Graph Representation Learning for Cross-Domain Recommendation

Tendai Mukande[1,2]

[1]School of Computing, Dublin City University, Dublin, Ireland
[2]SFI Centre for Research Training in Machine Learning, Dublin, Ireland

Abstract
This paper discusses the current challenges in modeling real-world recommendation scenarios and proposes a Graph Neural Network-based Representation Learning framework for heterogeneous data. By incorporating a graph attention mechanism, larger weights will be assigned to the most important neighbours to improve link prediction performance. To resolve data sparsity issues, experiments would be conducted on cross-domain recommendation.

Keywords
Graph Neural Networks, Graph Representation Learning, Cross-Domain Recommendation, Graph Attention Networks.

1. Introduction

The heterogeneity of real-world data result in modelling challenges for personalized recommendation. Recommender systems have been shown to be indispensable in various information access systems by facilitating decision-making as well as overcoming information overload [1]. Recommendation systems retrieve information from various sources to output personalized suggestions [2]. They also influence consumer behaviour which helps to improve business profit [3]. Graph Neural Networks (GNNs) have been considered as the state-of-the-art approach to represent complex interconnections between users and items in real world scenarios [4]. Using GNNs, user-item interactions can be represented into lower dimensional vector spaces where multiple relations (edges) connect user/items (nodes) [5].

Real-world recommendation scenarios involve multiple users, products and behaviour ("*view*", "*purchase*", "*mention*, "*add-to-cart*" among others), which can represented as a Graph Neural Network with multi-path relations. A good graph representation model for such scenarios should maintain the structural and semantic information for efficient implementation of the downstream tasks such as link prediction in personalized recommendation.

This research aims to resolve the following challenges, which are among the research interest areas in recommendation systems:

PhD Symposium on Future Directions in Information Access: 12th European Summer School on Information Retrieval, July 15–23, 2022, Lisbon, Portugal.
tendai.mukande2@mail.dcu.ie (T. Mukande)

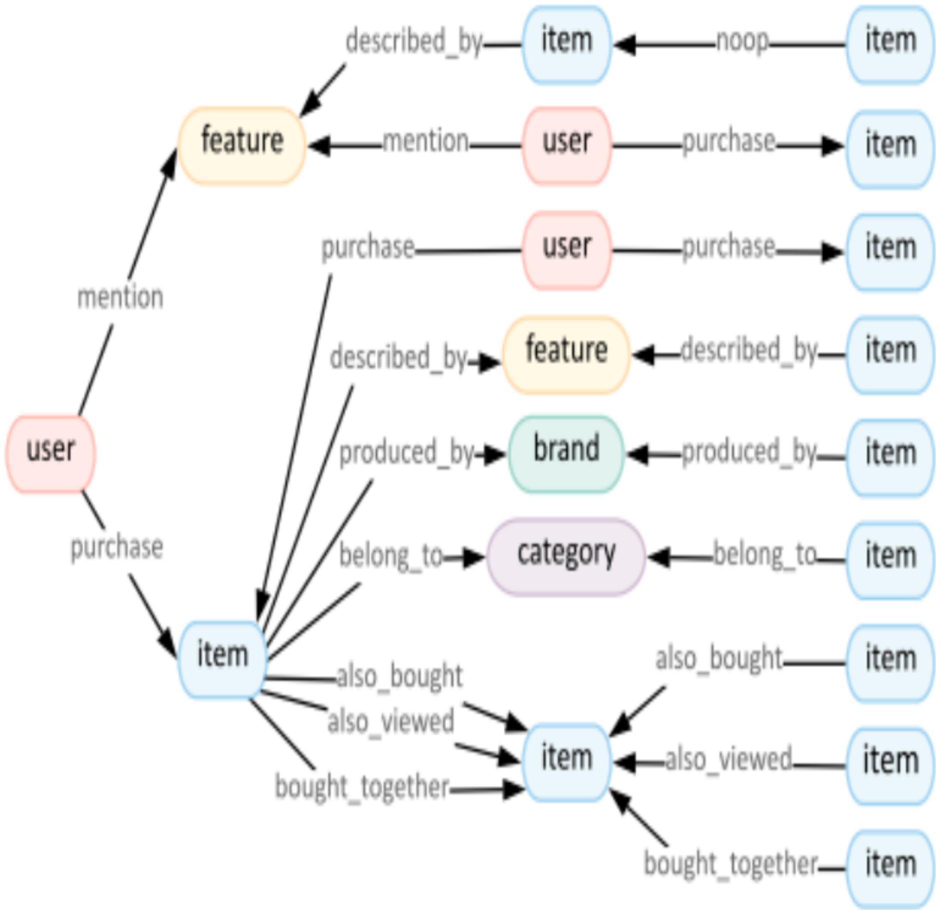

Figure 1: Real-world recommendation scenarios consist of disparate patterns. Heterogeneous Representation Learning is needed to preserve the structural and semantic information to facilitate efficient downstream tasks implementation. Image Credit [6].

Challenge 1: Effective Heterogeneous Information Representation

Real world recommendation scenarios involve disparate user and product data patterns. Due to this phenomenon, learning representations effectively while maintaining the rich semantic and robust structural information is challenging. Moreover, many state-of-the-art models are inefficient when handling scenarios involving heterogeneous data spanning multiple domains [7]. In addition, the models for heterogeneous datasets are complex.

Challenge 2: Preservation of the model structure and semantics at lower computational cost [8]

Apart from the model complexity indicated in Challenge 1, processing heterogeneous data involves many operations and retraining, which results in high computational cost [9]. There is therefore need for a not-so complex model to efficiently process the heterogeneous data [10].

Challenge 3: Data sparsity issues

In addition to the model complexity and efficient representation issues, using heterogeneous data may result in data sparsity issues as new and infrequent scenarios usually have insufficient data for recommendation. While most state of the art models are effective on homogeneous data, they have limited applicability when it comes to heterogeneous data [11]. As a result, the data sparsity issues are frequent phenomena, especially with new and infrequent situations.

To resolve the highlighted challenges, this research project will focus on developing a Graph Neural Network-based framework, which consists of a heterogeneous representation learning layer, and attention layers to improve the link prediction performance by giving more weight to important node relations. To resolve the sparsity issues, domain adaptation will be proposed.

The main contributions to resolve the highlighted challenges are summarized as below:

1. Construction of a Heterogeneous Graph Representation Learning model which captures structural and semantic information in the dataset to address Challenge 1.

2. Integration of graph attention layers to the proposed representation learning model to address Challenge 2, whereby computations would be mainly focused on the sub-graph related to the particular recommendation task instead of the whole Graph Neural Network representation. Most of the prior research works on heterogeneous graph attention have not been applied to multi-domain scenarios.

3. Use of a shared graph model for multi-domain relation inference to address Challenge 3 by the inter-domain user-product relations. Using the model proposed by Cui et al. [12] as a baseline, experiments will be conducted using the Graph Transformer Network to identify the best meta-paths for link prediction across different product domains.

2. RELATED WORK

2.1. Graph Representation Learning

Due to their ability to handle heterogeneous data, there is growing interest in the use of GNN methods in recommender systems, particularly systems [13], Heterogeneous Graph Representation Learning has been shown by recent studies to be highly useful in dimensionality reduction, particularly for tasks which involve abundant data [14].

For the effective use of GNNs for Heterogeneous Graph Representation Learning in recommendation, the following issues should be considered: proper graph construction, which requires the appropriate representation of nodes (entities) and edges (relations); adaptive design of the recommendation GNN information aggregation [10]; model optimisation including the loss functions and sampling for the particular tasks [15]; as well as the trade off between the model efficiency and computational cost [16].

Dai et al. [17] proposes Heterogeneous Graph Representation Learning for high-order academic graph connectivity exploration in knowledge aware recommendation. Results from their study show that their model is efficient at alleviating bias in networks with limited sensitive information. They use adversarial debiasing for fair node representations on the GNN. Using their FairGNN model as a baseline, more experiments can be done to study representation learning for other datasets.

2.2. Graph Attention Networks

To alleviate data sparsity issues and improve the recommendation accuracy; modelling multi-user behaviour such as *view*, *collect* and *add-to-cart*; in addition to *purchase* and *review* is essential [7]. While the goal is to recommend the products that would be eventually purchased, there is also need to take into account the multi-user behaviour that is related to the ultimate recommendation. In the work by Hong et al. [18], behaviour changes due to changing trends were modelled with high accuracy. They proposed a novel Heterogeneous Graph Structural Attention Neural Network (HetSANN) which encodes Heterogeneous Information Graph structural information without using meta-paths to learn lower dimensional vector representations. Their work is a good foundation which can be applied to other heterogeneous datasets. The authors in [19] investigate the influence of multi-modality information exchange on user preference. They proposed the Multi-Modal Graph Convolutional Network which uses parallel interaction graphs and treats all neighbour information equally. They do not, however, capture noisy information and adaptive user preference [1]. Another interesting approach in graph attention is by Tao et al. [20] in which they investigate adaptive user-item interaction. They propose the Multi-modal Graph Attention Network, in which they use the gated attention mechanism to determine weights based on varying user preferences, thereby capturing complex interaction patterns and different user behaviours which are essential for recommendation accuracy improvement.

Yang et al. [21] propose a Heterogeneous Graph Attention network which samples data into a unified sub-space and uses graph attention for multi- target domain prediction. This is a promising research direction, particularly when integrated to multi-domain models for complex user-item interactions.

Other interesting works on graph attention networks include the Heterogeneous Graph Attention network (HGAT) by Wang et al [11], who use the approach for multiple domain semantic transfer with label propagation. The HGAT network is learned to map multi-domain samples into a unified subspace and multi-domain relationship optimization while the authors in [22] studied the Heterogeneous Graph Transformer to model heterogeneous web graphs using heterogeneous attention.

2.3. Cross-Domain Recommendation

Cross-Domain Recommendation (CDR) has been used to resolve data sparsity problems in recommendation systems [21]. CDR can be classified into single-target and multi-target CDR [23] depending on the number of domains involved whereby information is transferred from the source domain(s) to the target domain(s) [21]. Relation inference across multiple domains is one way to ensure that sufficient data is available for a particular recommendation task, which can improve the performance [12].

These following works are an interesting direction in domain adaptation, particularly where heterogeneous data is involved. More study needs to be done, however, particularly on extending them to multi-behaviour, multi-target personalized recommendation scenarios. Long et al. [24] proposed Joint Adaptation Networks (JAN) to jointly align multiple layers distribution across domains, which greatly facilitates the the Deep Adaptation Network ability, at the same time simplifying the training process. A novel Deep Reconstruction Classification Network (DRCN) has also been explored by Ghifary et al [25], which learns semantic alignment for multi-target domain adaptation.

3. RESEARCH OBJECTIVES

The main purpose of this research project is to address the limitations of state-of-the-art models in heterogeneous recommendation scenarios. A novel unified, framework for recommendation would be developed, with a focus on Heterogeneous Graph Representation Learning, multi-behaviour graph attention and Cross-Domain Recommendation. The following research questions would be addressed model:

Research Question 1: How do we effectively represent heterogeneous data into a lower-dimensional embedding space?

Research Question 2: How do we reduce the model computational cost?

Research Question 3: What is the proposed model performance evaluation against state-of-the art baselines for personalized recommendation tasks?

4. PROPOSED APPROACH

The proposed approach in this project would be divided into the following tasks: 1. Model design and 2. Evaluation.

Model Design: The model design would mainly involve the development of a heterogeneous graph representation of the multiple user-item interactions shown in Figure 1. A single heterogeneous graph network would be used to represent the multi- user-product product interactions. Moreover, the model would also be able to capture the different semantic (behavioural) and structural (user/item) features in the graph structure. To address Challenge 1 and the first Research Question, the proposed model aims to prioritize the end-to-end stages of the recommendation pipeline by efficient representation of the semantic and structural features. To address Challenge 2 and Research Question 2, the model should be computationally efficient in providing high accuracy multi-scenario recommendations. However, the priority is on effective representation and processing.

For self-supervised attention to the recommendation model, a Heterogeneous Graph Attention Network for attention is proposed. From the complex relations, the transformer network learns the most suitable meta-paths for the complex interactions across the inter-linked domains, thus alleviating data sparsity issues and address Challenge 3 and Research Question 3. The proposed approach extends to the HeroGRAPH model [12], but instead of recurrent attention, different weights are assigned to relations based on their importance using the Heterogeneous Graph Transformer network.

Evaluation: Experiments will be conducted on the personalized recommendation performance evaluation. The proposed baselines are as follows: Bayesian Personalized Ranking [26], GraphSAGE-pool [27] and the HeroGRAPH model [28].

5. EXPECTED RESULTS

The output from this work would be a unified, computationally efficient GNN-based recommendation model for cross-domain recommendation. While the main focus in this project is resolving the challenges on heterogeneous data representation and data sparsity, in future, the framework could be extended to resolve issues such as improving robustness to adversarial attacks for instance, in credit-scoring models by spammers.

6. CONCLUSION

In this paper, the proposed design of a Graph Neural Network-based framework for Cross-Domain Recommendation is discussed. The expected results for the project is to resolve data sparsity and high computational cost issues in Recommender Systems.

Acknowledgments

Special thanks to Dr. Annalina Caputo, Prof. Noel O'Connor and Dr. Ruihai Dong for the feedback and guidance in this project.

This publication has emanated from research supported in part by a grant from Science Foundation Ireland under Grant number 18/CRT/6183. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References

[1] P. K. Singh, P. K. D. Pramanik, A. K. Dey, P. Choudhury, Recommender systems: an overview, research trends, and future directions, International Journal of Business and Systems Research 15 (2021) 14–52.

[2] D. Jannach, G. Adomavicius, Recommendations with a purpose, in: Proceedings of the 10th ACM conference on recommender systems, 2016, pp. 7–10.

[3] D. Jannach, M. Jugovac, Measuring the business value of recommender systems, ACM Transactions on Management Information Systems (TMIS) 10 (2019) 1–23.

[4] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, M. Sun, Graph neural networks: A review of methods and applications, AI Open 1 (2020) 57–81.

[5] Z. Tian, Y. Liu, J. Sun, Y. Jiang, M. Zhu, Exploiting group information for personalized recommendation with graph neural networks, ACM Transactions on Information Systems (TOIS) 40 (2021) 1–23.

[6] Y. Xian, Z. Fu, S. Muthukrishnan, G. De Melo, Y. Zhang, Reinforcement knowledge graph reasoning for explainable recommendation, in: Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval, 2019, pp. 285–294.

[7] C. Chen, W. Ma, M. Zhang, Z. Wang, X. He, C. Wang, Y. Liu, S. Ma, Graph heterogeneous multi-relational recommendation, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 35, 2021, pp. 3958–3966.
[8] C. Li, X. Peng, Y. Niu, S. Zhang, H. Peng, C. Zhou, J. Li, Learning graph attention-aware knowledge graph embedding, Neurocomputing 461 (2021) 516–529.
[9] X. Wang, Y. Lu, C. Shi, R. Wang, P. Cui, S. Mou, Dynamic heterogeneous information network embedding with meta-path based proximity, IEEE Transactions on Knowledge and Data Engineering (2020).
[10] W. Luo, H. Zhang, X. Yang, L. Bo, X. Yang, Z. Li, X. Qie, J. Ye, Dynamic heterogeneous graph neural network for real-time event prediction, in: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2020, pp. 3213–3223.
[11] X. Wang, H. Ji, C. Shi, B. Wang, Y. Ye, P. Cui, P. S. Yu, Heterogeneous graph attention network, in: The world wide web conference, 2019, pp. 2022–2032.
[12] Q. Cui, T. Wei, Y. Zhang, Q. Zhang, Herograph: A heterogeneous graph framework for multi-target cross-domain recommendation., in: ORSUM@ RecSys, 2020.
[13] S. Zhang, L. Yao, A. Sun, Y. Tay, Deep learning based recommender system: A survey and new perspectives, ACM Computing Surveys (CSUR) 52 (2019) 1–38.
[14] C. Shi, X. Wang, S. Y. Philip, Heterogeneous graph representation learning and applications, 2022.
[15] Z. Liu, C. Chen, X. Yang, J. Zhou, X. Li, L. Song, Heterogeneous graph neural networks for malicious account detection, in: Proceedings of the 27th ACM International Conference on Information and Knowledge Management, 2018, pp. 2077–2085.
[16] C. Gao, X. Wang, X. He, Y. Li, Graph neural networks for recommender system, in: Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining, 2022, pp. 1623–1625.
[17] E. Dai, S. Wang, Learning fair graph neural networks with limited and private sensitive attribute information, arXiv preprint arXiv:2009.01454 (2020).
[18] H. Hong, H. Guo, Y. Lin, X. Yang, Z. Li, J. Ye, An attention-based graph neural network for heterogeneous structural learning, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 34, 2020, pp. 4132–4139.
[19] Y. Wei, X. Wang, L. Nie, X. He, R. Hong, T.-S. Chua, Mmgcn: Multi-modal graph convolution network for personalized recommendation of micro-video, in: Proceedings of the 27th ACM International Conference on Multimedia, 2019, pp. 1437–1445.
[20] Z. Tao, Y. Wei, X. Wang, X. He, X. Huang, T.-S. Chua, Mgat: Multimodal graph attention network for recommendation, Information Processing & Management 57 (2020) 102277.
[21] X. Yang, C. Deng, T. Liu, D. Tao, Heterogeneous graph attention network for unsupervised multiple-target domain adaptation, IEEE Transactions on Pattern Analysis and Machine Intelligence (2020).
[22] Z. Hu, Y. Dong, K. Wang, Y. Sun, Heterogeneous graph transformer, in: Proceedings of The Web Conference 2020, 2020, pp. 2704–2710.
[23] F. Zhu, Y. Wang, C. Chen, J. Zhou, L. Li, G. Liu, Cross-domain recommendation: challenges, progress, and prospects, arXiv preprint arXiv:2103.01696 (2021).
[24] M. Long, H. Zhu, J. Wang, M. I. Jordan, Deep transfer learning with joint adaptation

networks, in: International conference on machine learning, PMLR, 2017, pp. 2208–2217.
[25] M. Ghifary, W. B. Kleijn, M. Zhang, D. Balduzzi, W. Li, Deep reconstruction-classification networks for unsupervised domain adaptation, in: European conference on computer vision, Springer, 2016, pp. 597–613.
[26] S. Rendle, C. Freudenthaler, Z. Gantner, L. Schmidt-Thieme, Bpr: Bayesian personalized ranking from implicit feedback, arXiv preprint arXiv:1205.2618 (2012).
[27] W. Hamilton, Z. Ying, J. Leskovec, Inductive representation learning on large graphs, Advances in neural information processing systems 30 (2017).
[28] C. Zhang, D. Song, C. Huang, A. Swami, N. V. Chawla, Heterogeneous graph neural network, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019, pp. 793–803.