

Un nouveau modèle pour l'évolution de réseaux sociaux personnels

(A new evolution model for personal social networks)

Sarra Djemili, Claudia Marinica^[0000-0003-2388-9095], Maria Malek^[0000-0002-7282-8021], and Dimitris Kotzinos^[0000-0002-3678-4092]

ETIS UMR8051, CY Cergy Paris University, ENSEA, CNRS, F-95000, Cergy, France
sarahdjemili@yahoo.fr,
{claudia.marinica, maria.malek, dimitrios.kotzinos}@cyu.fr

Abstract. Online personal networks (OPNs) are social networks composed of one focal individual, called ego, and several other individuals, called alters, connected to the ego directly or indirectly; the alters are positioned at a maximum distance k from the ego. Moreover, OPNs evolve over time, and understanding their evolution would allow to understand how the connection between the individuals change over time. For online social networks, several evolution models were proposed in the literature, such as preferential attachment and probabilistic models, or others integrating Markov models or spectral methods. But one important question that needs to be addressed is the suitability of these evolution models for online personal networks which carry different characteristics both structural and behavioral than social networks. This paper proposes to explore the structural features of OPNs and the influence on their evolution, in order to develop a new evolution model based on the clique structure of the personal collaboration networks.

1 Introduction

Les réseaux personnels en ligne (OPNs) sont des réseaux sociaux en ligne (OSNs) composés d'un individu central (l'égo) et de plusieurs individus (les alters) avec lesquels l'égo est connecté de manière directe ou indirecte [3]; les alters se trouvent à une distance maximale k de l'égo (on parle de réseaux de niveau k). Les réseaux personnels évoluent dans le temps et l'évolution d'un réseau est définie comme une séquence de réseaux statiques dans laquelle chaque réseau statique correspond à l'état du réseau à un temps donné. Lors de son évolution d'un temps t à un temps $t + 1$, le réseau se voit subir un certain nombre de modifications : ajout/suppression de nœuds, ajout/suppression de liens, modification de poids, etc. La compréhension de l'évolution des réseaux personnels permettrait aux sociologues de comprendre comment les liens entre les individus changent dans le temps dans divers domaines.

Dans le cadre des OSNs, un grand nombre de modèles d'évolution ont été proposés. Parmi les premiers, le modèle d'évolution basé sur l'attachement préférentiel [2] considère qu'un nouveau nœud arrivé dans un réseau aura plus de probabilité à se connecter à un nœud déjà fortement connecté. Une grande partie des modèles d'évolution proposés par la suite a été basée sur ce premier modèle; néanmoins, des modèles originaux ont été proposés comme l'utilisation des chaînes de Markov pour modéliser la dynamique, ou, un modèle plus récent qui utilise la structuration en cliques des réseaux pour estimer leur évolution [7].

Par contre, étant donné que l'étude des OPNs est très récente, il n'est toujours pas clair si leur évolution est comparable à celle des réseaux sociaux dans leur globalité, et donc si les modèles d'évolution des OSNs sont applicables aux OPNs. Par exemple, les auteurs de [1] ont étudié l'évolution de la taille et de la structure d'un OPN limité aux connexions directes entre l'égo et les alters et ils ont conclu que le nombre de relations actives de l'égo est constant à cause de la capacité cognitive limitée d'un individu. De plus, ce nombre correspond au nombre de Dunbar [5] égal à 150.

Dans ce travail, le domaine d'application correspond aux réseaux de collaboration de la base de données de DBLP¹. Les réseaux de collaboration ont été

¹ <https://aminer.org/citation>

étudiés sous différents aspects : tout d’abord, Newman [6] a étudié la structure des réseaux (nombre moyen de papiers par auteur, etc.) sans tenir compte de leur dynamique; plus tard, l’évolution de ces réseaux a été étudiée [2] et les auteurs ont conclu que l’évolution suit le processus d’attachement préférentiel, et que la moyenne de chemins les plus courts diminue dans le temps.

Dans notre papier [4], nous avons étudié l’évolution des réseaux personnels de collaboration en analysant l’évolution dans le temps des valeurs d’un ensemble de métriques. Cette étude nous a permis de comprendre que l’évolution des réseaux personnels de collaboration ne suivent pas les modèles d’évolution basés sur l’attachement préférentiel car le degré des nœuds ne suit pas une loi de puissance, et le chemin le plus court en moyenne est nettement plus important que 6, qui est la valeur retrouvée dans la littérature. De plus, nous avons validé le fait que les réseaux personnels étudiés ont la forme d’un graphe particulier, nommé *graphe de type moulin*, qui a la particularité d’être très clusterisé en local et beaucoup moins clusterisé globalement. Cette observation nous a permis d’étudier de plus près la structuration des réseaux en cliques maximales et nous avons conclu à la fois que les réseaux étudiés sont composés d’un nombre important de cliques maximales et que les nouveaux nœuds arrivés dans un réseau ont plus la tendance de se connecter à des sous-cliques qu’à des cliques maximales.

Ces analyses nous confortent dans la nécessité de proposer un nouveau modèle d’évolution adapté aux réseaux personnels de collaboration. Ainsi, nous proposons dans ce papier une version préliminaire d’un nouveau modèle d’évolution prenant en compte la structuration en cliques du réseau personnel de collaboration. Le modèle développé est une extension du modèle génératif basé sur la superposition de cliques proposé dans [7].

2 Le modèle d’évolution basé sur la superposition de cliques pour les réseaux personnels de collaboration

Le modèle d’évolution proposé par [7] est un modèle génératif de réseaux sociaux qui prend en compte leur organisation en cliques et sous-cliques, et il est basé sur le concept de superposition de cliques.

2.1 Le concept de superposition de cliques

Pour illustrer la superposition de cliques, nous allons considérer que, au temps t , nous avons une clique et, au temps $t + 1$, un nouveau nœud arrive dans la clique; alors le nouveau nœud se connectera à tous les membres de la clique et ensemble ils vont former une clique plus grande. De plus, si nous supposons que, au temps t , le nouveau nœud fait déjà partie d’autres cliques par rapport à son activité passée, les cliques se superposent dans ce nouveau nœud.

Par exemple, dans la figure 1, au temps t nous avons les cliques $C1$ et $C2$ (gauche). Au temps $t + 1$ (droite), le nœud 3 (de la clique $C2$) rejoint la clique $C1$, et donc il se connectera à tous les nœuds de la clique $C1$, ainsi les cliques $C1$ et $C2$ sont superposées dans le nœud 3.

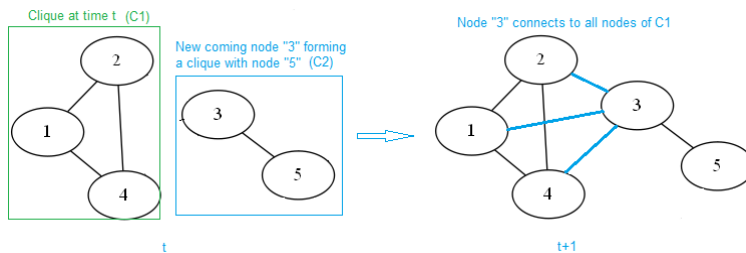


Fig. 1: Un exemple du concept de superposition de cliques.

2.2 Résultats sur la prédiction de réseaux personnels de collaboration

Pour analyser les résultats de prédiction sur les réseaux personnels de collaboration de DBLP suite à l'application de du modèle d'évolution basé sur la superposition de cliques, nous avons effectué un pré-traitement sur les données existantes et nous travaillons actuellement sur les auteurs ayant publié pour la première fois en 2004, formant la base de données $dataset_1$ (560 auteurs). Nous allons montrer dans la suite le résultat de prédiction pour un des réseaux. Comme indiqué dans l'introduction, nous nous intéressons ici à la prédiction d'un réseau au temps $t + 1$, en connaissant le réseau au temps t et le nombre de nœuds que le réseau devrait avoir au temps $t + 1$.

Dans la figure 2a est présenté le réseau personnel de collaboration de l'égo portant l'identifiant 2023 en 2005 qui est composé de 10 nœuds et 13 liens; la valeur du $k = 2$. La figure 2b présente le réseau du même égo en 2006 qui est composé de 28 nœuds et 47 liens. Cela indique que le modèle d'évolution devra ajouter au réseau initial 18 nœuds et 34 liens.

Le modèle d'évolution basé sur la superposition de cliques, prenant en entrée le réseau personnel en figure 2a, prédit le réseau présenté en figure 3. Nous pouvons remarquer plusieurs points divergents entre le réseau réel et le réseau prédit. Ces différences sont liées à la présence de nœuds isolés, à la profondeur du réseau personnel prédit, aux paramètres du modèle et ainsi qu'à la performance de certains algorithmes faisant partie du modèle.

3 Conclusion

Cet article présente une version préliminaire d'un nouveau modèle pour l'évolution des réseaux personnels de collaboration. Ce nouveau modèle est basé sur les analyses qu'on a pu réaliser précédemment, mais également sur un modèle déjà existant.

References

1. Arnaboldi, V., Conti, M., Passarella, A., Dunbar, R.: Dynamics of personal social relationships in online social networks: A study on twitter. In: Proceedings of the

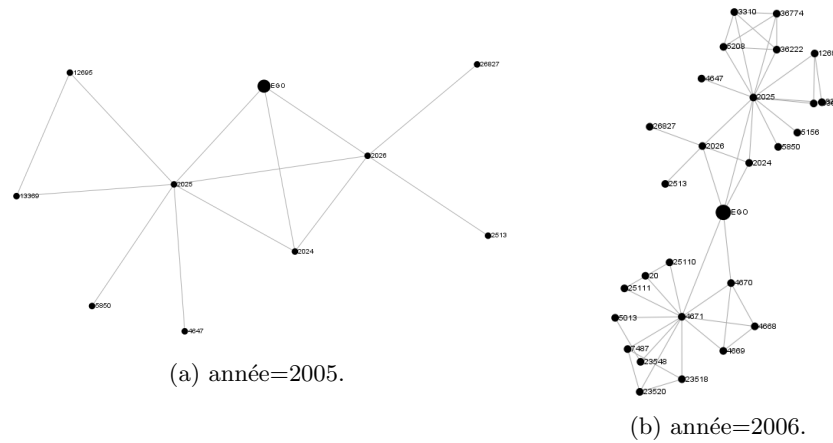


Fig. 2: Réseau personnel de collaboration du nœud $EGO = 2023$ avec $k = 2$ de l'année 2005 (gauche) et de l'année 2006 (droite).

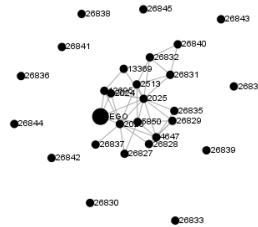


Fig. 3: Réseau prédit par le modèle basé sur la superposition de cliques.

First ACM Conference on Online Social Networks. pp. 15–26. COSN '13, ACM, New York, NY, USA (2013)

2. Barabasi, A., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., Vicsek, T.: Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications* **311**(3-4), 590–614 (Aug 2002)
3. Djemili, S., Marinica, C., Malek, M., Kotzinos, D.: A Definitions' Framework for Personal/Egocentric Online Social Networks. In: *MARAMI'16*. Cergy-Pontoise, France (Oct 2016)
4. Djemili, S., Marinica, C., Malek, M., Kotzinos, D.: Personal networks of scientific collaborators: A large scale experimental analysis of their evolution. In: *Information Search, Integration, and Personalization - 11th International Workshop, ISIP 2016*, Lyon, France, November 1-4, 2016, Revised Selected Papers. pp. 116–139 (2016)
5. Dunbar, R.I.: Neocortex size as a constraint on group size in primates. *Journal of human evolution* **22**(6), 469–493 (1992)
6. Newman, M.: The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 404–9 (02 2001)
7. Yan, F., Cai, S., Zhang, M., Liu, G., Deng, Z.: A clique-superposition model for social networks. *Science China Information Sciences* **56**(5), 1–19 (2013)