Beyond-Accuracy Perspectives on Graph Neural Network-Based Models for Behavioural User Profiling

Erasmo Purificato

Otto von Guericke University Magdeburg & Leibniz Institute for Educational Media | Georg Eckert Institute

email: erasmo.purificato@ovgu.de – personal website: www.erasmopurif.com

Abstract

The presented doctoral research aims to develop a **GNN-based behavioural user profiling framework** focusing on three *beyond-accuracy* perspectives:

- **Privacy**, to study how to provide methods for graph data without neither exposing personal user information nor corrupting the profiles creation.
- **Fairness**, to provide user representations that are free of any inherited discrimination by developing debiasing approaches to be applied on state-of-the-art GNN-based user profiling models.
- **Explainability**, to produce human-understandable descriptions of the framework results by designing an adaptive and personalised user interface which provides tailored explanations to the end-users, depending on their specific user profiles.

Background and Motivation

The main goal of **user profiling** is to infer an individual's interests, personality traits or behaviours from generated data to create an efficient user representation, i.e. a *user model*. Modern systems focus on profiling users' data based on individuals' actions and interactions (*implicit user profiling*). This approach is also referred to as **behavioural user profiling**.

A natural way to model these behaviours is through *graphs*, where edges can easily describe the interactions between users, represented by nodes. In this light, **Graph Neural Networks** (GNNs) are the perfect class of methods to deal with data represented by graph data structures.

Existing GNN-based approaches evaluate user profiling models based on the effectiveness of a classification task at predicting a user's personal characteristics.

The ***focus*** of the thesis is to look *beyond* the usual accuracy-based approaches and develop a graph neural network-based model for behavioural user profiling whose resulting *fair* and *privacy-preserving* user models are applied as the input of a downstream recommender system, acting as a bridge with an *adaptive and personalised user interface* implemented to provide tailored explanations to the end-users depending on their specific user profiles.

Research goals for the GNN-based behavioural user profiling framework

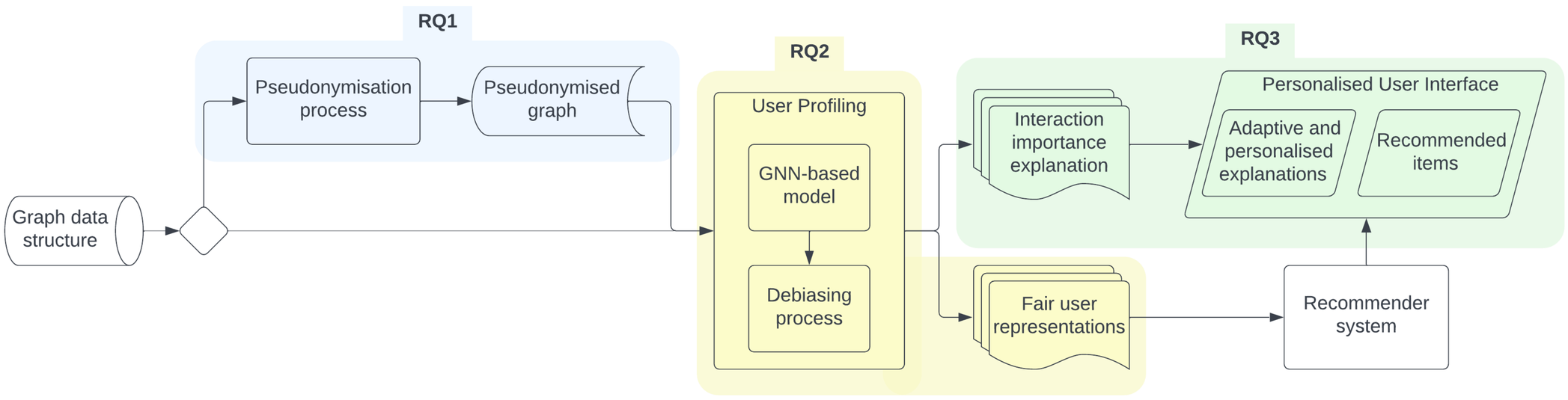

Figure 1. Overall architecture of the framework. Highlighted areas represent the modules developed to address each corresponding research question.

RQ1 – Privacy
How can we guarantee personal data protection on a graph data structure while avoiding affecting user models construction and retaining the performance of the recommender system built upon them?

RQ2 – Fairness
How do we build fair user representations from GNN-based user profiling models to keep the input of the downstream recommender debiased?

RQ3 – Explainability
How can we personalise user interfaces to adapt the explanations to the needs, demands and requirements of different end-user profiles, considering their distinct knowledge, background and expertise?

Results and contributions to date

- *Privacy* perspective (**RQ1**): pseudonymisation approach for privacy-preserving recommendations on academic graph data to retain system performance [1].
- *Fairness* perspective (**RQ2**): assessment of fairness in state-of-the-art GNN-based models (i.e. CatGCN and RHGN) for behavioural user profiling tasks, in terms of disparate impact and disparate mistreatment. Work currently submitted at CIKM'22.
- *Explainability* perspective (**RQ3**): user study among in the academic domain to evaluate how two different explainable UIs, providing explanations for the outcomes of a recommender system for scientific papers, are perceived, in terms of understandability, trust and user satisfaction [2].
- *Explainability* perspective (**RQ3**): organised the First Workshop on Adaptive and Personalised Explainable User Interfaces (APEx-UI 2022) at IUI'22 [3].
- *Fairness* and *Explainability* perspectives (**RQ2 – RQ3**): developed a system to show how the use of Responsible AI techniques can lead to the growth of a domain expert's trust and reliance on an AI system. Proposed a novel "*Trust&Reliance Scale*" to evaluate XAI systems [4].

References

[1] E. Purificato, S. Wehnert, and E. W. De Luca. *Dynamic Privacy-Preserving Recommendations on Academic Graph Data*. In Computers 10, 9 (2021), 107.

[2] E. Purificato, B. Aiyer, P. Karanam, M. Pattadkal, and E. W. De Luca. *Evaluating Explainable Interfaces for a Knowledge Graph-Based Recommender System*. In Proceedings of the 8th Joint Workshop on Interfaces and Human Decision Making for Recommender Systems, co-located with RecSys'21. 73–88.

[3] E. Purificato, C. Musto, P. Lops, and E. W. De Luca. *First Workshop on Adaptive and Personalized Explainable User Interfaces (APEx-UI 2022)*. In 27th International Conference on Intelligent User Interfaces (IUI '22 Companion). ACM, New York, NY, USA, 1–3.

[4] E. Purificato, F. Lorenzo, F. Fallucchi, and E. W. De Luca. *The Use of Responsible Artificial Intelligence in the Context of Loan Approval Processes*. In International Journal of Human-Computer Interaction (2022), 1-20.

Supervisor

Prof. Dr.-Ing. Ernesto W. De Luca

email: ernesto.deluca@ovgu.de

Doctoral Consortium, UMAP'22: 30th ACM Conference on User Modeling, Adaptation and Personalization – July 4-7, 2022, Barcelona, Spain