

Editorial du numéro spécial "Méthodes statistiques pour les réseaux".

Title: Editorial of the Special Issue "Networks and Statistics"

Jean-Jacques Daudin¹

Les réseaux intéressent de plus en plus de statisticiens et le nombre d'articles sur le sujet croît très rapidement. Ce numéro spécial participe de cet essor. Il contient des articles de revue mais aussi des contributions nouvelles notamment dans le domaine émergent de l'analyse statistique des réseaux dynamiques.

L'article d'Antoine Channarond est basé sur les premiers chapitres de sa thèse. Il présente les éléments de base indispensables pour entrer dans le domaine. Il définit les notions de degré, transitivité, *coefficient de clustering* et décrit les propriétés probabilistes du modèle de Erdos-Renyi. Ce modèle suppose que les sommets sont indifférenciés. L'article définit également plusieurs modèles 'sommets hétérogènes, comme le modèle "petit-monde" de Kleinberg-Watts-Strogatz, celui de l'"attachement préférentiel" d'Albert et Barabasi, les modèles exponentiels et ceux 'variables latentes. On y trouve une vision large du domaine qui ne prétend pas être exhaustive. L'article est focalisé sur les modèles et leurs propriétés ce qui est déjà beaucoup vu le développement très rapide du domaine, l'inférence statistique est juste esquissée. Pour compléter le tour d'horizon, le lecteur intéressé peut aussi consulter [Kolaczyk and Csárdi \(2014\)](#), [Vu et al. \(2013\)](#), [Matias and Robin \(2014\)](#), [Daudin \(2011\)](#) et [Leger et al. \(2014\)](#).

Le regroupement en classes des sommets ayant le même comportement topologique est une voie d'analyse des graphes très utilisée actuellement. Elle a déjà été abordée dans l'article précédent et est traitée dans les deux articles suivants.

L'article de Madalina Olteanu et Nathalie Villa-Vialaneix propose d'utiliser un algorithme de type carte auto-organisatrice pour classifier les sommets d'un graphe. L'article décrit le paquet R qui réalise la classification et illustre la méthode avec deux exemples dont l'un est un graphe de réactions chimiques avec plusieurs milliers de sommets.

Un graphe bipartite est constitué de deux sous-ensembles de sommets. Les seuls liens possibles vont d'un sommet d'un sous-ensemble 'un sommet de l'autre sous-ensemble. Le co-clustering consiste 'construire simultanément des classes de sommets dans chaque sous-ensemble. L'article de Vincent Brault et Mahendra Mariadassou est une revue des méthodes et modèles utilisés dans ce but, avec une focalisation sur le modèle des Blocs Latents (LBM).

La plupart des travaux actuels sur l'analyse statistique de graphes concerne des graphes statiques (un seul graphe). De plus en plus de graphes dynamiques (plusieurs graphes indexés par le temps)

¹ AgroParisTech, 16 rue C. Bernard, 75005 Paris
E-mail : daudin@agroparistech.fr

sont maintenant disponibles. C'est un domaine émergent car il y a besoin de modèles prenant en compte ces nouvelles données. Ce thème est traité dans les deux derniers articles de ce numéro spécial.

L'article de Viviana Amati, Felix Schvnenberger et Tom Snijders définit le modèle dit "stochastic actor-oriented model" et donne une nouvelle méthode des moments généralisés pour en faire l'inférence. Ce modèle permet d'étudier l'évolution d'un réseau et permet d'intégrer des co-variables.

L'article de Rawyia Zreik, Pierre Latouche et Charles Bouveyron propose une extension du modèle Random Subgraph model (RSM) au cas dynamique. Dans le modèle RSM, le graphe est divisé en sous-graphes connus. Chaque sous-graphe est divisé en classes latentes de sommets. Les mêmes classes se retrouvent dans tous les sous-graphes, mais avec des proportions différentes. L'article illustre la méthode sur l'évolution du graphe des relations entre les salariés du groupe ENRON qui a fait faillite en 2001.

Les méthodes statistiques ont été appliquées depuis longtemps aux réseaux sociaux et génomiques. Nous avions prévu un article sur l'analyse des réseaux écologiques. Ce projet a été abandonné. À défaut, les lecteurs intéressés pourront consulter [Gauzens et al. \(2015\)](#), [Daudin et al. \(2010\)](#), [Miele et al. \(2014\)](#) ou [Leger et al. \(2015\)](#).

Je remercie les auteurs de leurs contributions et l'éditeur en chef pour son aide précieuse.

Références

- Daudin, J.-J. (2011). A review of statistical models for clustering networks with an application to a ppi network. *Journal de la Société Française de Statistique*, 152(2) :111–125.
- Daudin, J.-J., Pierre, L., and Vacher, C. (2010). Model for heterogeneous random networks using continuous latent variables and an application to a tree-fungus network. *Biometrics*, 66(4) :1043–1051.
- Gauzens, B., Thébault, E., Lacroix, G., and Legendre, S. (2015). Trophic groups and modules : two levels of group detection in food webs. *Journal of The Royal Society Interface*, 12(106).
- Kolaczyk, E. and Csárdi, G. (2014). *Statistical Models for Network Graphs*, volume 65 of *Use R!* Springer New York.
- Leger, J.-B., Daudin, J.-J., and Vacher, C. (2015). Clustering methods differ in their ability to detect patterns in ecological networks. *Methods in Ecology and Evolution*, 6(4) :474–481.
- Leger, J.-B., Vacher, C., and Daudin, J.-J. (2014). Detection of structurally homogeneous subsets in graphs. *Statistics and Computing*, 24(5) :675–692.
- Matias, C. and Robin, S. (2014). Modeling heterogeneity in random graphs through latent space models : a selective review. *ESAIM : Proc.*, 47 :55–74.
- Miele, V., Picard, F., and Dray, S. (2014). Spatially constrained clustering of ecological networks. *Methods in Ecology and Evolution*, 5(8) :771–779.
- Vu, D. Q., Hunter, D. R., and Schweinberger, M. (2013). Model-based clustering of large networks. *Ann. Appl. Stat.*, 7(2) :1010–1039.