

TMD-MINER : Une nouvelle approche pour la détection des diffuseurs dans un système communautaire

Mohamed Nidhal Jelassi^{*,**}, Christine Largeron^{*}, Sadok Ben Yahia^{**}

^{*} Laboratoire Hubert CURIEN, Université Jean Monnet, Saint-Etienne, France.
Christine.Largeron@univ-st-etienne.fr,

^{**} Faculté des Sciences de Tunis, Université de Tunis-El Manar, Tunis, Tunisie.
Sadok.Benyahia@fst.rnu.tn

Résumé. Plusieurs méthodes ont été développées ces dernières années pour détecter, dans un réseau social, les membres qualifiés, selon les auteurs, d'influenceurs, de médiateurs, d'ambassadeurs ou encore d'experts. Dans cet article, nous proposons un nouveau cadre méthodologique permettant d'identifier des diffuseurs dans le contexte où seule l'information sur l'appartenance des membres du réseau à des communautés est disponible. Ce cadre, basé sur une représentation du réseau sous forme d'hypergraphe, nous a permis de formaliser la notion de diffuseur et d'introduire l'algorithme TMD-MINER, dédié à la détection des diffuseurs et basé sur les itemsets essentiels.

1 Introduction

C'est en s'appuyant sur des représentations et des concepts issus de la théorie des graphes que les réseaux sociaux ont été étudiés en sciences sociales dès les années soixante (Moreno (1934), Cartwright et Harary (1977)). Parmi les questions essentielles que l'analyse de réseau s'efforce de traiter figure l'identification d'individus occupant un rôle déterminant dans le réseau. Avec l'explosion des réseaux sociaux sur Internet, des travaux plus récents se sont attachés à repérer des acteurs qualifiés selon les auteurs d'influenceurs, de médiateurs, d'ambassadeurs ou encore d'experts (Opsahl et Hogan (2010); Scripps et al. (2007a,b); Agarwal et al. (2008)).

En particulier, plusieurs algorithmes ont été présentés récemment pour résoudre le problème de recherche d'influenceurs, redéfini comme un problème de maximisation (Richardson et Domingos (2002); Kempe et al. (2003)). Cependant, ces algorithmes exploitent la matrice d'adjacence associée au graphe décrivant le réseau ; ce qui suppose connues les relations entre les acteurs pris deux à deux. Or, dans de nombreuses applications, cette information n'est pas forcément accessible. Par contre, on sait à quelle(s) communauté(s) appartient un acteur. On cherche alors à identifier des acteurs, appelés diffuseurs, en nombre le plus limité possible, qui appartiennent à plusieurs groupes et qui du fait de cette position sont susceptibles d'assurer des échanges d'un groupe à un autre. Il peut s'agir par exemple des chercheurs qui ont participé à la conférence EGC et ceux qui ont assisté à Coria ou encore des clients ayant acheté différentes