

Apprendre à extraire de la connaissance de données d'interactions (graphes)

Objectifs et contenu

L'analyse statistique des réseaux d'interaction (ou graphes) a commencé par l'étude de petits réseaux (de l'ordre de la dizaine de noeuds) dans les années 60 en Sciences Sociales et connaît de nos jours un fort développement dans des domaines très variés (internet, biologie, réseaux sociaux, etc) avec des données de bien plus grande taille (quelques centaines, milliers voire millions de noeuds). Nous commencerons par présenter le graphes aléatoires qui sont l'objet mathématique sous-jacent et le stockage informatique des données (matrices d'adjacence ou listes de voisins). Nous présenterons ensuite les statistiques descriptives élémentaires des réseaux (composantes connexes, diamètre, degrés, comptages de motifs, coefficient de clustering, etc) ainsi que le problème de la visualisation des données. Enfin, nous aborderons le thème de la classification des noeuds (algorithmes spectraux, community detection, statistic block model, etc).

Intervenant(s)

Catherine Matias (CNRS, LPSM, Sorbonne Université).

Page personnelle : <http://cmatias.perso.math.cnrs.fr>

Page du laboratoire : <https://www.lpsm.paris>

Public visé

Non précisé.

Prérequis

Connaissance basique du logiciel R.

Durée et emploi du temps

3 jours.

La matinée est consacrée à la présentation des concepts et algorithmes tandis que l'après-midi est une mise en pratique sur machine (sous le logiciel R).

Infrastructure requise

Non précisée.

