

Apprentissage Joint de la Représentation et du Clustering avec un Réseau Convolutif sur Graphe

Chakib Fettal^{*,**}, Lazhar Labiod^{*}, Mohamed Nadif^{*}

^{*}Université de Paris, CNRS, Centre Borelli UMR9010

^{**}Informatique Caisse des Dépôts et Consignations
<prénom.nom>@u-paris.fr

Résumé. Nous proposons un modèle pour l'apprentissage simultané de la représentation et le partitionnement (ou clustering) de graphes attribués. Basé sur un simple réseau convolutif sur graphe, notre modèle effectue le clustering en minimisant la différence entre la représentation réduite des données convoluées et la reconstruction des centroïdes calculés dans l'espace de dimension réduite. Nous montrons l'efficacité du modèle par rapport à l'état de l'art sur différents jeux de données de graphes attribués sur la tâche du clustering. Cet article est un résumé du papier (Fettal et al., 2022).

1 Introduction

L'apprentissage de nouvelles représentations, à la fois fidèles aux données initiales tout en faisant en sorte qu'elles aient une structure favorable au *clustering*, conduit souvent à un meilleur partitionnement profitable à l'utilisateur (Yamamoto et Hwang, 2014; Allab et al., 2016, 2018; Labiod et Nadif, 2021). Dans le contexte des graphes attribués, qui sont des graphes dont les nœuds ont des attributs, cette démarche n'est pas exploitée. Plus encore, les méthodes de partitionnement de graphes attribués souffrent d'une grande complexité spatiale et/ou computationnelle. Contrairement à la plupart des approches existantes, cet article vise à surmonter cette faiblesse en considérant un apprentissage simultané d'un plongement (embedding) et d'un *clustering* de graphe. En effet, en raison de la structure non euclidienne des graphes ainsi que la grande dimension des diverses caractéristiques décrivant les nœuds, la recherche de représentations euclidiennes appropriées qui incorporent l'information sur la structure et les caractéristiques de ces graphes est un défi intéressant pour l'apprentissage automatique (Hamilton et al., 2017). La littérature récente propose d'apprendre ces représentations automatiquement. De telles méthodes d'apprentissage de représentations visent à représenter les nœuds dans un espace de faible dimension où la proximité des représentations des nœuds doit être suffisamment similaire à celle de la représentation originale du graphe. Récemment, les *Réseaux convolutifs sur graphe* ou *Graph Convolutional Networks* (GCNs) (Kipf et Welling, 2017) ont suscité beaucoup d'attention en raison de leur capacité à apprendre des représentations de graphes de haute qualité, et par extension, sont efficaces pour différentes tâches liées aux graphes telles que la classification supervisée de nœuds, la prédiction de liens, et le partitionnement de nœuds qui est la tâche sur laquelle notre article se concentre. Cependant, ce