

Méthodes spectrales pour la détection de groupes d'affinités dans les grands réseaux

Projet de conception N°

Xiaoyi MAI (A5-18)

Contexte du projet

La détection de communautés consiste, sur la base d'une observation des différentes interactions (liens) entre les noeuds du réseau, à grouper ces derniers en communautés de telle sorte qu'il y ait très peu d'interactions entre communautés et beaucoup d'interactions à l'intérieur des communautés.[1] Il est utilisé dans plusieurs applications aujourd'hui : réseaux sociaux (pour la création de groupes d'affinités entre les abonnés), sites de recommandation (recommander des produits aux clients qui ont les mêmes goûts), biologie, etc. Les algorithmes spectraux, une des méthodes d'approximation les plus puissantes et les plus utilisées, consistent à définir une matrice d'affinité du graphe (matrice d'adjacence, matrice de modularité, matrice Laplacienne) sur laquelle opérer, à extraire un certain nombre K de ses vecteurs propres correspondant à K valeurs propres isolées (valeurs propres dont la valeur absolue est relativement grande comparée à la majorité des valeurs propres) puis à faire de la classification non-supervisée des entrées de ces vecteurs propres en utilisant un algorithme classique de classification de points (k-means ou EM par exemple).

Description

Il s'agit dans ce projet de comparer les méthodes spectrales basées sur les matrices d'affinité classiques (matrice d'adjacence, matrice de modularité, matrice Laplacienne), qui ne marchent pas bien lorsque le réseau est sparse c'est-à-dire lorsque le degré moyen des noeuds reste constant lorsque la taille du réseau grandit et un nouveau algorithme basé sur la matrice appelée Bethe Hessian [3], qui marche beaucoup mieux dans ce cas. L'étude des algorithmes de détection de communautés se fait sur le "Stochastic Block model"[2], le modèle le plus connu et utilisé dans la littérature.

Étapes du projet :

1. Étude bibliographique sur les méthodes spectrales pour la détection de communautés dans les grands graphes.
2. Mise en œuvre des algorithmes sous Matlab.
3. Étude des performances.

Domaines concernés, mots-clefs

Détection de communautés, algorithmes spectraux, Stochastic Block model

Références

- [1] Santo Fortunato. Community detection in graphs. *Physics Reports*, 486(3) :75–174, 2010.
- [2] Raj Rao Nadakuditi and Mark EJ Newman. Graph spectra and the detectability of community structure in networks. *Physical review letters*, 108(18) :188701, 2012.
- [3] Alaa Saade, Florent Krzakala, and Lenka Zdeborová. Spectral clustering of graphs with the bethe hessian. In *Advances in Neural Information Processing Systems*, pages 406–414, 2014.