

Encadrant :

MAI Xiaoyi

MEJEAN Alexandre

MAISON Jonas

|  |
| --- |
| CentraleSupélec Campus de Gif, Année 2016-2017  Rapport projet de conception N°SS403 :  Méthodes spectrales pour la détection de groupes d’affinités dans les grands réseaux |

GitHub : <https://github.com/Jonas1312/CommunityDetection>

Sommaire

[Sommaire 2](#_Toc476925326)

[Introduction 3](#_Toc476925327)

[Génération de graphes aléatoires 4](#_Toc476925328)

[Algorithmes spectraux 4](#_Toc476925329)

[Analyse des résultats 4](#_Toc476925330)

[Conclusion 4](#_Toc476925331)

[Références 5](#_Toc476925332)

Introduction

Les graphes permettent de décrire une multitude de systèmes complexes issus de divers domaines : réseaux sociaux, moteurs de recherche, réseaux biologiques, circuits électroniques, jeux en réseaux et encore bien d’autres applications. On modélise en général ces graphes par un ensemble de nœuds, qui représentent les entités du réseau, et d’arêtes, qui représentent les relations entre ces entités. L’existence de communautés dans un réseau correspond à la présence de certains groupes de nœuds qui sont très fortement connectés entre eux par rapport aux autres nœuds du graphe. On appelle partitionnement de graphe la détection de communautés non chevauchantes.

Il existe actuellement dans la littérature une multitude de travaux portant sur la détection de communautés dans les graphes et ce domaine d’étude est en pleine expansion depuis la création des premiers réseaux sociaux. La plupart de ces méthodes consistent à déterminer une partition des sommets du graphe optimisant un certain critère de qualité d’un partitionnement. Ce problème d’optimisation est NP-complet, et il est donc difficile d’obtenir des résultats en un temps raisonnable sur des graphes de plusieurs milliers de nœuds.

Dans le cadre de notre projet nous avons abordé les algorithmes spectraux, qui sont aussi très utilisés dans la détection de communauté dans les graphes et qui prennent en compte les propriétés spectrales de ceux-ci. Le partitionnement spectral utilise les vecteurs et les valeurs propres d’une matrice d’affinité définie à partir d’un graphe et réalise ensuite de la classification non-supervisée pour trouver des partitions. Nous aborderons ici diverses matrices d’affinités mais plus particulièrement la matrice appelée Bethe Hessian, qui a la particularité de donner de meilleurs résultats sur les graphes sparses. Nous avons réalisé ces études sur un modèle de génération de graphe aléatoire, appelé le « stochastic block model », très connu et utilisé dans la littérature.

Nous aborderons donc dans un premier temps la génération de graphes aléatoires, puis nous présenterons les diverses matrices d’affinités que nous avons utilisé pour le partitionnement du graphe à l’aide des méthodes spectrales, et enfin nous étudierons les performances.

Génération de graphes aléatoires

Sbm, analyse des graphes, analyse des différentes matrices de proba, compléxité, stockage du graphe en mémoire, graphes dense ou sparse, notion de degré d’un graphe, utilisation de graphe non orientés, sans attributs

Algorithmes spectraux

Matrices utilisées, explication, histogramme valeurs propre, kmeans, bethe hessian, choix de « r » bethe hessian

Analyse des résultats

Conclusion

Références