

Using spatial indexes for labeled network analysis

Thibault Laurent ¹, Nathalie Villa-Vialaneix ^{2,3}

¹ Toulouse School of Economics,
Université Toulouse 1 - Manufacture des tabacs,
21 allées de Brienne, 31000 Toulouse - France
Thibault.Laurent@univ-tlse1.fr

² Institut de Mathématiques de Toulouse
Université Toulouse III (Paul Sabatier)
118 route de Narbonne, 31062 Toulouse cedex 9 - France
nathalie.villa@math.univ-toulouse.fr

³ IUT de Perpignan, Département STID
Domaine universitaire d'Auriac
Avenue du Dr Suzanne Noël, 11000 Carcassonne – France

Résumé

Un nombre croissant de données sont modélisées sous la forme d'un graphe, éventuellement pondéré : réseaux sociaux, réseaux biologiques...

Dans de nombreux exemples, ces données relationnelles peuvent être accompagnées d'une information supplémentaire, ou étiquette, sur les nœuds du graphe : il peut s'agir de l'appartenance à telle ou telle organisation pour un réseau social ou bien l'appartenance à une famille de protéines pour les réseaux d'interactions de protéines. Dans tous les cas, une question importante est de savoir si la distribution des valeurs de cette étiquette est influencée par la structure même du réseau. Nous proposons des outils d'exploration de cette question, basés sur des tests issus du domaine de la statistique spatiale. L'utilisation de ces tests est illustrée au travers de plusieurs exemples, tous issus du domaine des réseaux sociaux.

Mots-clés: données relationnelles, réseaux sociaux, I de Moran, statistiques de comptage, test de permutations, diagramme de Moran, sommets influents