Classification et visualisation de graphes avec SOMbrero

Madalina Olteanu*[1] and Nathalie Villa-Vialaneix[2]

[1]SAMM (Statistique, Analyse et Modélisation Multidisciplinaire), EA4543 – Université Paris 1 – France
[2]INRA, UR875, MIAT (MIAT) – Institut national de la recherche agronomique (INRA) : UR875 – F-31326 Castanet Tolosan, France

Résumé

Récemment, les données structurées et notamment les graphes ont connu un intérêt croissant. Celles-ci ont en effet de multiples applications en sciences humaines et sociales, en biologie ou en informatique. Pour comprendre les structures complexes modélisées par les graphes, une approche courante consiste à combiner classification des sommets du graphe avec visualisation. Les cartes auto-organisées (SOM) sont une classe d'algorithmes de classification non-supervisée dont le principe est de projeter de manière non linéaire les données sur une carte de dimension faible (généralement une grille rectangulaire à deux dimensions) tout en préservant la structure topologique des données. Dans cette proposition de communication, nous proposons de présenter le package R SOMbrero. Celui-ci contient une version stochastique de l'algorithme SOM adaptée à des données décrites par des matrices de dissimilarités, appelée SOM relationnel. En particulier, le package permet l'utilisation de l'algorithme pour des graphes à partir de (dis)similarités calculées entre sommets.

*Intervenant