République Algérienne Démocratique et Populaire
Ministère de l'enseignement Supérieur et de la Recherche Scientifique
Université de Constantine 2
Faculté des Technologies de l'Information et de la Communication
Département d'Informatique Fondamental et ses Applications

Année : 2014
N° d'ordre : ..... / 2014
N° de série : ...../2014

Thèse

Pour l'obtention du diplôme de Docteur en 3éme cycle LMD

Spécialité : Informatique

Détection de Communautés et Identification de Leaders dans les Réseaux Sociaux de Confiance :

GRAPSY : Plateforme de Recommandation Collaborative

Nassira CHEKKAI

Soutenue le .../.../ ... devant le jury composé de :

Mohamed Batouche	*Professeur, Université Constantine 2*	Président
Salim Chikhi	*Professeur, Université Constantine 2*	Rapporteur
Hamid Seridi	*Professeur, Université de Guelma*	Examinateur
Mohamed Khiredine Kholladi	*Professeur, Université d'El Oued*	Examinateur
Abdelhafid Zitouni	*Maître de Conférences A,Université Constantine 2*	Examinateur
Hamamache Kheddouci	*Professeur, Université Lyon 1*	Invité

الممسوحة ضوئيا بـ CamScanner

Table des matières

Résumé I
Abstract II
Introduction générale 1

I : Concepts de Base et État de l'Art 6

Chapitre 1 : Réseaux Sociaux de Confiance et Systèmes de Recommandation Collaboratifs 7
1.Introduction 10
2.Aperçu sur les réseaux sociaux 10
2.1. Historique 10
2.2.Définition 11
2.3.Concepts fondamentaux 11
2.3.1.Structure sociale 11
2.3.2.Acteurs 11
2.3.3.Liens sociaux 12
2.3.4.Capital social 12
2.4.Caractéristiques des réseaux sociaux 12
2.4.1.Profil 12
2.4.2.Collectivité 12
2.4.3.Messagerie 12
2.4.4.Partage de fichiers 12
2.4.5.Gratuité et ouverture 13
2.5.Types de réseaux sociaux 13
3. Aperçu sur les réseaux sociaux à base de confiance 13
3.1.Définition 13
3.2.Concepts de base 14
3.2.1.Confiance 14
3.2.2.Crédibilité 14
3.3.Structure d'un réseau social de confiance 14
3.4.Modèles de collection des informations de la confiance sociale 15
3.4.1.Attitude 15
3.4.2.Comportement 15
3.4.3.Expérience 15
3.5.Niveaux de confiance 16
3.5.1.Confiance totale 16
3.5.2.Confiance partielle 16
3.5.3.Pas de confiance 16
3.6.Caractéristiques des réseaux sociaux de confiance 16
3.6.1.Contexte spécifique 16
3.6.2.Transitivité 16
3.6.3.Propagation 16
3.6.4.Co-citation 16
3.6.5.Dynamicité 16
3.6.6.Couplage 16

الممسوحة ضوئيا بـ CamScanner

3.6.7.Réciprocité 17
3.6.8.Utilité 17
4.Aperçu sur les systèmes de recommandation collaboratifs 17
4.1.Définition 17
4.2. Notions de base 17
4.2.1 .Profil 17
4.2.2.Communauté 18
4.2.3.Confiance 18
4.2.4.Similarité 18
4.3.Principe de la recommandation collaborative 19
4.3.1.Evaluation de recommandations 20
4.3.2.Formation de communautés 20
4.3.3.Production de recommandations 20
4.4.Mesures d'évaluation de la qualité d'un système de recommandation 20
4.4.1.Erreur absolue moyenne 21
4.4.2.Rappel 21
4.4.3.Précision 21
4.4.4.Temps de calcul 21
4.5.Défis de la recommandation collaborative 21
4.5.1.Problème de manque d'information 21
4.5.2.Problème de surcharge d'information 22
4.5.3.Problème de détection de communautés 22
4.5.4.Problème du Mouton Gris 22
4.5.5.Problème d'Intimité 22
4.5.6.Problème de démarrage à froid 22
5.Conclusion 23

Chapitre 2 : Analyse des Réseaux Sociaux 24

1.Introduction 26
2.Définition 26
3.Cycle de vie de l'analyse des réseaux sociaux 26
3.1. Fondements de l'approche 26
3.2. Évolution de la méthode 27
3.3.Période de développement 27
4.Notions de base en analyse des réseaux sociaux 27
5.Représentations de l'analyse des réseaux sociaux 28
5.1.Représentation graphique 28
5.1.1.Notions de graphes utilisées dans l'analyse des réseaux sociaux 28
5.1.2.Classes de graphes utilisées dans l'analyse des réseaux sociaux 29
5.1.3.Types de matrices 29
5.2.Représentation sémantique 30
5.2.1.Groupes sémantiques d'acteurs sociaux 31
6.Indicateurs de connectivité dans l'analyse des réseaux sociaux 31
6.1.Centralité 31
6.2.Densité 32
6.3.Fragmentation 33
6.4.Collage et pontage 33

الممسوحة ضوئيا بـ CamScanner

7.Branches de l'analyse des réseaux sociaux..... 33
7.1.Analyse de l'ensemble du réseau..... 34
7.2.Analyse de positions..... 34
7.3.Analyse de sous-groupes..... 34
8.Conclusion..... 35

Chapitre 3 : Détection de Communautés et Identification de Leaders..... 36

1.Introduction..... 38
2.Détection de communautés et identification de leaders dans les RS..... 38
2.1.Détection de communautés..... 38
2.1.1.Communauté..... 38
2.1.2.Détection de communautés..... 38
2.1.3.Méthodes de détection de communautés dans les réseau sociaux..... 39
2.1.4.Critères d'évaluation de la détection de communautés..... 42
2.1.5.Défis de la détection de communautés..... 42
2.2.Identification de leaders..... 42
2.2.1.Leaders..... 42
2.2.2.Détection de leaders..... 42
2.2.3.Méthodes de détection de leaders..... 43
2.2.4.Mesures évaluant l'efficacité de l'identification de leaders..... 45
2.2.5.Défis de l'identification de leaders..... 45
3.Détection de communautés et de identification de leaders dans les systèmes de recommandation collaboratifs..... 45
3.1.Classification des techniques de résolution du problème de démarrage à froid..... 46
3.2.Principe de recommandation basée « Leader »..... 47
3.3.Classification des techniques de recommandation basée « Leader »..... 48
3.3.1.Techniques sans-Graphes..... 48
3.3.2.Techniques basées « Graphes »..... 49
4.Conclusion..... 52

Chapitre 4 : Paramètres de Graphes..... 53

1.Introduction..... 55
2.Aperçu sur les Paramètres de Graphes..... 55
2.1.Historique..... 55
2.2.Définition..... 56
2.3.Domaines d'application..... 57
2.4.Principaux paramètres de graphes..... 57
2.4.1. Nœuds Critiques et Arêtes Critiques..... 57
2.4.2.Nombre Chromatique et Indice Chromatique..... 57
2.4.3. Longueur du Plus Long Cycle..... 58
2.4.4.Ensemble de Sommets de Feedback..... 58
2.4.5.Ensemble Minimal de Sommets Dominants..... 59
2.4.6.Couverture Minimum de Sommets..... 59
3.Paramètre de Nœuds Critiques..... 59
3.1.Définitions..... 60
3.2.Notation..... 60

الممسوحة ضوئيا بـ CamScanner

3.3.Applications du paramètre de nœuds critiques 61
3.3.1.L'analyse des réseaux sociaux 61
3.3.2.Réseaux de vaccination 62
3.3.3.Brouillage des réseaux 62
3.3.4.Réseaux de mafia 62
3.3.5.Ingénierie des transports 62
3.4.Variantes du paramètre de nœuds critiques 62
3.4.1.Première variante : Nombre de sommets à supprimer 63
3.4.2.Seconde variante: Taille de composantes connexes 63
3.4.3.Troisième variante : Mesure d'efficacité 64
3.4.4.Quatrième variante : Points d'articulation 64
4.Paramètre du Stable Maximal 65
4.1.Définition 65
4.2.Notation 66
4.3.Domaines d'application 66
4.3.1.Communications radios 66
4.3.2.Réseaux filaires 66
4.3.3.Détection d'intrusions 66
4.3.4.Gestion de commandes 66
4.4.Paramètres connexes 66
4.4.1.Noeuds critiques 67
4.4.2.Clique maximale 67
4.4.3.Couverture de sommets 67
5.Conclusion 67

II : Contributions 68

Chapitre 5 : Méthodes à base de Nœuds Critiques pour la Détection de Communautés et l'Identification de Leaders 69

1.Introduction 71
2.Motivation 71
3.Méthodes proposées 72
3.1.Première méthode : Stable maximal 72
3.1.1.Enchainement de la méthode proposée 73
3.1.2.Algorithme global proposé 77
3.2.Seconde méthode : Les points d'articulation 78
3.2.1.Algorithme global proposé 79
3.3.Troisième méthode : Efficacité 80
3.3.1.Algorithme global proposé 82
4.Conclusion 83

Chapitre 6 : Validation des Méthodes à base de Nœuds Critiques pour la Détection de Communautés et l'Identification de Leaders 85

1.Introduction 87
2. Résultats expérimentaux 87

الممسوحة ضوئيا بـ CamScanner

2.1.L'exemple typique de "Trust Recommendation" 87
2.2.Réseau social de "Les-Misérables" 88
2.3.Réseau social de "Zachary's Karate Club" 89
2.4.Graphes générés aléatoirement 91
3.Discussion 92
4.Conclusion 93

Chapitre 7 : GRAPSY : Plateforme à base de Leaders Sociaux de Confiance pour la Recommandation Collaborative de Papiers de Recherche 94

1.Introduction 96
2.Motivation 96
3.Description de notre plateforme GRAPSY 97
4.Fonctionnement de GRAPSY 98
4.1.Phase centralisée 98
4.2.Phase distribuée 99
5.Composants de GRAPSY 99
5.1.Module d'Utilisateur 100
5.1.1.Sous-module de Crédibilité 100
5.1.2.Sous-module de Communication 101
5.1.3.Sous-module de Confiance 101
5.1.4.Sous-module d'Adaptation 103
5.2. Module de Document 103
5.3.Module de Recherche de documents 103
6.Approche basée leaders de confiances pour la résolution de problèmes de démarrage à froid d'un nouvel utilisateur et un nouvel article 104
6.1.Modélisation graphique de l'approche proposée 104
6.1.1.Types de nœuds 105
6.1.2.Types d'arêtes 106
6.2.Calcul de la crédibilité d'un utilisateur 106
6.2.1.Connexion au système 106
6.2.2.Ancienneté dans le système 107
6.2.3.Nombre de documents notés et proposés dans le système 107
6.2.4.Expérience de l'utilisateur 108
6.2.5.Personnalité de l'utilisateur 108
6.3.Détection de communautés et identification de leaders de confiance 109
6.3.1.Détection de communautés et de connecteurs 109
6.3.2.Identification de délégués de communautés 110
6.4.Évolution du graphe 111
7.Conclusion 112

Chapitre 8 : Développement et Validation de GRAPSY 114

1.Introduction 116
2.Simulation de GRAPSY avec la technique Multi-Agents 116
3.Résultats et discussion 120
4.Conclusion 125
Conclusion et Perspectives 126

الممسوحة ضوئيا بـ CamScanner

Résumé

À l'heure actuelle, les réseaux sociaux de confiance sont devenus un fait social primordial. L'objectif majeur de ces réseaux est de regrouper des personnes partageant des intérêts communs en se basant sur la confiance mutuelle entre eux. L'analyse des réseaux sociaux, basée principalement sur la théorie des graphes, consiste à étudier diverses facettes de ces réseaux, dont les principales sont la détection de communautés et l'identification de leaders.

Dans cette thèse, et dans le cadre d'une première contribution, nous présentons trois méthodes visant à détecter les communautés d'intérêt et identifier les leaders dans les réseaux sociaux de confiance. Pour atteindre ces objectifs, nous modélisons ces réseaux par des graphes pondérés et nous utilisons le paramètre de graphes "nœuds critiques" avec la notion de "composantes connexes" associée aux "scores de confiance". Ainsi, nos deux premières méthodes détectent l'ensemble des nœuds critiques les plus influents à travers les composantes connexes fortement liées et les scores de confiance élevés. La première spécifie les conditions sous lesquelles les nœuds critiques doivent être contrôlés en utilisant le paramètre du stable maximal, et la seconde permet d'identifier l'ensemble de points d'articulations qui sont en mesure de jouer le rôle des nœuds critiques. Quand à la troisième méthode, elle s'appuie sur une mesure physique dite "l'efficacité du réseau" et les scores de confiance dominants; dans le but d'atteindre l'ensemble des nœuds critiques influents.

Comme seconde contribution de ce travail de recherche, nous nous sommes intéressés aux systèmes de recommandation collaboratifs en tant que cas d'étude. Ceux-ci s'inscrivent dans le cadre des systèmes communautaires et des réseaux sociaux de confiance. Dans cette optique, nous présentons une nouvelle plateforme "GRAPSY" pour la recommandation collaborative des papiers de recherche. GRAPSY vise à atténuer les problèmes de démarrage à froid en couplant plusieurs disciplines. D'une part, elle exploite des techniques avancées de graphes qui font partie de l'informatique et les mathématiques. D'autre part, et afin de déterminer le degré de sociabilité de l'utilisateur, elle introduit un test psychométrique de sociabilité relevant du domaine de la psychologie et la sociologie.

Mots clés

Réseau sociaux de confiance, Détection de communautés, Identification de leaders, Graphes pondérés, Paramètres de graphes, Nœuds critiques, Composantes connexes, Scores de confiance, Stable maximal, Points d'articulation, Mesure d'efficacité, Systèmes de recommandation collaboratif, Problème de démarrage à froid, Test de sociabilité.

الممسوحة ضوئيا بـ CamScanner

Abstract

Nowadays, trust social networks have become a primordial social interest. The main objective of these networks is to bring together people who share common interests based on the mutual trust between them. The social networks analysis, mainly founded on graph theory, studies various aspects of these networks such as the communities' detection and the leaders' identification.

In this thesis, we present, in a first contribution, three methods aiming at detecting the communities of interest and identifying the influential actors in trust social networks. To this end, we model these networks by weighted graphs and we use the graph parameter "critical nodes" with the notions of "connected components" and "confidence scores". Thus, our first two methods detect the sets of most influential critical nodes through the connected components which are strongly linked and the greatest trust scores. Indeed, the first method specifies conditions under which critical nodes should be controlled using maximal stable parameter, while the second extracts the sets of the most powerful articulation points which are able to play critical nodes' role. Furthermore, our third method identifies the sets of most influential critical nodes using a physical measure "the network efficiency" and dominating trust scores.

In the second contribution of this thesis, we were interested in collaborative recommender systems as a case study. These systems are a part of community systems and trust social networks. In this context, we present a new platform "GRAPSY" for collaborative recommendation of research papers. GRAPSY solves the cold start problems by coupling different disciplines. On one hand, it uses advanced techniques of graphs that are part of computer and mathematics sciences. On the other hand, it employs a psychometric questionnaire of sociability, which is a sociological psychological test, in order to specify the sociability degree of a user.

Key words

Trust social networks, Communities' detection, Leaders' identification, Weighted graphs, Graph parameters, Critical nodes, Connected Components, Trust scores, Maximal stable, Articulation Points, Efficiency measure, Collaborative recommender systems, Cold start problem, Sociability test.

الممسوحة ضوئيا بـ CamScanner

ملخص

في وقتنا الحاضر، غدت الشبكات الاجتماعية الموثوقة حدثا اجتماعيا جوهريا لأن هدفها الرئيسي هو تجميع الأفراد الذين تربطهم اهتمامات مشتركة مبنية على الثقة المتبادلة. تحليل الشبكات الإجتماعية ـ القائم أساسا على نظرية البيانات ـ يهدف لدراسة جوانب مختلفة لهذه الشبكات، مثل إشكاليتي استخراج جماعات الاهتمام وتحديد الأفراد المؤثرين.

في هذه الأطروحة، وفي إطار مساهمتنا الأولى، نعرض ثلاثة مناهج لتحديد جماعات الاهتمام والأفراد المؤثرين داخل الشبكات الاجتماعية الموثوقة. من أجل ذلك، نمثل هذه الشبكات ببيانات موزونة مستعملين "عامل العقد النقدية" مع مفهوم "المجتمعات الموصولة" المرتبط "بمؤشرات الثقة". منهجانا الأولان يسمحان باستخراج مجموعة العقد النقدية الأكثر تأثيرا بواسطة المكونات الموصولة الأشد ارتباطا ومؤشرات الثقة الأكبر قيمة. في هذا السياق، يسمح المنهج الأول بتحديد شروط مراقبة العقد النقدية مستندا إلى عامل الثابت الأقصى، ويهدف المنهج الثاني إلى تحديد مجموعة النقاط المفصلية النافذة والقادرة على تمثيل العقد النقدية الناجعة. بينما يهدف المنهج الثالث إلى بلوغ العقد النقدية المؤثرة عن طريق استعمال مقياس فيزيائي يسمى "فعالية الشبكة" ومؤشرات الثقة المهيمنة.

في مساهمتنا الثانية، اهتممنا بأنظمة التوصية التعاونية كدراسة حالة. هذه الأخيرة تشكل جزءا لا يتجزأ من أنظمة الجماعات و الأنظمة الاجتماعية الموثوقة. في هذا الإطار، نقدم نظام GRAPSY المختص في التوصية التعاونية لوثاق البحث العلمي والرامي إلى حل مشاكل الانطلاق البارد معتمدا على مزج ميادين عديدة. فالنظام يوظف من جهة تقنيات بيانية متقدمة ذات صلة بميدان الإعلام الآلي والرياضيات، ويستعمل من جهة أخرى اختبارا اجتماعيا بسيكولوجيا لمعرفة درجة الانبساط لدى المستخدم.

كلمات مفتاحية

الشبكات الاجتماعية الموثوقة، استخراج الجماعات، تحديد القادة، البيانات الموزونة، عوامل البيان، العقد النقدية، المركبات الموصولة، مؤشرات الثقة، الثابت الأقصى، النقاط المفصلية، مقياس الفعالية، أنظمة التوصية التعاونية، مشكل الانطلاق البارد ،الإختبار الاجتماعي.

الممسوحة ضوئيا بـ CamScanner