

Projet de recherche – Master informatique

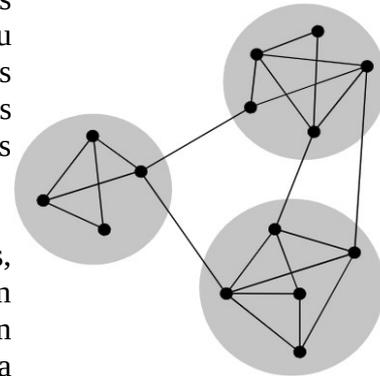
Communautés de graphes fermés transitivement

Chercheur référent :

Antoine Huchet – antoine.huchet@univ-lr.fr

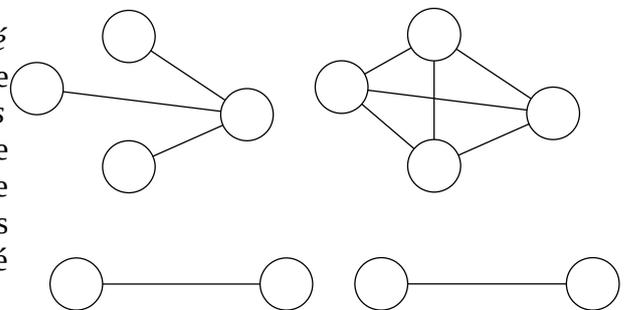
Contexte scientifique :

La structure communautaire d'un graphe représente des groupes de sommets entre lesquels il y a plus d'arêtes que vers le reste du graphe. Ceci sous-entend une interaction plus forte entre les sommets connectés qu'entre les sommets qui ne sont pas directement connectés par une arête. Dans la figure ci-contre, les communautés correspondent aux disques gris.



Or, dans de nombreux cas, comme dans un réseau d'ordinateurs, même si deux ordinateurs ne sont pas directement connectés, un signal peut transiter d'un ordinateur à l'autre en passant par un troisième ordinateur. Ainsi, si nous souhaitons étudier la transmission de l'information dans un graphe d'ordinateurs, il peut être pertinent de considérer un graphe dans lequel les arêtes représentent une possibilité de connexion, plutôt qu'une connexion directe.

Dans un tel graphe, appelé un graphe *fermé transitivement*, deux sommets sont reliés par une arête si et seulement s'il existe un chemin entre ces deux sommets. Dans la figure ci-contre, le graphe de gauche correspond au graphe de départ, et le graphe de droite est le graphe fermé correspondant. Les arêtes sont pondérées pour indiquer la proximité entre deux sommets.



Nous avons observé que la structure communautaire d'un graphe quelconque, et celle de son graphe fermé est différente. L'opération de calcul de la fermeture d'un graphe est une opération coûteuse en temps. En partant d'un graphe quelconque, nous souhaitons étudier la possibilité (ou non) de trouver la structure communautaire du graphe fermé correspondant, sans calculer la fermeture transitive.

Partie état de l'art :

L'objectif de cette partie est de faire un état de l'art autour des algorithmes de détection de communautés, et des graphes fermés transitivement.

La méthode de Louvain [1] est une méthode de détection de communautés couramment utilisée. Quels autres algorithmes sont utilisés pour la détection de communautés ?

Quels algorithmes permettent de calculer la fermeture transitive d'un graphe ?

Pour finir, est-ce que la détection de communautés sur des graphes fermés a déjà été étudié ? Existe-t-il des algorithmes pour ce problème spécifique ?

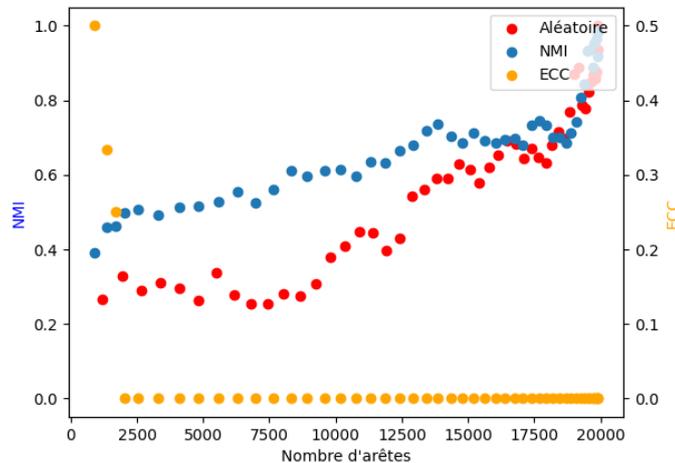
Partie réalisation pratique :

Une piste qui permettrait de trouver la structure communautaire du graphe fermé serait de calculer une fermeture partielle du graphe de départ, puis de faire une détection de communautés (à l'aide d'algorithmes existants) sur le graphe fermé partiellement plutôt que sur le graphe fermé. Ainsi, une partie du calcul (coûteux) permettant de fermer le graphe est évitée, au prix d'une baisse de la qualité des communautés.

Supposons que nous connaissions le graphe de départ, et les communautés du graphe fermé. Le but de la partie pratique est de trouver un ordre dans lequel fermer le graphe de départ : quelles arêtes ajouter dans la fermeture partielle.

La figure ci-dessous montre la qualité des communautés (appelé NMI, sur l'axe y, à gauche) en fonction du nombre d'arêtes dans la fermeture partielle, sur l'axe x. Les arêtes ont été ajoutées dans un ordre bien choisi. Nous observons qu'autour de 8000-9000 arêtes, la qualité des communautés est supérieure à la qualité des communautés quand les arêtes sont ajoutées dans un ordre aléatoire.

Je vous fournirai du code python permettant de générer cette figure. Le but sera de trouver un ordre plus performant, puis idéalement un critère permettant de décider quand arrêter d'ajouter des arêtes.



Bibliographie

1: Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E, Fast unfolding of communities in large networks, 2008