

Comparaison de la capacité des plongements de graphes à capturer les propriétés des réseaux

Rémi Vaudaine*, Rémy Cazabet**, Christine Largeron*

*Univ Lyon, UJM-Saint-Etienne, CNRS, Institut d'Optique Graduate School
Laboratoire Hubert Curien UMR 5516, F-42023, SAINT-ETIENNE, France
{remi.vaudaine,christine.largeron}@univ-st-etienne.fr

**Univ Lyon, UCBL, CNRS, LIRIS UMR 5205, F-69621, Lyon, France
remy.cazabet@gmail.com

1 Motivations

Les graphes sont utiles pour modéliser des systèmes complexes dans un large éventail de domaines. Parmi les approches conçues pour les étudier, le plongement de graphes consiste à encoder les noeuds, des parties du graphe ou un graphe entier dans un espace de faible dimension tout en préservant certaines propriétés structurelles. Puisqu'il permet à toute la gamme de techniques d'exploration de données et d'apprentissage automatique qui nécessitent des vecteurs d'être appliquée aux données relationnelles, il peut profiter à de nombreuses applications.

Plusieurs articles de l'état de l'art consacrés au plongement de graphes récemment publiés Cui et al. (2019); Goyal et Ferrara (2017); Cai et al. (2018); Zhang et al. (2020); Wu et al. (2020), comportent une étude comparative des performances des méthodes à résoudre des tâches spécifiques. Or, nous considérons qu'une telle évaluation des modèles de plongement de graphes guidée par une tâche (prédiction de liens / classement / détection de communautés) induit un biais lié au choix de l'algorithme employé pour traiter la tâche. Si les performances obtenues sur la tâche sont mauvaises cela est-il dû à l'algorithme choisi pour traiter la tâche ou à l'algorithme d'embedding? En réalité, l'évaluation guidée par la tâche ignore le fait que les algorithmes de plongement sont conçus, explicitement ou implicitement, pour préserver certaines propriétés structurelles particulières. Leur utilité pour une tâche donnée dépend donc de la manière dont ils réussissent à capturer ces propriétés. Par exemple, dans le cas de l'attachement préférentiel, un classifieur aura du mal à prédire de nouveaux liens si les représentations de deux nœuds voisins ne sont pas proches dans l'espace de plongement. Dans cet article, nous adoptons donc un point de vue différent des évaluations précédentes axées sur les tâches en étudiant directement, à travers une étude comparative expérimentale, la capacité des algorithmes de plongement de graphes à capturer des propriétés spécifiques, qui sont la proximité de premier ordre des nœuds, l'équivalence structurelle (proximité de second ordre) ou l'équivalence isomorphe. Pour ce faire, nous proposons différentes mesures permettant d'évaluer dans quelle mesure les méthodes de plongement les encodent.