



HAL
open science

Application de GNN sur des graphes non-attribués : Benchmark de performances

Ikram Boukharouba, Florence Sèdes, Christophe Bortolaso, Florent Mouysset

► **To cite this version:**

Ikram Boukharouba, Florence Sèdes, Christophe Bortolaso, Florent Mouysset. Application de GNN sur des graphes non-attribués : Benchmark de performances. 40ème Congrès annuel INFormatique des ORganisations et Systèmes d'Information et de Décision (INFORSID 2022), May 2022, Dijon, France. à paraître. hal-03783370

HAL Id: hal-03783370

<https://hal.science/hal-03783370>

Submitted on 11 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Application de GNN sur des graphes non-attribués : Benchmark de performances.

Ikram Boukharouba^{1,2}, Florence Sèdes¹, Christophe Bortolaso², Florent Mouisset².

¹ IRIT, Université Toulouse III Paul Sabatier, Toulouse/France

² Berger Levrault, Labège/France

ikram.boukharouba@irit.fr

Résumé : Les Graph Neural Networks (GNNs) sont des méthodes d'apprentissage profond construisant un modèle lié à un graphe. Cet apprentissage rend possible la classification et la prédiction de nœuds ou d'arêtes manquantes. Cependant, ces approches produisent des résultats significatifs lorsque les features sur les nœuds sont suffisamment nombreuses, ce qui n'est pas le cas pour les problèmes de terrain. En effet, du fait de l'origine des données et de leur génération non maîtrisée, les graphes résultants contiennent peu de features. En outre certaines contraintes légales, peuvent limiter l'utilisation des données. Dans cet article, nous montrons que des graphes issus de traces d'activité utilisateur (logs) comportent également de telles limites. Notre démonstration commence par une synthèse de l'état de l'art justifiant notre utilisation des GNNs pour la tâche de classification. Par la suite, nous détaillons les expérimentations conduites sur notre corpus de données de terrain (logs issus du produit Sedit produit par Berger-Levrault). Les résultats démontrent les limites effectives des GNNs. Enfin, certaines approches pouvant proposer des solutions à ce problème sont avancées.

ABSTRACT. Graph Neural network (GNN) is a deep learning method that builds a model linked to a graph. This learning allows the classification and prediction of nodes and/or missing edges. However, these approaches produce noteworthy results only when the features on the nodes are enough numerous. In real-world, especially industrial problems, these features may not always be as numerous which is not the case for "real-world problems". Indeed, due to the origin of the data and their uncontrolled generation process, the resulting graphs may contain few features or even none. In addition, certain legal constraints may limit the use of data. In this article, we show that graphs derived from traces of user activity (logs) also have such limits. Our demonstration begins with a summary of the state of the art justifying the use of GNNs for the classification task in our use case. Thereafter, we detail the experiments conducted on our corpus of field data (logs from the Sedit product produced by Berger-Levrault). The results demonstrate the effective limitations of GNNs. Finally, some approaches that can offer solutions to this problem are put forward.

MOTS-CLES : Traces d'activité utilisateurs, Logs, Application industrielle, Graphe de navigation, Graph Neural Network (GNN), Graphe non-attribué.

KEYWORDS : Traces of user activity, Logs, Industrial application, Navigation graph, Graph Neural Network (GNN), Non-attributed graph.
