Sequential Recommendation Models: A Graph-based Perspective

ANDREAS PEINTNER, University of Innsbruck, Austria

Recommender systems (RS) traditionally leverage the users' rich interaction data with the system, but ignore the sequential dependency of items. Sequential recommender systems aim to predict the next item the user will interact with (e.g., click on, purchase, or listen to) based on the preceding interactions of the user with the system. Current state-of-the-art approaches focus on transformer-based architectures and graph neural networks. Specifically, graph-based modeling of sequences has been shown to be state-of-the-art by introducing a structured, inductive bias into the recommendation learning framework. In this work, we outline our research into designing novel graph-based methods for sequential recommendation.

CCS Concepts: • **Information systems → Recommender systems**.

Additional Key Words and Phrases: Recommender Systems, Sequential Recommendation, Graph Neural Network

ACM Reference Format:
Andreas Peintner. 2023. Sequential Recommendation Models: A Graph-based Perspective. In *Seventeenth ACM Conference on Recommender Systems (RecSys '23), September 18–22, 2023, Singapore, Singapore.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3604915.3608776

1 INTRODUCTION

The most widely used traditional RS approaches include content-based and collaborative filtering systems [20]. Collaborative filtering systems predict the users' preference based on the interests of other, similar users: If Users A and B have a similar interest in one or multiple items, then they are likely to have similar interests for other items too. Content-based systems model the users' preferences only based on positive interactions and aim to match similar items, e. g., if the user listens to songs of a certain singer on a music platform, it will more likely recommend songs from the same singer. These conventional RS model the user-item interactions in a static way and ignore any temporal information contained in the interaction sequence such as timestamps or order. Therefore, such RS are only able to capture the general preference of the user. In contrast, sequential recommendation (SR) systems suggest succeeding items or whole sequences of possible interest to the user by modeling the sequential dependencies in the user-item interaction history. SR emphasizes the dynamics in the interaction sequence and uses long-term and short-term dependencies to capture the current preference of a user to provide more accurate recommendations [32].

Fig. 1. An example of SR: User A booked a flight, a hotel and rented a car. What will be his next action?

Using SR methods as a recommendation model has distinct advantages over general recommender systems. In real-world scenarios, interactions mostly happen successively and are not isolated from each other. Figure 1 shows an example of a shopping spree of User A. In this scenario (the user is booking a holiday), each action depends on the prior ones and so all interactions are sequentially dependent: As a next action User A might book tickets for a tourist attraction. This example also shows that user-item interactions usually happen in a certain sequential context. Additionally, the preference of the user and the popularity of different items are dynamic (e. g., music or clothing) rather than static over time due to personal development and trends [23]. These typical characteristics of online interaction sequences are captured by SR systems, but are hard to model with traditional RS.

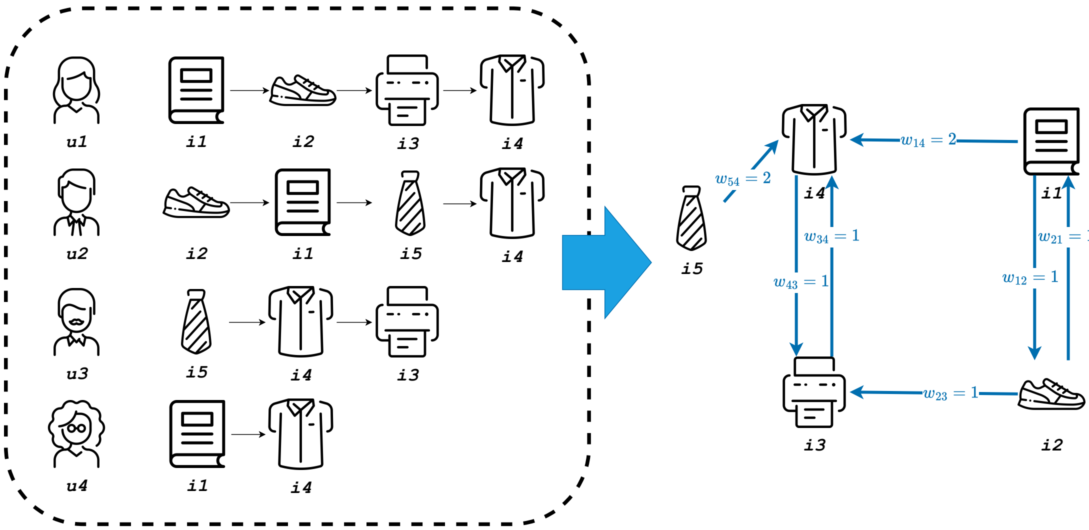

Fig. 2. Example of global transition graph construction from observed user behavior sequences. Edge weights correspond to the number of appearances of the item-item transitions in the user sequences. Note that the edge weights usually get normalized before used in training the GNN.

Current state-of-the-art models in SR comprise the usage of Recurrent Neural Networks (RNNs) [15, 26], Attention [11, 17] and Graph Neural Networks (GNNs) [33, 36, 39] to model the interaction sequences. In our research, we focus on GNN-approaches that construct global item graphs including all user-item interactions and learn the sequential item embedding from its neighborhood in the graph as opposed to methods that represent each user interaction sequence as a directed graph of items. An example of such global item graph construction is given in Figure 2. In this example the global item graph is constructed from four user interaction sequences where each transition between items increases the weight of the corresponding directed edge between those items (nodes) in the graph. Nevertheless, there are many possibilities for how to integrate the GNN framework into the task of sequential recommendation. We have analyzed relevant work in graph representation learning as well as sequential recommendation and identified following important gaps in graph-based SR that we aim to tackle in the proposed PhD project:

- (G1) Recent work in SR based on GNNs mostly ignore item features to improve the item representation in the model [21].
- (G2) Graph-based sequential models usually only consider the order in the interaction sequence and fully neglect the dwelling time or the time difference between interaction sequences [1, 5, 16].

(G3) GNNs are prone to the over-smoothing effect, where node representations converge to the same value over multiple layers, and also introduce additional computational complexity [3, 12].
(G4) Datasets in SR potentially include noisy relations (e. g., user misclicks on an item) and can introduce misleading information into the learning process. Filtering those noisy relations on the other hand leads to an increased data sparsity, which is already severely present in the original setting of SR [8, 38].
(G5) Current works in explainability of recommender systems rely on graph-based representations [2], but struggle to provide intuitive explanations due to the lack of feature-rich datasets.

To summarize, the core goals of our research comprise extending graph representations with additional feature information as well as improving the graph construction and learning process for informative item embeddings. Our research will investigate and aim to fill those described gaps in graph-based sequential recommendation.

2 RELATED WORK

In this section, we cover important research related to learning graph and node embeddings as well as sequential recommendation. Additionally, we will indicate potential gaps in the research and align them with our overall research goal of improving graph-based sequential recommendation.

2.1 Graph and Node Embeddings

Graph embedding aims to generate low-dimensional vector representations of the graph's nodes which preserve topology and leverage node features. Non-deep learning methods are mainly based on random walks to explore node neighborhoods [6, 22, 27]. With Graph Convolutional Networks (GCNs) [13, 30], more sophisticated graph embedding methods than random-walk-based approaches were introduced: To scale GCNs to large graphs, the layer sampling algorithm [7] generates embeddings from a fixed node neighborhood. Current state-of-the-art methods in self-supervised/semi-supervised learning of representations rely on contrastive methods which base their loss on the difference between positive and negative samples. Deep Graph Infomax (DGI) [31] contrasts node and graph encodings by maximizing the mutual information between them. Hassani and Khasahmadi [9] propose multi-view representation learning by contrasting first-order neighbor encodings with a general graph diffusion. Contrastive learning methods usually require a large number of negative examples and are, therefore, not scalable for large graphs. The approach by Thakoor et al. [29] learns by predicting substitute augmentations of the input and circumventing the need of contrasting with negative samples. In GraFN [14] a semi-supervised node classification framework leverages few labeled nodes to learn discriminative node representations and ensures nodes from the same class are grouped together.

The aforementioned methods can easily incorporate external item feature information as initial node embeddings, but are rarely used in the SR domain. Additionally, none of the existing methods appear to be specifically designed for the task of auto-tagging, which aims to predict relevant labels or tags for a given item [34] and is becoming increasingly important to generate or enrich recommendation datasets (cf. gaps (G1) and (G5)).

2.2 Sequential Recommendation

The initial phase of sequential recommendation focuses on discovering short-term item representations and interaction patterns. Markov decision processes are used in early works to model the interaction sequences. In FPMC [24], first-order Markov chains capture sequential patterns while matrix factorization models long-term user preferences. Also, convolutional neural networks (CNNs) have been found to be useful, where items are seen as images and short-term sequential patterns are learned via convolutional filters [28]. Xu et al. [41] combine CNNs with long-short-term memory to extract additional complex long-term dependencies. In HGN [18], a feature and instance gating mechanism is used to capture long- and short-term user interests. Other studies apply the attention mechanism to obtain and fuse different levels of interaction information [25, 42].

Self-attention and Transformer-based architectures are widely used for sequential recommendation models. SAS-Rec [11] applies the self-attention mechanism to identify relevant interactions from the user's history. Others use custom Transformer models to provide more personalized recommendations [4, 35]. In FDSA [43], heterogeneous features of items are integrated via feature sequences, and self-attention is applied to jointly model item and feature transition patterns. S^3-Rec [45] utilizes self-supervised learning to enhance the item representations via pre-training methods.

Hsu and Li [10] extract a local subgraph from a user-item pair and apply self-attention to encode long-term and short-term temporal patterns. MA-GNN [19] captures the item contextual information within a short-term period with a graph neural network and utilizes a shared memory network to model long-range dependencies. Work in [5] utilizes temporal graph representations to model continuous-time recommendation, where user and item embeddings are generated for any unseen future timestamps. Zhang et al. [44] extract augmented sequences representations from an item transition graph for a contrastive learning objective.

In session-based recommendation (SBR), a subtask of sequential recommendation, user profiles, and long-term interaction histories are no longer available. Most recent works in SBR are based on GNNs: As the first to introduce the concept of representing sessions as graphs, SR-GNN [37] models each session as a directed, unweighted graph and applies a gating mechanism to generate session representations. This work is extended by a self-attention mechanism in GCSAN [40] to effectively capture long-range dependencies. Incorporating collaborative knowledge into GNN-based methods leads to a new line of research. GCE-GNN [33] learns item embeddings on a session level as well as on a global level and uses a soft-attention mechanism to fuse the learned item representations. Chen and Wong [3] tackle the long-range dependency (over-smoothing) problem of session graphs by introducing a lossless encoding scheme and a shortcut graph attention layer. Xia et al. [38] introduce a dual-channel hypergraph to capture beyond-pairwise relations and apply self-supervised learning to maximize the mutual information between both session representations.

Recent research in the field of graph-based sequential recommendation has several limitations and room for improvement. Unlike earlier approaches that attempted to clean noisy data, there is little research on developing GNNs that can learn from noisy data without compromising performance (cf. gap (G4)). Additionally, there has been a recent push towards using more computationally complex GNN models that can better capture the structure and relationships within graphs. However, this increased complexity comes at the cost of greater computational resources (cf. gap (G3)). Another area of focus has been on addressing data sparsity, particularly in the context of contrastive learning (CL). Although CL has shown promise in learning representations from sparse data, there is still considerable room for improvement in this area (cf. gap (G4)).

3 RESEARCH OBJECTIVES

Our analysis of the field of graph-based sequential recommendation identified various gaps and issues as shown in the previous sections. To fill the previously identified research gaps, our research will seek to address the following research questions and provide valuable contributions in this field:

RQ1: How can graphs effectively be applied to incorporate item feature information in the setting of SR?. Graphs can be used in different ways in SR: To model the interaction sequences as separate graphs or to generate global item and user graphs based on the co-occurrences of item interactions, social networks, or knowledge graphs. Each node in a graph can be initially described via item feature information as opposed to simple one-hot encoding. As a first work to answer this research question, we proposed GCNext, a graph-based unsupervised learning approach to pre-train item embeddings with item feature information [21]. This pre-training approach can be used as an extension to any sequential model, be it nearest-neighbor methods or neural network models, in a plug-in fashion. To generate the pre-trained item embeddings, a global item co-occurrence graph is constructed from which the item embeddings are learned via a custom graph-encoder architecture based on attentional convolutions [30]. These graph-based item embeddings are used to initialize the item embedding tables of the corresponding neural network model. For extending nearest-neighbor methods we integrate the learned embeddings via session similarity computation based on the cosine distance of the graph-based embeddings. The evaluation performed on three session-based recommendation datasets showed that our approach significantly boosts the performance of the underlying sequential models.

RQ2: How can graph-based methods tackle the noisy and sparse data problem? Current graph-based methods [33, 36, 38] capture the topological structure of the sequence graph and rely on multi-hop information aggregation in GNNs to exchange information along edges. Consequently, graph-based models suffer from over-smoothing (node representations converge to the same value) if the number of layers is larger than three [3, 12]. Additionally, graph-based methods are prone to noisy item relations in the training data and introduce high complexity for large item catalogs. We propose to explicitly model the multi-hop information aggregation mechanism over multiple layers via shortest-path edges based on knowledge from the sequential recommendation domain. Our approach does not require multiple layers to exchange information and ignores unreliable item-item relations. Furthermore, to address inherent data sparsity, we apply supervised contrastive learning by mining data-driven positive and hard negative item samples from the training data. This work is submitted to the 17th ACM Conference on Recommender Systems and is currently under review.

RQ3: How can we effectively incorporate temporal information in the graph structure? User interaction sequences are usually not only ordered sequentially but also contain the timestamp per user-item interaction. From this information, we can infer the dwelling time or periodicity of items which potentially increases the recommendation performance. However, most of the current SR systems ignore this valuable information and only rely on the order of items in a sequence [11, 28]. To tackle this research question we plan to examine hyper-graphs in the setting of sequential recommendation and capture time information with personalized temporal point processes to model periodicity and mutual excitation of items.

RQ4: How can we incorporate item features to increase recommendation performance and explainability? As described in RQ1, each item can be described by features based on its content or meta-data. These features can support the learning process of the model by providing additional information per item. Additionally, known item features allow us to gain deeper knowledge about the insides of the model and explain its recommendation more coherently. Our research goal is to extend a large, feature-rich music dataset with emotional features based on techniques from the semi-supervised graph learning domain [13, 14] and use this dataset to generate explainable and more personalized recommendations.

4 CONCLUSION AND NEXT STEPS

In this paper, we analyzed recent works in graph-based sequential recommendation and identified various issues and research gaps as part of the ongoing PhD project. To contribute to this field of research, we formulate research questions and seek to answer them in a profound and rigorous fashion. As our research is in an advanced state and already investigated RQ1 and RQ2, we plan to tackle RQ3 and RQ4 as next steps.

ACKNOWLEDGMENTS

Special thanks to Dr. Eva Zangerle for the feedback and guidance in this project. This research was funded in whole, or in part, by the Austrian Science Fund (FWF) [P33526].

REFERENCES

[1] Veronika Bogina, Tsvi Kuflik, Dietmar Jannach, Mária Bieliková, Michal Kompan, and Christoph Trattner. 2023. Considering temporal aspects in recommender systems: a survey. *User Model. User Adapt. Interact.* 33, 1 (2023), 81–119.

[2] Hongxu Chen, Yicong Li, Xiangguo Sun, Guandong Xu, and Hongzhi Yin. 2021. Temporal meta-path guided explainable recommendation. In *Proceedings of the 14th ACM international conference on web search and data mining*. 1056–1064.

[3] Tianwen Chen and Raymond Chi-Wing Wong. 2020. Handling Information Loss of Graph Neural Networks for Session-based Recommendation. In *KDD '20: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, CA, USA, August 23-27, 2020*, Rajesh Gupta, Yan Liu, Jiliang Tang, and B. Aditya Prakash (Eds.). ACM, 1172–1180.

[4] Gabriel de Souza Pereira Moreira, Sara Rabhi, Jeong Min Lee, Ronay Ak, and Even Oldridge. 2021. Transformers4Rec: Bridging the Gap between NLP and Sequential / Session-Based Recommendation. In *RecSys '21: Fifteenth ACM Conference on Recommender Systems*. ACM, 143–153.

[5] Ziwei Fan, Zhiwei Liu, Jiawei Zhang, Yun Xiong, Lei Zheng, and Philip S Yu. 2021. Continuous-time sequential recommendation with temporal graph collaborative transformer. In *Proceedings of the 30th ACM international conference on information & knowledge management*. 433–442.

[6] Aditya Grover and Jure Leskovec. 2016. node2vec: Scalable Feature Learning for Networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 2016*. ACM, 855–864.

[7] William L. Hamilton, Zhitao Ying, and Jure Leskovec. 2017. Inductive Representation Learning on Large Graphs. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017*. 1024–1034.

[8] Qilong Han, Chi Zhang, Rui Chen, Riwei Lai, Hongtao Song, and Li Li. 2022. Multi-Faceted Global Item Relation Learning for Session-Based Recommendation. In *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2022*. ACM, 1705–1715.

[9] Kaveh Hassani and Amir Hosein Khas Ahmadi. 2020. Contrastive Multi-View Representation Learning on Graphs. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020 (Proceedings of Machine Learning Research, Vol. 119)*. PMLR, 4116–4126.

[10] Cheng Hsu and Cheng-Te Li. 2021. RetaGNN: Relational Temporal Attentive Graph Neural Networks for Holistic Sequential Recommendation. In *WWW '21: The Web Conference 2021, Virtual Event / Ljubljana, Slovenia, April 19-23, 2021*. ACM / IW3C2, 2968–2979.

[11] Wang-Cheng Kang and Julian J. McAuley. 2018. Self-Attentive Sequential Recommendation. In *IEEE International Conference on Data Mining, ICDM 2018*. IEEE Computer Society, 197–206.

[12] Thomas N. Kipf and Max Welling. 2017. Semi-Supervised Classification with Graph Convolutional Networks. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net.

[13] Thomas N. Kipf and Max Welling. 2017. Semi-Supervised Classification with Graph Convolutional Networks. In *5th International Conference on Learning Representations, ICLR 2017, Conference Track Proceedings*.

[14] Junseok Lee, Yunhak Oh, Yeonjun In, Namkyeong Lee, Dongmin Hyun, and Chanyoung Park. 2022. GraFN: Semi-Supervised Node Classification on Graph with Few Labels via Non-Parametric Distribution Assignment. In *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022*. ACM, 2243–2248.

[15] Jing Li, Pengjie Ren, Zhumin Chen, Zhaochun Ren, Tao Lian, and Jun Ma. 2017. Neural Attentive Session-based Recommendation. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, CIKM 2017*. ACM, 1419–1428.

[16] Jiacheng Li, Yujie Wang, and Julian J. McAuley. 2020. Time Interval Aware Self-Attention for Sequential Recommendation. In *WSDM '20: The Thirteenth ACM International Conference on Web Search and Data Mining, Houston, TX, USA, February 3-7, 2020*, James Caverlee, Xia (Ben) Hu, Mounia Lalmas, and Wei Wang (Eds.). ACM, 322–330.

[17] Qiao Liu, Yifu Zeng, Refuoe Mokhosi, and Haibin Zhang. 2018. STAMP: Short-Term Attention/Memory Priority Model for Session-based Recommendation. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018*. ACM, 1831–1839.

[18] Chen Ma, Peng Kang, and Xue Liu. 2019. Hierarchical Gating Networks for Sequential Recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2019, Anchorage, AK, USA, August 4-8, 2019*, Ankur Teredesai, Vipin Kumar,

Ying Li, Rómer Rosales, Evimaria Terzi, and George Karypis (Eds.). ACM, 825–833.

[19] Chen Ma, Liheng Ma, Yingxue Zhang, Jianing Sun, Xue Liu, and Mark Coates. 2020. Memory Augmented Graph Neural Networks for Sequential Recommendation. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020.* AAAI Press, 5045–5052.

[20] Prem Melville and Vikas Sindhwani. 2010. Recommender systems. *Encyclopedia of machine learning* 1 (2010), 829–838.

[21] Andreas Peintner, Marta Moscati, Emilia Parada-Cabaleiro, Markus Schedl, and Eva Zangerle. 2022. Unsupervised Graph Embeddings for Session-based Recommendation with Item Features. In *CARS: Workshop on Context-Aware Recommender Systems (RecSys '22).*

[22] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. 2014. DeepWalk: online learning of social representations. In *The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '14.* ACM, 701–710.

[23] Massimo Quadrana, Paolo Cremonesi, and Dietmar Jannach. 2018. Sequence-aware recommender systems. *ACM Computing Surveys (CSUR)* 51, 4 (2018), 1–36.

[24] Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. 2010. Factorizing personalized Markov chains for next-basket recommendation. In *Proceedings of the 19th International Conference on World Wide Web, WWW, 2010.* ACM, 811–820.

[25] Qiaoyu Tan, Jianwei Zhang, Ninghao Liu, Xiao Huang, Hongxia Yang, Jingren Zhou, and Xia Hu. 2021. Dynamic Memory based Attention Network for Sequential Recommendation. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021.* AAAI Press, 4384–4392.

[26] Yong Kiam Tan, Xinxing Xu, and Yong Liu. 2016. Improved Recurrent Neural Networks for Session-based Recommendations. In *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems, DLRS@RecSys 2016.* ACM, 17–22.

[27] Jian Tang, Meng Qu, Mingzhe Wang, Ming Zhang, Jun Yan, and Qiaozhu Mei. 2015. LINE: Large-scale Information Network Embedding. In *Proceedings of the 24th International Conference on World Wide Web, WWW 2015.* ACM, 1067–1077.

[28] Jiaxi Tang and Ke Wang. 2018. Personalized Top-N Sequential Recommendation via Convolutional Sequence Embedding. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, WSDM 2018, Marina Del Rey, CA, USA, February 5-9, 2018*, Yi Chang, Chengxiang Zhai, Yan Liu, and Yoelle Maarek (Eds.). ACM, 565–573.

[29] Shantanu Thakoor, Corentin Tallec, Mohammad Gheshlaghi Azar, Remi Munos, Petar Veličković, and Michal Valko. 2021. Bootstrapped Representation Learning on Graphs. In *ICLR 2021 Workshop on Geometrical and Topological Representation Learning.*

[30] Petar Velickovic, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. 2018. Graph Attention Networks. In *6th International Conference on Learning Representations, ICLR 2018, Conference Track Proceedings.* OpenReview.net.

[31] Petar Velickovic, William Fedus, William L. Hamilton, Pietro Liò, Yoshua Bengio, and R. Devon Hjelm. 2019. Deep Graph Infomax. In *7th International Conference on Learning Representations, ICLR 2019.* OpenReview.net.

[32] Shoujin Wang, Liang Hu, Yan Wang, Longbing Cao, Quan Z. Sheng, and Mehmet Orgun. 2019. Sequential Recommender Systems: Challenges, Progress and Prospects. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19.* 6332–6338.

[33] Ziyang Wang, Wei Wei, Gao Cong, Xiao-Li Li, Xian-Ling Mao, and Minghui Qiu. 2020. Global context enhanced graph neural networks for session-based recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval.* 169–178.

[34] Minz Won, Andres Ferraro, Dmitry Bogdanov, and Xavier Serra. 2020. Evaluation of CNN-based Automatic Music Tagging Models. *CoRR* abs/2006.00751 (2020).

[35] Liwei Wu, Shuqing Li, Cho-Jui Hsieh, and James Sharpnack. 2020. SSE-PT: Sequential Recommendation Via Personalized Transformer. In *RecSys 2020: Fourteenth ACM Conference on Recommender Systems, Virtual Event, Brazil, September 22-26, 2020.* ACM, 328–337.

[36] Shu Wu, Yuyuan Tang, Yanqiao Zhu, Liang Wang, Xing Xie, and Tieniu Tan. 2019. Session-Based Recommendation with Graph Neural Networks. In *Proceedings of the AAAI Conference on Artificial Intelligence.* 346–353.

[37] Shu Wu, Yuyuan Tang, Yanqiao Zhu, Liang Wang, Xing Xie, and Tieniu Tan. 2019. Session-Based Recommendation with Graph Neural Networks. *Proceedings of the AAAI Conference on Artificial Intelligence* 33, 01, 346–353.

[38] Xin Xia, Hongzhi Yin, Junliang Yu, Yingxia Shao, and Lizhen Cui. 2021. Self-Supervised Graph Co-Training for Session-based Recommendation. In *CIKM '21: The 30th ACM International Conference on Information and Knowledge Management.* ACM, 2180–2190.

[39] Chengfeng Xu, Pengpeng Zhao, Yanchi Liu, Victor S. Sheng, Jiajie Xu, Fuzhen Zhuang, Junhua Fang, and Xiaofang Zhou. 2019. Graph Contextualized Self-Attention Network for Session-based Recommendation. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19.* 3940–3946.

[40] Chengfeng Xu, Pengpeng Zhao, Yanchi Liu, Victor S. Sheng, Jiajie Xu, Fuzhen Zhuang, Junhua Fang, and Xiaofang Zhou. 2019. Graph Contextualized Self-Attention Network for Session-based Recommendation. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019.* ijcai.org, 3940–3946.

[41] Chengfeng Xu, Pengpeng Zhao, Yanchi Liu, Jiajie Xu, Victor S. Sheng, Zhiming Cui, Xiaofang Zhou, and Hui Xiong. 2019. Recurrent Convolutional Neural Network for Sequential Recommendation. In *The World Wide Web Conference, WWW 2019, San Francisco, CA, USA, May 13-17, 2019.* ACM, 3398–3404.

[42] Lu Yu, Chuxu Zhang, Shangsong Liang, and Xiangliang Zhang. 2019. Multi-Order Attentive Ranking Model for Sequential Recommendation. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019.* AAAI Press, 5709–5716.

[43] Tingting Zhang, Pengpeng Zhao, Yanchi Liu, Victor S. Sheng, Jiajie Xu, Deqing Wang, Guanfeng Liu, and Xiaofang Zhou. 2019. Feature-level Deeper Self-Attention Network for Sequential Recommendation. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019*. ijcai.org, 4320–4326.

[44] Yixin Zhang, Yong Liu, Yonghui Xu, Hao Xiong, Chenyi Lei, Wei He, Lizhen Cui, and Chunyan Miao. 2022. Enhancing Sequential Recommendation with Graph Contrastive Learning. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022*, Luc De Raedt (Ed.). ijcai.org, 2398–2405.

[45] Kun Zhou, Hui Wang, Wayne Xin Zhao, Yutao Zhu, Sirui Wang, Fuzheng Zhang, Zhongyuan Wang, and Ji-Rong Wen. 2020. S3-Rec: Self-Supervised Learning for Sequential Recommendation with Mutual Information Maximization. In *CIKM '20: The 29th ACM International Conference on Information and Knowledge Management, Virtual Event, Ireland, October 19-23, 2020.* ACM, 1893–1902.