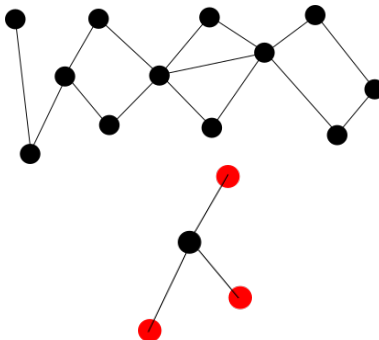


Plan

- 1 Ego-réseau et une analyse ego-centrée ?
- 2 Echantillonnage et ego-réseaux
- 3 Analyse globale pour des ego-réseaux (ou des boules de neige)
- 4 Analyse ego-centrée
 - Indicateurs de la position du noeud dans le réseau
 - Relier la position du noeud et caractéristiques du noeud
 - Prédiction de lien

D'un réseau à des réseaux ego-centrés

Définition : Le réseau ego-centré d'un noeud est le réseau formé par un noeud et ses voisins.



Qu'est-ce qu'une analyse ego-centrée ?

Dans des analyses ego-centrées, on s'intéresse typiquement à des indicateurs/ mesures sur les noeuds (ex : degrés) .

En calculant des indicateurs sur les noeuds, on peut ensuite les incorporer dans des méthodes d'analyse statistiques plus classiques.

Ex : Pour expliquer le nombre d'espèces cultivées dans une ferme, ajouter à des variables plus classique (ex : taille de la ferme) le degré de la ferme dans le réseau d'échanges.

Les analyses égo-centrées sont surtout intéressantes si les questions d'intérêt portent sur les noeuds.

Cet exposé

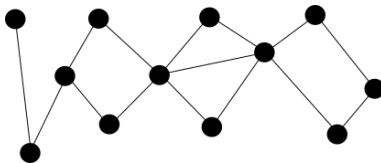
- ▶ Quelles analyses lorsque le réseau faiblement échantillonné de type ego-centré ou snow-ball ?
- ▶ Procédures d'analyse Ego-centrées

Plan

- 1 Ego-réseau et une analyse ego-centrée ?
- 2 Echantillonnage et ego-réseaux
- 3 Analyse globale pour des ego-réseaux (ou des boules de neige)
- 4 Analyse ego-centrée
 - Indicateurs de la position du noeud dans le réseau
 - Relier la position du noeud et caractéristiques du noeud
 - Prédiction de lien

Echantillonnage de réseaux

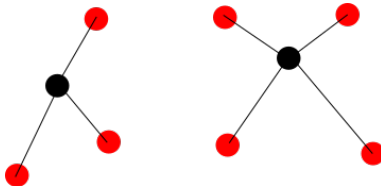
Le réseau est rarement complètement observé.



On s'intéresse ici à des données pour lesquelles l'effort d'échantillonnage est faible.

Echantillonnage d'ego-réseaux

Observations : Collection d'ego-réseaux



Deux types de noeuds

- ▶ Noeud explorés
- ▶ Noeud découverts mais non explorés (en rouge)

Difficulté : Si l'effort d'échantillonnage est faible (au regard de la taille du réseau),

Echantillonnage en boule de neige

Objectif : Mieux découvrir localement le réseau

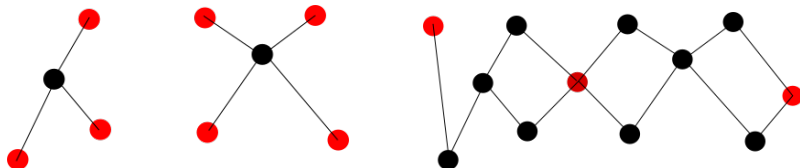


FIGURE – Etapes 1 et 2 de la boule de neige

Principe : Echantillonnage d'ego-réseaux puis exploration de certains noeuds découverts à la première étape,...

Plan

- 1 Ego-réseau et une analyse ego-centrée ?
- 2 Echantillonnage et ego-réseaux
- 3 Analyse globale pour des ego-réseaux (ou des boules de neige)
- 4 Analyse ego-centrée
 - Indicateurs de la position du noeud dans le réseau
 - Relier la position du noeud et caractéristiques du noeud
 - Prédiction de lien

Objectifs

- ▶ Quelles caractéristiques globales peut-on estimer avec des ego-reseaux ?
- ▶ Quelles caractéristiques globales peut-on estimer avec un échantillonnage boule de neige ?

Librairie utilisée : *igraph*

Distribution des degrés

Caractéristiques :

- ▶ Toujours possible à estimer.
- ▶ Attention à ne pas compter les degrés des noeuds non visités !
- ▶ (biais vers les hauts degrés pour l'échantillonnage boule de neige)

Homophilie

Homophilie ou assortativité : Propension d'individus semblables (ex : âges, liens de parenté,...) à être connectés.

Caractéristiques :

- ▶ Nécessite de connaître les caractéristiques (ex : âge, liens de parenté) des individus non visités.
- ▶ Sinon ... uniquement estimable à partir des couples d'individus visités.

Reciprocité

Pour un réseaux dirigé, propension de deux noeuds à avoir deux liens réciproque.

Caractéristiques :

- ▶ Nécessite de connaître les arêtes partant et arrivant des noeuds visités.
- ▶ Sinon ... uniquement estimable à partir des couples d'individus visités.

Transitivité

Propension de deux noeud reliés à un même noeud à être relié entre eux.

Caractéristiques :

- ▶ Privilégier au moins un échantillonnage boule de neige à 2 étapes.

Plan

- 1 Ego-réseau et une analyse ego-centrée ?
- 2 Echantillonnage et ego-réseaux
- 3 Analyse globale pour des ego-réseaux (ou des boules de neige)
- 4 Analyse ego-centrée
 - Indicateurs de la position du noeud dans le réseau
 - Relier la position du noeud et caractéristiques du noeud
 - Prédiction de lien

Principe général :

- 1 Calculer des indicateurs de la position du noeuds dans le du réseau
- 2 Comparer ces indicateurs à d'autres caractéristiques du noeuds

Différents indicateurs d'importance/centralité des noeuds

Sensible = nécessite un effort d'échantillonnage important !

Mesures de centralité

- 1 Degree [Peu sensible]
- 2 Eigen-centrality [Sensible !]
- 3 Closeness [Sensible !]
- 4 Betweenness [Très sensible ! !]

Lier centralité et caractéristiques du noeud

Comparer à d'autres caractéristique du noeud (âge, richesse), en utilisant des méthodes statistiques multivariées classique :

- ▶ Régression linéaire, régression logistique
- ▶ tests du χ^2 .
- ▶ Analyse en composantes principales, ...

Prédiction de lien

Il ne s'agit pas à proprement d'une 'ego-analyse'.

Objectif : Essayer de prédire/expliciter la présence de lien à partir de caractéristique des noeuds (ex : Age des protagonistes) ou du couple de noeud (ex : lien de parenté)

Outils : Régression logistique (ou linéaire)

- ▶ Eric D. Kolaczyk, Gábor Csárdi (auth.)- **Statistical Analysis of Network Data with R**-Springer-Verlag New York (2014)