

Modularisation et Recherche de Communautés dans les réseaux complexes par Unification Relationnelle

Patricia Conde-Céspedes*, Jean-François Marcotorchino**,

*Laboratoire de Statistique théorique et Appliquée 4 place Jussieu 75252 Paris cedex 05
patricia.conde_cespedes@upmc.fr

**Thales Communications et Sécurité, TCS, 4 Avenue des Louvresses, 92230 Gennevilliers
jeanfrancois.marcotorchino@thalesgroup.com,

Résumé. Un graphe étant un ensemble d'objets liés par une certaine relation typée, le problème de "modularisation" des grands graphes (qui revient à leur partitionnement en classes) peut, alors, être modélisé mathématiquement via l'Analyse Relationnelle. Cette modélisation permet de comparer sur les mêmes bases un certain nombre de critères de découpage de graphe c'est-à-dire de modularisation. Nous proposons une réécriture Relationnelle des critères de modularisation connus tels le critère de Newman-Girvan, le critère de Mancoridis-Gansner, le critère de Zahn-Condorcet, etc. Cette approche permet de faciliter leur compréhension, et d'interpréter plus clairement leurs finalités en y associant la preuve de leur utilité dans certains contextes pratiques.

1 Introduction à la recherche de communautés

De nos jours, l'étude des réseaux est devenue à la mode grâce essentiellement à l'utilisation massive des "Réseaux Sociaux" comme FaceBook, Twitter, LinkedIn, etc. Ces réseaux génèrent eux-mêmes des quantités considérables de données, ce qui contribue à accroître leur complexité. Cet afflux massif de données provient principalement d'Internet, des télécommunications et des technologies numériques en général. Toutes les disciplines, notamment la physique, la biologie et les sciences sociales ont été confrontées à cette arrivée d'énormes quantités de données à gérer. C'est donc la manipulation de cette masse de données latentes qui a rajouté un niveau supplémentaire de complexité à l'exploration de tels réseaux.

Les réseaux sont des objets très variés que l'on rencontre dans de nombreuses configurations, applicatives en particulier partout où l'on peut (ou doit) définir des connexions entre entités. Tous ces réseaux ont en commun de pouvoir être décrits par un graphe où les "sommets" représentent les différents éléments en interaction et les "arêtes" représentent l'influence d'un sommet sur un autre. Par exemple : en biologie moléculaire (l'interaction entre protéines), en informatique (les sites internet et les liens entre eux ou les réseaux client-serveur), en physique (les circuits électriques), etc.