe_j) = \max$

numerodecommunaute \leftarrow 1

Ajouter Pe_j dans communauté $M(\text{numerodecommunaute})$

while (*listecandidats* \neq \emptyset) **do**

while ((*listecandidats* \neq \emptyset) and (*arretCouvertureTotale*())) **do**

$Pe_j \leftarrow \text{trouverLaPersonneComplementaire}(\text{numerodecommunaute})$

 Ajouter Pe_j dans $M(\text{numerodecommunaute})$

end while

if (*listecandidats* \neq \emptyset) **then**

numerodecommunaute \leftarrow *numerodecommunaute* + 1

 Ajouter Pe_k dans $M(\text{numerodecommunaute})$ | $Gain(Pe_k) = \max$

end if

end while

trouverLaPersonneComplementaire(*numerodecommunaute*) :

calculer pour $C = M(\text{numerodecommunaute})$, Pe_i | $\text{complémentarité}(Pe_i, C) = \max$

Sortie : Pe_i

3 Expérimentation

Nous testons notre algorithme sur trois graphes de tailles différentes. Tout d'abord nos expériences ont été effectuées sur un graphe biparti représentant 18 dames qui ont

participé à 14 événements Sociaux (voir Freeman (2003)). Nous testons ensuite l'algorithme sur un graphe biparti issu de Facebook représentant le partage de 700 photos entre 274 utilisateurs de facebook. Enfin nous expérimentons notre algorithme sur sur jeu de données beaucoup plus grand "Pubmed" contenant ¹ plus de 82000 auteurs et plus de 40000 articles scientifiques que nous nommons "Auteur-Articles". Les résultats sur les trois graphes sont présentés dans la figure 1.

jeux de données :	taille(nb personnes)	nombre de communautés trouvées
Women Events	18	3
Facebook	274	102
Auteur-Articles	82457	12309

FIG. 1 – *Graphes partitionnés en communautés de personnes complémentaires*

Discussion

Notre algorithme place les personnes les plus complémentaires dans une communauté pour chacun des graphes expérimentés. Nous pouvons démontrer par la suite que notre algorithme est calculable. Pour la suite de nos travaux nous cherchons à répondre à différentes questions. Quelle est la meilleur communauté? Les personnes sont elles dans la bonne communauté? Que se passe-t-il si on déplace une personne vers d'autres communautés? Le déplacement aura-t-il un impact sur les autres communautés? Comment juger de la stabilité du système? La suite des travaux sera de définir une mesure de qualité **locale** et **globale** des communautés de personnes complémentaires et critère de stabilité du système.

Références

- Crampes, M. et M. Plantié (2015). Complémentarité de personnes partageant des propriétés dans les Réseaux Sociaux. In *IC 2015 Ingénierie des connaissances*, Rennes, France.
- Fortunato, S. (2009). Community detection in graphs. *Physics Reports* 486(3-5), 103.
- Freeman, L. C. (2003). Finding social groups : A meta-analysis of the southern women data. In *Dynamic Social Network Modeling and Analysis. The National Academies*, pp. 39—97. Press.
- Milgrom, P. et J. Roberts (1990). Rationalizability, learning, and equilibrium in games with strategic complementarities. *econometrica*. In *Journal of the Econometric Society*. Citeseer.

Summary

Search of communities in graphs made the subject of many research tasks. The existing algorithms are based on the concept of similarity, distances or modularity. In this paper we propose a graph partitioning of communities based on people complementarity in a community.

1. <http://www.ncbi.nlm.nih.gov/pubmed>