

# TP 2 : Igraph/R - Prévision de liens

Rushed Kanawati

## Résumé

L'objectif de ce TP est de se familiariser avec l'outil `igraph` : un outil d'analyse et de visualisation de graphes.

## Exercices

- 1 Télécharger sur le site <http://lipn.fr/~kanawati/datasets> les deux fichiers `dblp72-75.gml`, `dblp72-77.gml`. Les deux fichiers représentent l'évolution des relations de collaboration entre chercheurs dans la période 1975-1977.
- 2 Les deux graphes sont-ils connexes ? pour chaque graphe donner le nombre de composantes connexes et l'histogramme des tailles des composantes.  
NB. Penser à utiliser la fonction `clusters`
- 3 Développer une fonction qui renvoie le sous-graphe connexe maximale d'un graphe.
- 4 On considère maintenant seulement le sous-graphe connexe maximal du graphe `dblp72-75`. Donner les caractéristiques topologiques de ce graphe. Quel est le nombre de liens potentiels (qui peuvent se former) dans ce graphe ?  
NB. Penser à utiliser la fonction `graph.complementary`
- 5 Refaire la même question pour le graphe égo-centré sur le chercheur le plus central selon la centralité d'intermédiarité, puis selon la centralité de proximité.
- 6 En comparant avec les nouveaux liens formés entre les nœuds du graphe sélectionné mais dans la période 75-77 donner le ration de déséquilibre entre le nombre de liens positifs et les liens négatives pour chaque cas.
- 7 Évaluer les performances des mesures dyadiques suivantes pour la prévision de liens pour chaque graphe extrait précédemment : Jaccard, Adamic-Adar, Attachement préférentiel, longueur de plus court chemin, Katz-tronquée. Les performances sont à évaluer en termes de précision et de rappel.