

Report on the 1st International Workshop on Graph-Based Approaches in Information Retrieval (IRonGraphs 2024) at ECIR 2024

Ludovico Boratto
University of Cagliari
Italy
ludovico.boratto@acm.org

Daniele Malitesta
Université Paris-Saclay,
CentraleSupélec, Inria, France
daniele.malitesta@centralesupelec.fr

Mirko Marras
University of Cagliari
Italy
mirko.marras@acm.org

Giacomo Medda
University of Cagliari
Italy
giacomo.medda@unica.it

Cataldo Musto
University of Bari
Italy
cataldo.musto@uniba.it

Erasmus Purificato
Otto von Guericke University
Magdeburg, Germany
erasmo.purificato@acm.org

Abstract

The First International Workshop on Graph-Based Approaches in Information Retrieval (IRonGraphs 2024) was held as a physical (in-person) event on March 24, 2024, in conjunction with the 46th European Conference on Information Retrieval (ECIR 2024) in Glasgow (Scotland). The scientific program included paper, spotlight, and poster presentations. Two keynote talks were delivered by Francesco Fabbri (Spotify, Spain) and Ruihong Qiu (University of Queensland, Australia). This report presents an overview of the activities conducted during the workshop and the main topics covered.

Date: 24 March 2024.

Website: <https://irongraphs.github.io/ecir2024/>.

1 Introduction

On March 24, 2024, the First International Workshop on Graph-Based Approaches in Information Retrieval (IRonGraphs 2024)¹ [Boratto et al., 2024] was held in-person, in conjunction with the 46th European Conference on Information Retrieval (ECIR 2024) in Glasgow (Scotland). The workshop was organized by the Department of Mathematics and Computer Science at the University of Cagliari (Italy), CentraleSupélec, Inria at Université Paris-Saclay (France), the Department of Computer Science at the University of Bari (Italy), and the Otto von Guericke University Magdeburg (Germany).

The workshop aimed to provide the community with a dedicated forum regarding the broader field of graph learning for information retrieval [Blanco and Lioma, 2012; Wu et al., 2023]. In

¹<https://irongraphs.github.io/ecir2024/>.

terms of attendance, around 30 people followed the venue. The two keynote speeches were held by Francesco Fabbri (Spotify, Spain) and Ruihong Qiu (University of Queensland, Australia). The scientific program included paper presentations, a spotlight session, and poster presentations. The workshop covered several micro- and macro-aspects in graph-based information retrieval spanning, among others, graph models for personalized recommendation, knowledge graphs, adversarial learning, explainability, reproducibility, and evaluation challenges.

2 Summary of the keynote talks

2.1 Graph Foundation Model for Personalization

In this talk, Francesco Fabbri (Spotify, Spain) discussed the importance of introducing and defining the novel concept of a Graph Foundation Model (GFM) for the personalized recommendation of digital audio content at a large scale. The talk emphasized the application of graph learning techniques alongside large language models to address various tasks in the domain of talk audio space, such as podcast and audiobook personalization.

2.2 Graph Learning Methods in Session-based Recommendations and Legal Case Retrieval

In this talk, Ruihong Qiu (University of Queensland, Australia) discussed the utility of graph learning in information retrieval, focusing on two case scenarios: session-based recommendation and legal case retrieval. As for the former, user history can be represented as a graph capturing user behavior. As for the latter, legal cases are transformed into graphs to maintain contextual information across the various parties involved. Overall, the talk suggested that any information retrieval problem can benefit from a proper graph representation to harness valuable information.

3 Summary of the papers

3.1 KGUF: Simple Knowledge-aware Graph-based Recommender with User-based Semantic Features Filtering

The paper discusses how Knowledge Graph Collaborative Filtering (KGCF) models currently leverage Knowledge Graphs (KGs) to enhance recommendation accuracy by mining hidden user intents. However, the authors suggest that simpler approaches might be as effective, especially when considering user historical preferences. To this end, they introduce KGUF, a KGCF model that learns semantic features from KGs to refine item profiles. KGUF incorporates decision trees to model user profiles in a lightweight manner while being competitive with the state-of-the-art.

Authors: Salvatore Bui, Alberto Carlo Maria Mancino, Antonio Ferrara, Daniele Malitesta, Tommaso Di Noia, and Eugenio Di Sciascio.

3.2 The Impact of Source-Target Node Distance on Vicious Adversarial Attacks in Social Network Recommendation Systems

The paper addresses the vulnerability of social network recommendation systems to adversarial attacks, particularly when considering sparse attacks performed by frameworks like SAVAGE. In this respect, the authors investigate how the distance between the source and target nodes in the social network may impact the effectiveness of such attacks. On a Twitter dataset, the analysis finds that SAVAGE performs exceptionally well, especially for longer distances between the source and target nodes, while also minimally affecting the network structure.

Authors: Federico Albanese, Giovanni Trappolini, Lorenzo Scarlino and Fabrizio Silvestri.

3.3 The Effectiveness of Graph Contrastive Learning on Mathematical Information Retrieval

The paper deals with Mathematical Information Retrieval (MIR). Differently from traditional Information Retrieval problems, where word frequencies and syntactic structures may play a crucial role, MIR involves formulas and mathematical equations, where a proper representation of the notation structure is sought. As datasets for MIR are limited in the literature, the authors propose a graph contrastive learning (GCL) method to learn mathematical notations without the need for labeled data. Results show that the proposed approach can outperform TangentCFT.

Authors: Pei Syuan Wang and Hung-Hsuan Chen.

3.4 Identifying User Shopping Needs in Voice Product Questions for Proactive Shopping Recommendations

The paper highlights the gap in proactive recommendations from voice assistants to users, especially when it comes to Shopping Product Questions (SPQs), where a user's product-related query may indicate an underlying shopping need. To identify SPQs, the work proposes to analyze latent user behavior patterns inferred from past shopping history, using a novel Mixture-of-Experts (MoE) model. Results demonstrate the efficacy of the proposed approach both in an offline and online evaluation setting.

Authors: Besnik Fetahu, Nachshon Cohen, Elad Haramaty, Liane Lewin-Eytan, Oleg Rokhlenko and Shervin Malmasi.

3.5 Spotlight session

This was done through the following spotlight talks:

- CaseGNN: Graph Neural Networks for Legal Case Retrieval with Text-Attributed Graphs.
Authors: Yanran Tang, Ruihong Qiu, Yilun Liu, Xue Li, and Zi Huang.
- XSearchKG: A Platform for Explainable Keyword Search over Knowledge Graphs.
Authors: Leila Feddoul, Martin Birke, and Sirko Schindler.
- Performance Comparison of Session-Based Recommendation Algorithms Based on GNNs.
Authors: Faisal Shehzad and Dietmar Jannach.

-
- Event-specific Document Ranking through Multi-stage Query Expansion Using Knowledge Graphs.
Authors: Sara Abdollahi, Tin Kuculo, and Simon Gottschalk.
 - Hypergraphs with Attention on Reviews for Explainable Recommendation.
Authors: Theis E. Jendal, Trung-Hoang Le, Hady Lauw, Matteo Lissandrini, Peter Dolog, and Katja Hose.
 - Overcoming Recall Failures in Ad Hoc Search using Graphs.
Authors: Sean MacAvaney.

This special session aimed to provide a dedicated spot to those research works in the field of graph-based information retrieval that were presented in the main conference, or preliminary (but promising) research directions to share and discuss with the audience to receive useful feedback.

4 Conclusions

In conclusion, the International Workshop on Graph-Based Approaches in Information Retrieval had a positive debut in its first edition, incentivizing fruitful discussions and possible future collaborations on novel ideas regarding all the topics covered during the workshop. Tentative plans are being formed to organize the second workshop next year. Furthermore, a proceedings volume is currently under production and should be published soon via Springer’s “Communications in Computer and Information Science” series.

The organizers would like to thank the authors of the submitted papers and the spotlight presenters for their interest in the venue, the two keynote speakers for their insightful talks, the reviewers for their amazing help in shaping an interesting workshop program, and the attendees for their active participation. We are willing to share our experience and arrange joint initiatives in future venues. As a kind reminder, we suggest that interested researchers and practitioners follow our website² and X (former Twitter) profile³ and stay in touch via email, for ongoing and future initiatives.

References

- Roi Blanco and Christina Lioma. Graph-based term weighting for information retrieval. *Inf. Retr.*, 15(1):54–92, 2012.
- Ludovico Boratto, Daniele Malitesta, Mirko Marras, Giacomo Medda, Cataldo Musto, and Erasmo Purificato. First international workshop on graph-based approaches in information retrieval (irongraphs 2024). In *ECIR (5)*, volume 14612 of *Lecture Notes in Computer Science*, pages 415–421. Springer, 2024.
- Shiwen Wu, Fei Sun, Wentao Zhang, Xu Xie, and Bin Cui. Graph neural networks in recommender systems: A survey. *ACM Comput. Surv.*, 55(5):97:1–97:37, 2023.

²<https://irongraphs.github.io/ecir2024/>.

³<https://twitter.com/IRonGraphsWS>.