

# Mémoire

Présenté par

*MERAZKA Mustapha*

Pour l'obtention du diplôme de Magister

Filière : Informatique

Option : Cloud Computing

Thème

---

## Détection de communautés dans les réseaux sociaux

---

Soutenu le :

Devant le jury composé de :

Mr TARI Abdelkamel	MCA	Université de Bejaïa	Président
Mr NOUALI Omar	DR	CERIST, Alger	Rapporteur
Mr BOUKERRAM Abdallah	Professeur	Université de Bejaïa	Examineur
Mr TANDJAQUI Djamel	MRA	CERIST, Alger	Examineur
Mme ALIANE Hassina	MRB	CERIST, Alger	Invitée

Année Universitaire : 2013/2014

*A la mémoire de mon père*  
*A toute ma famille...*

# Remerciements

Je remercie Dieu le tout puissant qui m'a donné le courage et la force pour aboutir à l'accomplissement de ce travail.

J'adresse tout d'abord mes remerciements à mes encadreurs, pour avoir accepté de m'encadrer et m'avoir aidé à conduire ce travail jusqu'au bout : Madame ALIANE Hassina, Maitre de recherche B au CERIST et Monsieur NOUALI Omar, Directeur de recherche au CERIST.

Je remercie Mr TARI Abdekamel, Mr BOUKERRAM Abdellah, et Mr TANDJAOUI Djamel, les membres du jury pour l'honneur qu'ils m'ont attribué en acceptant d'évaluer et de juger ce modeste travail.

Je remercie mes amis pour leur aide appréciable, leurs encouragements continus et leur soutien moral ininterrompu.

Ce travail n'aurait jamais été possible sans le soutien et l'appui moral des membres de ma famille. Je les remercie tous.

Enfin, pour être sûr de n'oublier personne, que tous ceux, qui de près ou loin, ont contribué par leurs conseils, leurs encouragements ou leurs amitiés, à l'aboutissement de ce modeste travail, trouvent ici l'expression de ma profonde reconnaissance.

## RESUME

La capacité d'analyser et de trouver les communautés présentes dans le réseau peut aider à comprendre et à visualiser la structure des réseaux. Des entités sociales interagissent de diverses manières en formant des groupes, envoyer des messages, partager des articles, ce joindre à un groupe de discussion etc. De nombreuses approches ont été proposées pour découvrir les structures de communautés dans les réseaux. Ces approches sont pour la plupart dédiées pour détecter les communautés disjointes. Dans ce mémoire, nous présentons une approche pour la détection de communautés chevauchantes dans les réseaux sociaux. Cette approche est basée sur la propagation de labels à travers un échange d'informations entre les nœuds du réseau sur l'appartenance aux communautés. Nous avons mis en place une évaluation de notre approche et nous avons présenté les résultats. Les résultats obtenus comparés à ceux de la littérature sont jugés satisfaisants.

**Mots clés:** réseaux sociaux, détection de communautés chevauchantes, propagation de labels, théorie de graphes

## ABSTRACT

The ability to analyze and find the communities in the network can help to understand and visualize the structure of networks. Social entities interact in various ways to form groups, send messages, share articles, join a discussion group etc. Many approaches have been proposed to discover community structure in networks. These approaches are mainly dedicated to detect communities disjoint. In this paper, we present an approach for the detection of overlapping communities in social networks. This approach is based on the propagation of labels through an exchange of information between the nodes of the network belonging to communities. We have implemented an evaluation of our approach and we presented the results. The results are considered satisfactory compared to those.

**Keywords:** social networks, overlapping community detection, label propagation, graph theory

# *Table des Matières*

Introduction générale.....	1
1. Contexte général.....	2
2. Problématique et objectif de l'étude.....	2
3. Contribution.....	3
4. Organisation du mémoire.....	3
Chapitre 1: Notions et définitions.....	4
1.1 Introduction.....	4
1.2 Modélisation par les graphes.....	4
1.3 Notions sur les réseaux complexes .....	5
1.4 Définitions sur les graphes.....	6
1.5 Les communautés et leurs propriétés.....	9
1.6 Structure de communautés .....	10
1.7 Nœud chevauchant ou non chevauchant.....	11
1.8 Mesures de modularités.....	11
1.8.1 Mesure de la modularité par Girvan Newman .....	12
1.8.2 Mesure de la modularité par Nepsuz et al.....	13
1.8.3 Mesure de la modularité par Chen et al .....	15
1.8.4 La modularité par Shen et al.....	15
1.8.5 Partition de densité par Ahn et al.....	16
1.8.6 La modularité par Lazar et al .....	17
1.9 Les communautés dans les réseaux sociaux.....	19
1.9.1 Le réseaux Zachary .....	19
1.9.2 Le réseau de collaboration entre scientifiques SFI .....	20
1.9.3 Le réseau de page Web .....	20
1.10 Détection de communautés.....	21
1.10.1 Définition formelle du problème .....	22
1.10.2 Intérêt de la détection de communautés et ses applications.....	23
1.11 Conclusion .....	24

Chapitre 2: La détection de communautés disjointes.....	25
2.1 Introduction.....	25
2.2 Les approches classiques .....	25
2.2.1 Partitionnement de graphes.....	26
2.2.2 Clustering hiérarchique.....	27
2.2.3 Clustering partitionnel.....	27
2.3 Les approches séparatives.....	28
2.3.1 L'approche de Girvan et Newman Edge Betweenness.....	29
2.3.1.1 Mesure de centralité d'intermédiarité.....	29
2.3.1.2 Description de l'algorithme de Girvan et Newman .....	30
2.4 Les approches agglomeratives.....	33
2.5 Les approches d'optimisation de la modularité.....	34
2.5.1 L'algorithme de Louvain.....	35
2.6 Les approches alternatives.....	36
2.6.1 Propagation de label LPA.....	36
2.7 Conclusion.....	40
Chapitre 3: La détection de communautés chevauchantes.....	41
3.1 Introduction.....	41
3.2 Percolation de cliques.....	41
3.3 Graphes de liens et partitionnement des liens.....	45
3.4 Extension locale et optimisation.....	46
3.5 Méthodes à base de propagation de labels.....	48
3.5.1 La méthode COPRA .....	48
3.5.2 SLPA .....	50
3.6 Conclusion.....	51
Chapitre 4: Contribution .....	52
4.1 Introduction .....	52
4.2 Schéma et description générale de l'approche proposée.....	52

4.3 La première phase.....	54
4.3.1 Opérateur de Propagation.....	54
4.3.2 Opérateur d'inflation.....	55
4.3.3 Opérateur Cutoff .....	56
4. 3.4 Opérateur mise à jour conditionnelle.....	56
4.3.5 Le critère de de stop.....	57
4.4 Deuxième phase.....	58
4.5 L'algorithme proposé.....	58
4.5.1 Un exemple illustratif.....	60
4. 5.3 La Complexité.....	65
4.6 Résultats et expérimentations.....	67
4.6.1 Le réseau <i>High school</i> .....	67
4.6.2 Le réseau <i>Karate club Zachary</i> .....	69
4. 6.3 Le réseau <i>Dolphin Social Network</i> .....	69
4.6.4 Le réseau <i>Les Misérables</i> .....	70
4.6.5 Le réseau <i>American College Football</i> .....	71
4. 6.6 Le réseau <i>Books about US Politics</i> .....	71
4.5.7 Le réseau <i>Political Blogs</i> .....	71
4.7 Densité de chevauchement.....	72
4.8 Evaluation de la méthode propose.....	74
4.8.1 Information mutuelle normalisée(NMI).....	74
4.9 Conclusion.....	77
Conclusion et perspectives.....	78
Conclusion .....	78
Perspectives.....	79
Références Bibliographiques.....	80