
Vers des algorithmes équitables pour graphes hétérogènes

Établissement/Laboratoire/Équipe d'accueil : Nantes Université/LS2N/DUKe

Encadrants : Marc GELGON (marc.gelgon@univ-nantes.fr), Hoel LE CAPITAINÉ (hoel.lecapitaine@univ-nantes.fr)

Contexte

L'analyse de phénomènes complexes modélisés par des graphes joue un rôle crucial dans de nombreux domaines d'application du monde réel où les décisions peuvent avoir un fort impact sur la société. En réalité, de nombreux graphes modélisent des situations complexes, dans lesquelles les nœuds peuvent représenter des entités différentes, et les liens entre les nœuds être de différents types, rendant ces graphes **hétérogènes**. Ainsi, la majeure partie des graphes en recommandation, les réseaux sociaux ou encore les réseaux académiques appartiennent à cette famille de graphes.

Par ailleurs, beaucoup de travaux ont montré que les modèles d'apprentissage pouvaient conduire à des inégalités de traitement, et mener à des décisions non équitables, voir (Mehrabi et al., 2021). La suppression simple des caractéristiques sensibles peut ne pas suffire à améliorer l'équité, car la corrélation entre ces caractéristiques et d'autres facteurs pourrait toujours introduire un biais. Les algorithmes d'apprentissage sur graphe ne sont pas épargnés par ce problème, et leur conception doit également tenir compte des spécificités de la structure, rendant non opérantes de nombreuses techniques de régularisation équitable proposées sur données dites *simples*.

De nombreux efforts ont été consacrés à l'amélioration de l'**équité** dans l'apprentissage des graphes (Chen et al., 2024). Plus précisément, des méthodes d'apprentissage fondées sur l'utilisation de réseaux de neurones sur graphes (GNNs) sensibles à l'équité ont émergé, se concentrant principalement sur la prévention de l'utilisation abusive des caractéristiques sensibles par le biais de régularisations ou de contraintes supplémentaires (Dai & Wang, 2022).

Problématique

Il n'existe aujourd'hui, à notre connaissance, aucune méthode de représentation équitable de graphes hétérogènes. Cette absence s'étend même à de simples graphes dirigés, non correctement représentés par les GNNs. D'un autre côté, l'incorporation de mécanismes d'attention et d'encodage de positionnement, à l'instar des grands modèles de langage (LLMs), conduit à des modèles appelés **Transformeurs sur Graphes** (TG), qui permettent d'obtenir des performances supérieures aux réseaux de neurones sur graphes usuels, avec une taille de modèle bien inférieure, que ce soit pour des graphes homogènes (Yun et al., 2019) ou hétérogènes (Hu et al., 2020).

Objectifs

Il est donc nécessaire de revenir sur le manque de méthodes de régularisation équitable sur des graphes hétérogènes. Le principe proposé est de s'inspirer des méthodes GNN récentes présentant ces régularisations, mais qui par la spécificité du graphe hétérogène, tiendra compte de la variété des nœuds et des relations entre les nœuds. Une extension vers l'intégration de ces concepts dans une architecture de type transformeur pourra être envisagée.

Bibliographie

- Chen, A., Rossi, R. A., Park, N., Trivedi, P., Wang, Y., Yu, T., Kim, S., Derroncourt, F., & Ahmed, N. K. (2024). Fairness-aware graph neural networks: A survey. *ACM Transactions on Knowledge Discovery from Data*, 18(6), 1-23.
- Dai, E., & Wang, S. (2022). Learning fair graph neural networks with limited and private sensitive attribute information. *IEEE Transactions on Knowledge and Data Engineering*, 35(7), 7103-7117.
- Hu, Z., Dong, Y., Wang, K., & Sun, Y. (2020). Heterogeneous graph transformer. *Proceedings of the Web Conference 2020*, 2704-2710.
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, 54(6), 1-35.
- Yun, S., Jeong, M., Kim, R., Kang, J., & Kim, H. J. (2019). Graph transformer networks. *Advances in Neural Information Processing Systems*, 32.