Recent Advances in Heterogeneous Relation Learning for Recommendation

Chao Huang

University of Hong Kong, Hong Kong
chaohuang75@gmail.com

Abstract

Recommender systems have played a critical role in many web applications to meet user's personalized interests and alleviate the information overload. In this survey, we review the development of recommendation frameworks with the focus on heterogeneous relational learning, which consists of different types of dependencies among users and items. The objective of this task is to map heterogeneous relational data into latent representation space, such that the structural and relational properties from both user and item domain can be well preserved. To address this problem, recent research developments can fall into three major lines: social recommendation, knowledge graph-enhanced recommender system, and multi-behavior recommendation. We discuss the learning approaches in each category, such as matrix factorization, attention mechanism and graph neural networks, for effectively distilling heterogeneous contextual information. Finally, we present exploratory outlook to highlight several promising directions and opportunities in heterogeneous relational learning frameworks for recommendation.

1 Introduction

In the era of information explosion, recommender systems are designed to help allieveate the information explosion and meet users personalized interests [He *et al.*, 2017]. For example, recommending products to users on e-commerce platforms (*e.g.*, Amazon, Ebay, Taobao) [Huang *et al.*, 2019], or suggesting movies for users on online video sites (*e.g.*, Netflix, Youtube) [Covington *et al.*, 2016]. To accurately capture user preference, one of most popular recommendation architectures is collaborative filtering (CF), which leverages the historical co-interact patterns of users on items and make recommendations [Mnih and Salakhutdinov, 2007].

In our computational online services, practical recommender systems are often associated with heterogeneous structural relations, *e.g.*, users' social connections [Wu *et al.*, 2019b; Xu *et al.*, 2020], knowledge graph of items [Cao *et al.*, 2019; Ma *et al.*, 2019b], multiplex user-item interactions [Xia *et al.*, 2020; Jin *et al.*, 2020]. Such heterogeneous relational

data provide opportunities for researchers and practitioners to not only understand and investigate user preferences better, but also present challenges to learn insightful information from them. For example, the incorporation of cross-user social relations could help characterize user preference, based on the social influences among users [Huang *et al.*, 2021b; Fan *et al.*, 2019]. In addition, knowledge graph could provide connectivity information to capture the semantic relatedness between items [Zhu *et al.*, 2020; Wang and others, 2019]. Hence, distilling and aggregating knowledge from such heterogeneous structural data would benefit recommendation designs with better personalization.

In the recommendation with heterogeneous relational context, the key challenge lies in the effective modeling of complex dependent structures from a diverse set of heterogeneous relations. In addition to the difficulty of discovering insightful knowledge from heterogeneous data sources, it is a significant challenge to effectively aggregate the heterogeneous knowledge in a unified way, to facilitate various user modeling applications. Existing heterogeneous relational learning frameworks should be able to address two important questions which corresponds to the aforementioned challenges: (1) How to effectively perform the relation representation learning from the heterogeneous structural data? (2) How to perform the information fusion based on the extracted knowledge via automatic machine learning frameworks, and endow the user preference modeling paradigms with heterogeneous context incorporation?

In this survey, we focus on the recent advances and new trends in heterogeneous relation learning for recommendation. A rich body of learning methods have been applied to solve this challenge in recommender systems, based on various techniques, including attention mechanism [Wu *et al.*, 2019c; Chen *et al.*, 2019a], graph neural network [Xia *et al.*, 2021a; Wang and others, 2019], knowledge base embedding [Cao *et al.*, 2019], and multi-task learning [Chen *et al.*, 2020; Gao *et al.*, 2019]. In the following, we break down this investigation into three parts: i) We identify key challenges of capturing heterogeneous relations for modeling user preferences; ii) We categorize the existing solutions into three groups, *i.e.*, social recommendation, knowledge graph-enhanced recommender systems, and multi-behavior recommendation; iii) We point out promising research directions in heterogeneous relation learning for recommendation. To our

best knowledge, this paper is the first survey that pays special attention to recent advances in user modeling and personalization with heterogeneous relational context.

2 Collaborative Filtering Based Recommendation

Collaborative filtering has emerged as the most popular paradigm to build recommender systems. At its core is to presume that behaviorally similar users likely to share similar interests over items, based on users' historical interaction behaviors (*e.g.*, clicks, purchase) [He *et al.*, 2017; Hofmann and Puzicha, 1999]. The most common collaborative filtering framework lies in representing users and items with low-dimensional latent features (*i.e.*, embeddings), and then making predictions based on the learned embeddings vectors [Han *et al.*, 2018; Liang *et al.*, 2018].

2.1 Matrix Factorization

There exist various types of collaborative filtering techniques which design different model structures for modeling user-item interactions. Specifically, matrix factorization is an early solution to decompose sparse matrix to low-rank latent embeddings, by retaining the dependencies of interactions among users [Mnih and Salakhutdinov, 2007]. For example, Singular Value Decomposition(SVD) [Billsus *et al.*, 1998] infers the missing elements in user-item rating matrix by factorizing it into a product of two low-rank matrices corresponding to user and item dimensions. Later on, several probabilistic variations are proposed to extend matrix factorization with the integration of the sophisticated Bayesian inference, such as Probabilistic Matrix Factorization (PMF) [Mnih and Salakhutdinov, 2007], Bayesian Probabilistic Matrix Factorization (BPMF), and Maximum Margin Matrix Factorization (MMMF) [Srebro *et al.*, 2005; Rennie and Srebro, 2005].

In a typical recommendation scenario, we have a set of users $u_i \in \mathcal{U}$ and items $v_j \in \mathcal{V}$ with the size of I and J, respectively. The user-item interaction matrix $\mathbf{Y}$ is generated based on users' implicit feedback on items, in which $y_{i,j} = 1$ if user u_i has interacts with item v_j (*e.g.*, click, watch, or purchase) and $y_{i,j} = 0$ otherwise. Matrix factorization models aim to project user and item into low-dimensional latent feature vectors, which can be formally represented as:

$$\hat{y}_{i,j} = f(i, j | \Theta) \tag{1}$$

where $\hat{y}_{i,j}$ is the predicted probability of interaction $y_{i,j}$ between user u_i and item v_j. The learned model parameters are denoted by Θ. $f(\cdot)$ is the encoding function which maps input user interactions $\mathbf{Y}$ to the estimated probability $\hat{y}_{i,j}$.

2.2 Neural Collaborative Filtering Architecture

With the advent of deep learning in various domains (*e.g.*, computer vision, nature language processing), neural network techniques have been utilized to enhance the collaborative filtering paradigm with non-linear feature interactions. To be specific, NCF [He *et al.*, 2017] and DMF [Xue *et al.*, 2017] leverage multi-layer perceptron to endow the neural collaborative filtering for user-item interaction modeling

with non-linearities. In CDAE [Wu *et al.*, 2016] and AutoRec [Sedhain *et al.*, 2015], the autoencoder architecture has been used as the mapping function to learn latent representations with the reconstruction of the user-item interaction input. In addition, inspired from restricted Boltzmann machines [Salakhutdinov *et al.*, 2007; Larochelle and Bengio, 2008], neural autoregressive models are adopted to distribution estimation for collaborative filtering [Zheng *et al.*, 2016; Du *et al.*, 2018]. These neural models perform autoregression for user and item dimensions, so as to capture the inter-correlations between users and items in an explicit way.

Motivated by the development of graph neural network and its strength in relational learning over graph-structured data, many efforts propose to guide the embedding learning over the generated user-item interaction graph [Klicpera *et al.*, 2019]. Recently emerged graph neural networks (GNNs) design various message passing schemes to encourage neighboring nodes to have similar representations [Kipf and Welling, 2017; Veličković *et al.*, 2018]. To adapt the GNNs to the collaborative filtering scenario, some studies employ the graph convolutions to aggregate embeddings of neighboring users or items, to refine the representation of the target node, like NGCF [Wang *et al.*, 2019b], PinSage [Ying *et al.*, 2018], DGCF [Wang *et al.*, 2020], GC-MC [Berg and others, 2017].

Heterogeneity of Relations. However, these above collaborative filtering models have thus far focused on representation learning for homogeneous relations for user modeling. As a matter of fact, practical recommendation scenarios involve a diversity of relations from both user and item dimensions [Shi *et al.*, 2018] (as shown in Figure 1). Given different recommendation scenarios, methods can be categorized as different dimensions of heterogeneous relation learning:

- **Social Connections among Users**. The increasing popularity of online social platforms enable the connections between users with similar interests, by integrating item recommendations and social network services [Wu *et al.*, 2019c], *e.g.*, preference of users on items may be influenced by their friends. Hence, one line of heterogeneous relational learning for recommender system is *Social Recommendation* which leverages social information to alleviate the data sparsity and improve the recommendation quality, with the joint modeling of user-user and user-item relations [Jiang *et al.*, 2012; Yang *et al.*, 2016].

- **Knowledge-aware Dependencies among Items**. In real-life applications, there exist explicit relations among different items with the consideration of external knowledge (*e.g.*, movie's director, product category, or venue's location) [Xin *et al.*, 2019]. Such item-wise dependencies can be used to construct the knowledge graph to present the heterogeneous structured information and reflect the semantic relatedness between items. The rich semantics and topology of knowledge graphs could be incorporated into the user embedding process as important side information, by exploiting the knowledge-aware dependency across items [Ma *et al.*, 2019b].

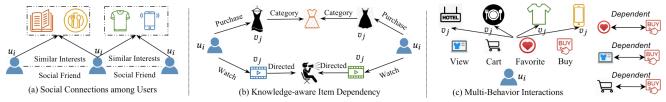- **Multi-Behavior User-Item Interactions**. In practice,

Figure 1: Illustration of relation heterogeneity learning for recommender systems.

users exhibit different interaction behaviors on items depending on the intention context [Chen *et al.*, 2020]. Let us consider the e-commerce platform as an example: users can click products when they are interested in, or tag these items as favourites if they like them, or make final purchases given the products fulfill their needs [Tanjim *et al.*, 2020]. To provide better purchase forecasting for a user, it is beneficial to investigate not only which products he/she has purchased before, but also what this user has browsed and tagged as his/her favorite items [Xia *et al.*, 2021b]. In such case, user-item interactions are heterogeneous in nature, involving different types of user behaviors (*e.g.*, click, tag-as-favorite, purchase).

To provide a better understanding, Figure 1 presents the illustrated examples corresponding to the above recommendation scenarios with heterogeneous relational context.

3 Heterogeneous Contextual Relation Learning for User Representation

In this section, we review the recent advances for user modeling and personalization with the modeling heterogeneous relational context from three perspectives (stated in Section 2).

3.1 Social Recommendation

Social recommendation aims to probe social relations among users into user-item interaction modeling paradigm, based on the observed phenomena that users' preference is likely to be affected by their social connected friends [Tang *et al.*, 2013]. In social-aware recommender systems, there exist two types of relations: user-user social connections (*e.g.*, friendship, or family) and user-item interactions (*e.g.*, click, or purchase).
Challenge. The key challenge of social recommendation lies in the effectively transforming the user-user and user-item relational structure into latent factor spaces, such that the semantics of heterogeneous relations can be well preserved.

Social-aware Factorization Methods. In light of the above challenge, the exploration of social recommendation can date back to the development of social relation-enhanced matrix factorization model. In particular, SoReg [Ma *et al.*, 2011] considers social signals as the regularization term to constrain the optimized objective in the matrix factorization architecture. Jamali *et al.* [Jamali and others, 2010] proposes to augment the matrix factorization with the trust influence modeling, so as to capture the social-aware user preference. Ma *et al.* [Ma *et al.*, 2008] develops a factor analysis method–SoRec which is built on the probabilistic matrix factorization, to jointly investigate users' social connections

and interactions. Furthermore, a social collaborative filtering model named TrustMF [Yang *et al.*, 2016] is developed to use the trust relations among users in addressing the data sparsity issue of traditional recommender systems. Nevertheless, those social-aware factorization methods model the two-dimensional interaction of user and item hidden factors, with the assumption of the linear relations between each dimensional latent space. Hence, their model performance can be hindered by the simply linear combination of latent factors [He *et al.*, 2017].

Attentive Social Recommender Systems. Attention mechanism has shown its effectiveness in a variety of tasks, *e.g.*, machine translation [Indurthi *et al.*, 2019] and human activity recognition [Ma *et al.*, 2019a]. The key idea of attention mechanism is to identify the influential parts with the attentive weights from the input data. Along with this line, several studies seek to capture the relationships between users based on neural attentive architectures. For instance, Chen *et al.* [Chen *et al.*, 2019a] proposes a two-stage attentional model (SAMN) to differentiate users' social relations. In SAMN framework, an aspect-level attention module is firstly introduced to generate the user-friend specific relation representation, and then another attention layer is proposed to capture the friend-level differences. Additionally, with the introduction of attention mechanism, an adaptive transfer learning scheme is designed to model the interplay between user and item domain for social recommendation [Chen *et al.*, 2019b]. To capture social contextual signals, Wu *et al.* [Wu *et al.*, 2019a] proposes a hierarchical attention approach to model factors (*i.e.*, past behaviors, social relations, user admiration) which impacts user preferences.

Graph Neural Network for Social Recommendation. Advances in graph neural networks (GNNs) have shed light on the development of social recommendation solutions [Huang *et al.*, 2021b], due to the strong capability of learning on graph structure data. Relations in social recommendation can be naturally represented as two explicit graph-based heterogeneous structures: i) user-item interactions; ii) user-user social connections. Along with this line, many efforts have developed graph neural networks with information propagation paradigm, to inject social signals into the user-item interaction modeling. We show the overall model flow of graph neural network-based social recommender systems in Figure 2. In general, current methods can be summarized into the following graph neural paradigms.

(i) **Graph Convolutional Network-based Models**. Inspired
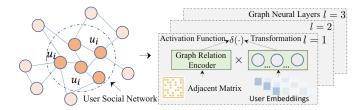
Figure 2: Graph neural network for social recommendation with relation aggregation under message passing paradigm.

by the prevalence of spectral graph learning framework–Graph Convolutional Network (GCN), DiffNet [Wu *et al.*, 2019b] simulates the influence process between users, by recursively propagating information over the social network for user and item representation refinement. The designed influence diffusion neural layers can be incorporated into the collaborative filtering architecture, to capture the heterogeneous relational context for social-aware recommendation. In RecoGCN [Xu *et al.*, 2019], the GCN is adopted to aggregate heterogeneous features over the constructed heterogeneous graph with the semantic-aware meta-paths from user-user and user-item edges. To address the sparsity problem, a GCN-based autoencoder is introduced to augment the data with the adversarial learning strategy [Yu *et al.*, 2020].

(ii) **Graph Attention Network-based Models**. To discriminate the importance of neighbouring feature aggregation, Graph Attention Network (GAT) [Veličković *et al.*, 2018] is proposed as a non-spectral learning method that leverages the node spatial information for embedding propagation. To alleviate the limitation of the mean-based information aggregator, GraphRec [Fan *et al.*, 2019] integrates attention mechanism into the graph neural network to model heterogeneous user latent factors from social space. To endow the social recommender system to capture multifaceted social relations, Wu *et al.* [Wu *et al.*, 2019c] devises a dual-stage graph attention to dynamic weigh social effects from user and item domain context. To inject temporal context into social recommendation, GAT is adopted to cooperate with a recurrent neural network to capture the time-aware social influence and user interests [Song *et al.*, 2019].

General GNN Social Recommendation Framework. Suppose $\mathbf{H}^l[u_i]$ is the representation of user u_i at the (l)-th GNN layer, the message passing paradigm from the (l)-th layer to the $(l+1)$-th layer can be formally represented as:

$$\mathbf{H}^{(l+1)}[u_i] = \text{Aggre}\Big(\text{Extract}(\mathbf{H}^l[u_i], \mathbf{H}^l[v_j]), \mathbf{H}^l[u_{i'}]\Big) \quad (2)$$

where $N_{u_i}^s$ and $N_{u_i}^v$ denote the set of user u_i's social connected users and items, respectively. $u_{i'} \in N_{u_i}^s$; $v_j \in N_{u_i}^v$. The key GNN operators: Extract$(\cdot)$ represents the heterogeneous neighbor information extractor from user-user and user-item dependent context. Aggre$(\cdot)$ is the aggregation function to fuse heterogeneous signals, including the mean and attention-based operator corresponding to the GCN and GAT message passing methods, respectively.

Summary. To adapt to the heterogeneous relational learning scenarios, all existing social recommendation models employ different structural information aggregation schemes, *e.g.*, attention mechanism, graph convolution and graph attention network. The nature of graph neural networks is able to jointly maintain user-user and user-item relational structures, for generating contextualized user and item embeddings.

3.2 Knowledge Graph-enhanced Recommendation

Knowledge Graphs (KGs) brings up a large amount of commercial benefits for various applications, ranging from question answering [Yang *et al.*, 2019a] and information extraction [Yang *et al.*, 2019b]. KGs provide world's factual information with heterogeneous relational structures (entity-relation-entity triplets), in which nodes and edges represent the entities and the corresponding entity-wise relations, respectively. In recent years, incorporating knowledge graph information into the recommendation has attracted much attention, to help improve the performance of recommender systems [Zhu *et al.*, 2020]. Such knowledge side information presents heterogeneous relations corresponding to different item attributes, which i) captures the rich semantic relatedness among items; ii) improves the explainability of recommender systems with knowledge graph connections.

Challenge. The main challenge in incorporating knowledge graph into recommendation lies in how to capture the semantic relatedness between items, and transform such underlying dependencies into the user/item representation paradigm with the heterogeneous relational context.

To address the aforementioned challenge, existing knowledge graph-enhanced recommendation techniques can be grouped into three main types: path-based approaches, regularization-based methods, and hybrid learning methods.

Path-based Approaches. The path-based approaches attempt to generate entity paths between graph nodes based on their high-order connections (as illustrated in Figure 3). Early personalized entity recommendation method [Yu *et al.*, 2014] considers the meta-path heterogeneous graph as external knowledge to alleviate the data scarcity issue. Motivated by the deep learning methods over heterogeneous information network, a co-attention mechanism is proposed to inject the meta-path based context into the recommendation architecture, so as to characterize the three-way knowledge-based interaction with the triple form of (user, meta-path, item) [Hu *et al.*, 2018]. To effectively identify useful meta-graph based features (*e.g.*, user→review→business), Zhao *et al.* [Zhao *et al.*, 2017] develops a group lasso regularized factorization machine, to characterize the knowledge-aware semantics. However, the success of most path-based methods largely relies on the effective design of meth paths for modeling heterogeneous relations among users and items, which requires domain-specific knowledge. Hence, such process may hardly be adapted to the complex knowledge graph structures with diverse categories of entities and relations.

Regularization-based Methods. Inspired by the strength of knowledge graph embedding techniques (*e.g.*, TransE [Bordes *et al.*, 2013]), several regularization-based methods
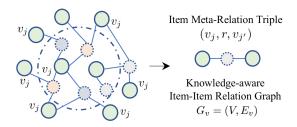
Figure 3: Knowledge graph-guided item meta-relations.

propose to regularize the user representation learning in recommendation with the designed additional knowledge graph loss. Following this line, regularization-based methods integrate the tasks of knowledge graph learning and recommendation under a multi-task learning paradigm. For example, KTUP [Cao *et al.*, 2019] simultaneously models item recommendation and knowledge graph completion, to learn user preference with the consideration of knowledge-aware relations. To explain the recommendation results, Ai *et al.* [Ai *et al.*, 2018] generates the interpretations based on embedded knowledge base, with the preservation of heterogeneous dependent structures. Due to the lack of modeling high-order connectivity between users and items, these methods cannot capture the long-range dependencies for user-item interactions and knowledge graphs.

Hybrid Learning Methods. Another line of KG-enhanced recommendation models learn user/item representations with the joint modeling of i) path-based entity dependence; and ii) knowledge graph structural semantics. For instance, KARN [Zhu *et al.*, 2020] proposes an attentive recurrent network to incorporate knowledge graph into the sequence learning framework of user's click records. A joint optimization framework is developed to extract inductive rules from item-centric knowledge graphs [Ma *et al.*, 2019b]. In addition, graph neural networks have been utilized for encoding the knowledge-aware relations to transform the knowledge graph into user interest modeling, such as KGNN-LS [Wang and others, 2019] and KGAT [Wang *et al.*, 2019a]. In particular, KGNN-LS is a knowledge-aware graph neural model with the integration of label smoothness regularization. This work is on the basis of the assumption that users are likely to share similar preferences on items which are connected in the knowledge graph. KGAT enables the high-order knowledge-aware relation learning under the graph attention architecture.

Summary. Overall, the limitation of path-based methods lies in the requirement of domain-specific knowledge for effective meta-path construction, which hinders the flexibility of those KG-enhanced recommendation solutions. The success of regularization-based methods replies on the output of knowledge base embedding. To effectively leverage external knowledge, it is important to design a main representation space which effectively bridges the recommendation and knowledge graph learning. Compared with the path-based approaches and regularization-based methods, the hybrid learning models have a better capability with respect to discovering complex knowledge-aware heterogeneous re-

lations. However, by involving the recursive embedding propagation, GNN-based methods have large computational cost by exploring the high-order connectivity among entities. Hence, how to achieve a nice trade-off between the modeling of long-range collaborative knowledge and model efficiency, remains a significant challenge.

3.3 Multi-Behavior Recommender System

In real-world applications, users often interact with items based on various types of behaviors, *e.g.*, click, tag-as-favorite, add-to-cart and purchase in online retailing systems [Xia *et al.*, 2020]. Hence, it is important to capture the collaborative signals from multi-behavioral patterns, to characterize diverse user-item relationships. In the behavior-aware recommendation settings, the behavior type we aim to forecast is regarded as *target behavior*, such as purchase in online retailing platforms or positive rating feedback in online review sites. Other types of user behavior are regarded as *auxiliary behavior*. For example, in e-commerce platforms, both add-to-cart and tag-as-favorite behavior could serve as indicators to model the purchase patterns of users.

Challenge. In the multi-behavior recommendation scenario, different types of behavior may reflect the interactive pattern between users and items from different perspectives. It is important to capture the unique context and semantics of heterogeneous user-item interactions when modeling the multi-behavioral relationships [Xia *et al.*, 2021a].Because diverse types of user-item interactions can be dependent in a complex way, it is important to design an effective learning model to capture the underlying multi-behavior dependencies.

Existing multi-behavior recommendation methods can be categorized into two key dimensions: multi-task behavior learning frameworks and graph-based recommender systems.

Multi-Task Behavior Learning Frameworks. The first type of studies tackle the multi-behavior recommendation under a multi-task learning architecture. Particularly, the modeling of multi-typed behavioral interactions is conducted with multiple tasks based on the shared embedding space. For example, To account for multi-typed user behaviors, Gao *et al*. [Gao *et al.*, 2019] integrates the neural collaborative filtering, to learn separate interaction encoding function for individual type of behaviors under a multi-task learning paradigm. Furthermore, in EHCF [Chen *et al.*, 2020], a transfer-based model is developed to correlate the forecasting of each type of user-item interaction. With the designed transfer learning approach, EHCF can not only capture heterogeneous relational structure between user and item, but also alleviate the data sparsity issue for inactive users.

Graph-based Multi-Behavior Recommender Systems. The second type of multi-behavior recommender systems are built upon the graph neural architectures. These methods firstly construct the heterogeneous user-item interaction graph by differentiating behaviors (as shown in Figure 4). Edges in the generated graph represent different types of user behaviors (*e.g.*, click, tag-as-favorite, purchase). Then, they leverage different graph relation encoders (*e.g.*, graph convolutional network and graph attention mechanism), to aggre-

gate multi-relational contextual signals during the message passing procedure. For example, MBGCN [Jin *et al.*, 2020] is a behavior-aware information propagation framework based on graph convolutional operations, which captures the semantics of multi-typed behaviors. Through extensive experiments, authors demonstrate the performance superiority of multi-behavior models as compared to the recommender systems with singular type of behaviors.
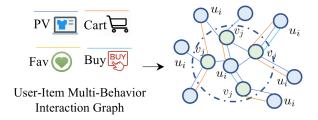

Figure 4: Multi-behavior user-item interaction graph.

To effectively and explicitly discriminate the influences of different types of behaviors, KHGT [Xia *et al.*, 2021a] builds a graph-structured transformer network to determine which types of user-item interactions contribute more in assisting the prediction of target behavior (*e.g.*, purchase). In those graph-based multi-behavior recommendation methods, the behavior-aware embeddings are propagated over the multi-behavior interaction graph to refine user and item representations with the high-order relational structures.

Summary. Different from the heterogeneous relations extracted from user or item domain as side information (*e.g.*, social connections, item knowledge-aware relations), edges between node pairs in the multi-behavior recommendation present the relation multiplicity. Such multiplex topological structures between user and item are investigated in the recently emerged multi-behavior recommender systems.

4 Future Challenges and Opportunities

This section presents a few potential opportunities for recommendation with heterogeneous relational context.

- **Learning to Recommend with Multi-Modal Representations**. The prevalence of web services allow the recommender systems to access online multi-modal content, such as item images/videos or textural description, user profile and reviews [Wei *et al.*, 2020]. Along with this line, learning informative multi-modal representations of users and items with the incorporation of multi-modal content data into heterogeneous relational learning paradigms, become an important theme of the recommendation frameworks.

- **Learning with Data Scarcity**. The effectiveness of current methods relies on sufficient relational data, and may not be able to learn quality embeddings when user behavior or heterogeneous relation data are scarce. They often require a sufficient amount of user-item interactions, or knowledge graph connections, so that the encoded representations is able to well capture the dependent structures between entities. However, many real-life systems involve a limited

amount of relational data. Therefore, to build effective heterogeneous relational learning models, it is important to develop effective learning architectures to perform the heterogeneous relation learning with data scarcity.

- **Temporal Context Injection**. In real-life online platforms, the temporal patterns of user-item interactions and the underlying time-aware behavior relationships, serve as a key feature dimension in many recommender systems [Huang *et al.*, 2021a]. Hence, it is necessary to design dedicated learning frameworks for modeling dynamic heterogeneous relational data with complex temporal context. However, recommendation models with temporal heterogeneous relations still face several challenges: i) from the local perspective, connections among users and items are inherently dynamic; ii) from the global perspective, it is important to capture the temporal long-term structures with the relational heterogeneity, which calls for dedicated efforts.

- **Interpretable User Modeling**. The model interpretation ability with heterogeneous relational learning is less explored in current recommender systems. Specifically, most of current approaches only focus on forecasting unknown user-item interactions, but ignore the inference of casual factors for predicted results. In practical recommendation scenarios, identifying the underlying factors which influence user interaction behaviors, is beneficial to show result explanations and further improve the recommendation performance. For example, performing accurate reasoning on cross-type behavior dependent structures to infer user preferences towards target items, will endow the multi-behavior recommender systems with the capacity of providing informative explanations for identifying multi-dimensional factors. Therefore, it would be interesting and desirable to leverage advanced techniques (*e.g.*, causal inference [Pearl, 2019]) and develop interpretable paradigms for heterogeneous relation learning-based recommender systems.

- **Real-Time Computational Frameworks**. Notwithstanding the promising performance of existing heterogeneous relation learning-based recommendation frameworks, most of them operate on the static observed user behaviors and side features. This makes those methods less suited to the practical applications, where new heterogeneous relational data (*e.g.*, clicks and purchases of users, dynamic user connections) arrive continually over time. Hence, real-time computational frameworks are important because they allow the heterogeneous relational recommender systems to infer user preference from streaming data.

5 Conclusion

This work introduces the problem of heterogeneous relational learning for recommendation and discuss its corresponding challenges. We present the recent advances in addressing these challenges. This survey not only provides a comprehensive study of user modeling techniques with heterogeneous context for researchers and practitioners in related fields, but also shed light on designing solutions to harness the power of heterogeneous data, and distill useful knowledge for various real-life user modeling and personalization applications.

References

[Ai *et al.*, 2018] Qingyao Ai, Vahid Azizi, et al. Learning heterogeneous knowledge base embeddings for explainable recommendation. *Algorithms*, 11(9):137, 2018.

[Berg and others, 2017] Rianne van den Berg et al. Graph convolutional matrix completion. *KDD Workshop*, 2017.

[Billsus *et al.*, 1998] Daniel Billsus, Michael J Pazzani, et al. Learning collaborative information filters. In *ICML*, volume 98, pages 46–54, 1998.

[Bordes *et al.*, 2013] Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, et al. Translating embeddings for modeling multi-relational data. In *NIPS*, pages 1–9, 2013.

[Cao *et al.*, 2019] Yixin Cao, Xiang Wang, Xiangnan He, Zikun Hu, et al. Unifying knowledge graph learning and recommendation: Towards a better understanding of user preferences. In *WWW*, pages 151–161, 2019.

[Chen *et al.*, 2019a] Chong Chen, Min Zhang, Yiqun Liu, and Shaoping Ma. Social attentional memory network: Modeling aspect-and friend-level differences in recommendation. In *WSDM*, pages 177–185, 2019.

[Chen *et al.*, 2019b] Chong Chen, Min Zhang, Chenyang Wang, Weizhi Ma, Minming Li, Yiqun Liu, et al. An efficient adaptive transfer neural network for social-aware recommendation. In *SIGIR*, pages 225–234, 2019.

[Chen *et al.*, 2020] Chong Chen, Min Zhang, Yongfeng Zhang, Weizhi Ma, et al. Efficient heterogeneous collaborative filtering without negative sampling for recommendation. In *AAAI*, volume 34, pages 19–26, 2020.

[Covington *et al.*, 2016] Paul Covington, Jay Adams, and Emre Sargin. Deep neural networks for youtube recommendations. In *Recsys*, pages 191–198, 2016.

[Du *et al.*, 2018] Chao Du, Chongxuan Li, Yin Zheng, Jun Zhu, and Bo Zhang. Collaborative filtering with user-item co-autoregressive models. In *AAAI*, volume 32, 2018.

[Fan *et al.*, 2019] Wenqi Fan, Yao Ma, Qing Li, Yuan He, Eric Zhao, Jiliang Tang, et al. Graph neural networks for social recommendation. In *WWW*, pages 417–426, 2019.

[Gao *et al.*, 2019] Chen Gao, Xiangnan He, Danhua Gan, Xiangning Chen, Fuli Feng, et al. Learning to recommend with multiple cascading behaviors. *TKDE*, 2019.

[Han *et al.*, 2018] Xiaotian Han, Chuan Shi, et al. Aspect-level deep collaborative filtering via heterogeneous information networks. In *IJCAI*, pages 3393–3399, 2018.

[He *et al.*, 2017] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. Neural collaborative filtering. In *WWW*, pages 173–182, 2017.

[Hofmann and Puzicha, 1999] Thomas Hofmann and Jan Puzicha. Latent class models for collaborative filtering. In *IJCAI*, volume 99, 1999.

[Hu *et al.*, 2018] Binbin Hu, Chuan Shi, Wayne Xin Zhao, and Philip S Yu. Leveraging meta-path based context for top-n recommendation with a neural co-attention model. In *KDD*, pages 1531–1540, 2018.

[Huang *et al.*, 2019] Chao Huang, Xian Wu, Xuchao Zhang, Chuxu Zhang, Jiashu Zhao, Dawei Yin, et al. Online purchase prediction via multi-scale modeling of behavior dynamics. In *KDD*, pages 2613–2622, 2019.

[Huang *et al.*, 2021a] Chao Huang, Jiahui Chen, Lianghao Xia, Yong Xu, et al. Graph-enhanced multi-task learning of multi-level transition dynamics for session-based recommendation. In *AAAI*, 2021.

[Huang *et al.*, 2021b] Chao Huang, Huance Xu, Yong Xu, Peng Dai, et al. Knowledge-aware coupled graph neural network for social recommendation. In *AAAI*, 2021.

[Indurthi *et al.*, 2019] Sathish Reddy Indurthi, Insoo Chung, et al. Look harder: A neural machine translation model with hard attention. In *ACL*, pages 3037–3043, 2019.

[Jamali and others, 2010] Mohsen Jamali et al. A matrix factorization technique with trust propagation for recommendation in social networks. In *Recsys*, pages 135–142, 2010.

[Jiang *et al.*, 2012] Meng Jiang, Peng Cui, et al. Social contextual recommendation. In *CIKM*, pages 45–54, 2012.

[Jin *et al.*, 2020] Bowen Jin, Chen Gao, Xiangnan He, et al. Multi-behavior recommendation with graph convolutional networks. In *SIGIR*, pages 659–668, 2020.

[Kipf and Welling, 2017] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *ICLR*, 2017.

[Klicpera *et al.*, 2019] Johannes Klicpera, Aleksandar Bojchevski, et al. Predict then propagate: Graph neural networks meet personalized pagerank. *ICLR*, 2019.

[Larochelle and Bengio, 2008] Hugo Larochelle and Yoshua Bengio. Classification using discriminative restricted boltzmann machines. In *ICML*, pages 536–543, 2008.

[Liang *et al.*, 2018] Dawen Liang, Rahul G Krishnan, Matthew D Hoffman, et al. Variational autoencoders for collaborative filtering. In *WWW*, pages 689–698, 2018.

[Ma *et al.*, 2008] Hao Ma, Haixuan Yang, et al. Sorec: social recommendation using probabilistic matrix factorization. In *CIKM*, pages 931–940. ACM, 2008.

[Ma *et al.*, 2011] Hao Ma, Dengyong Zhou, Chao Liu, Michael R Lyu, et al. Recommender systems with social regularization. In *WSDM*, pages 287–296, 2011.

[Ma *et al.*, 2019a] Haojie Ma, Wenzhong Li, et al. Attnsense: Multi-level attention mechanism for multimodal human activity recognition. In *IJCAI*, 2019.

[Ma *et al.*, 2019b] Weizhi Ma, Min Zhang, Yue Cao, et al. Jointly learning explainable rules for recommendation with knowledge graph. In *WWW*, pages 1210–1221, 2019.

[Mnih and Salakhutdinov, 2007] Andriy Mnih and Russ R Salakhutdinov. Probabilistic matrix factorization. *NIPS*, 20:1257–1264, 2007.

[Pearl, 2019] Judea Pearl. The seven tools of causal inference, with reflections on machine learning. *Communications of the ACM*, 62(3):54–60, 2019.

[Rennie and Srebro, 2005] Jasson DM Rennie and Nathan Srebro. Fast maximum margin matrix factorization for collaborative prediction. In *ICML*, pages 713–719, 2005.

[Salakhutdinov *et al.*, 2007] Ruslan Salakhutdinov, Andriy Mnih, et al. Restricted boltzmann machines for collaborative filtering. In *ICML*, pages 791–798, 2007.

[Sedhain *et al.*, 2015] Suvash Sedhain, Aditya Krishna Menon, et al. Autorec: Autoencoders meet collaborative filtering. In *WWW*, pages 111–112, 2015.

[Shi *et al.*, 2018] Chuan Shi, Binbin Hu, Wayne Xin Zhao, et al. Heterogeneous information network embedding for recommendation. *TKDE*, 31(2):357–370, 2018.

[Song *et al.*, 2019] Weiping Song, Zhiping Xiao, et al. Session-based social recommendation via dynamic graph attention networks. In *WSDM*, pages 555–563, 2019.

[Srebro *et al.*, 2005] Nathan Srebro, Jason Rennie, and Tommi S Jaakkola. Maximum-margin matrix factorization. In *NIPS*, pages 1329–1336, 2005.

[Tang *et al.*, 2013] Jiliang Tang, Xia Hu, et al. Social recommendation: a review. *SNAM*, pages 1113–1133, 2013.

[Tanjim *et al.*, 2020] Md Mehrab Tanjim, Congzhe Su, et al. Attentive sequential models of latent intent for next item recommendation. In *WWW*, pages 2528–2534, 2020.

[Veličković *et al.*, 2018] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *ICLR*, 2018.

[Wang and others, 2019] Hongwei Wang et al. Knowledge-aware graph neural networks with label smoothness regularization for recommender systems. In *KDD*, 2019.

[Wang *et al.*, 2019a] Xiang Wang, Xiangnan He, Yixin Cao, et al. Kgat: Knowledge graph attention network for recommendation. In *KDD*, pages 950–958, 2019.

[Wang *et al.*, 2019b] Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, and Tat-Seng Chua. Neural graph collaborative filtering. In *SIGIR*, pages 165–174, 2019.

[Wang *et al.*, 2020] Xiang Wang, Hongye Jin, An Zhang, Xiangnan He, Tong Xu, et al. Disentangled graph collaborative filtering. In *SIGIR*, pages 1001–1010, 2020.

[Wei *et al.*, 2020] Yinwei Wei, Xiang Wang, et al. Graph-refined convolutional network for multimedia recommendation with implicit feedback. In *MM*, 2020.

[Wu *et al.*, 2016] Yao Wu, Christopher DuBois, et al. Collaborative denoising auto-encoders for top-n recommender systems. In *WSDM*, pages 153–162, 2016.

[Wu *et al.*, 2019a] Le Wu, Lei Chen, Richang Hong, et al. A hierarchical attention model for social contextual image recommendation. *TKDE*, 32(10):1854–1867, 2019.

[Wu *et al.*, 2019b] Le Wu, Peijie Sun, Yanjie Fu, Richang Hong, et al. A neural influence diffusion model for social recommendation. In *SIGIR*, pages 235–244, 2019.

[Wu *et al.*, 2019c] Qitian Wu, Hengrui Zhang, Xiaofeng Gao, et al. Dual graph attention networks for deep latent representation of multifaceted social effects in recommender systems. In *WWW*, pages 2091–2102, 2019.

[Xia *et al.*, 2020] Lianghao Xia, Chao Huang, Yong Xu, Peng Dai, et al. Multiplex behavioral relation learning for recommendation via memory augmented transformer network. In *SIGIR*, pages 2397–2406, 2020.

[Xia *et al.*, 2021a] Lianghao Xia, Chao Huang, Yong Xu, Peng Dai, Xiyue Zhang, Hongsheng Yang, et al. Knowledge-enhanced hierarchical graph transformer network for multi-behavior recommendation. In *AAAI*, 2021.

[Xia *et al.*, 2021b] Lianghao Xia, Yong Xu, Chao Huang, Peng Dai, and Liefeng Bo. Graph meta network for multi-behavior recommendation. In *SIGIR*, 2021.

[Xin *et al.*, 2019] Xin Xin, Xiangnan He, et al. Relational collaborative filtering: Modeling multiple item relations for recommendation. In *SIGIR*, pages 125–134, 2019.

[Xu *et al.*, 2019] Fengli Xu, Jianxun Lian, et al. Relation-aware graph convolutional networks for agent-initiated social e-commerce recommendation. In *CIKM*, 2019.

[Xu *et al.*, 2020] Huance Xu, Chao Huang, Yong Xu, Lianghao Xia, Hao Xing, and Dawei Yin. Global context enhanced social recommendation with hierarchical graph neural networks. In *ICDM*, pages 701–710. IEEE, 2020.

[Xue *et al.*, 2017] Hong-Jian Xue, Xinyu Dai, et al. Deep matrix factorization models for recommender systems. In *IJCAI*, pages 3203–3209, 2017.

[Yang *et al.*, 2016] Bo Yang, Yu Lei, et al. Social collaborative filtering by trust. *TPAMI*, 39(8):1633–1647, 2016.

[Yang *et al.*, 2019a] Min Yang, Lei Chen, Xiaojun Chen, et al. Knowledge-enhanced hierarchical attention for community question answering with multi-task and adaptive learning. In *IJCAI*, pages 5349–5355, 2019.

[Yang *et al.*, 2019b] Shihui Yang, Jidong Tian, et al. Transms: Knowledge graph embedding for complex relations by multidirectional semantics. In *IJCAI*, 2019.

[Ying *et al.*, 2018] Rex Ying, Ruining He, et al. Graph convolutional neural networks for web-scale recommender systems. In *KDD*, pages 974–983, 2018.

[Yu *et al.*, 2014] Xiao Yu, Xiang Ren, et al. Personalized entity recommendation: A heterogeneous information network approach. In *WSDM*, pages 283–292, 2014.

[Yu *et al.*, 2020] Junliang Yu, Hongzhi Yin, Jundong Li, Min Gao, et al. Enhance social recommendation with adversarial graph convolutional networks. *TKDE*, 2020.

[Zhao *et al.*, 2017] Huan Zhao, Quanming Yao, et al. Meta-graph based recommendation fusion over heterogeneous information networks. In *KDD*, pages 635–644, 2017.

[Zheng *et al.*, 2016] Yin Zheng, Bangsheng Tang, et al. A neural autoregressive approach to collaborative filtering. In *ICML*, pages 764–773. PMLR, 2016.

[Zhu *et al.*, 2020] Qiannan Zhu, Xiaofei Zhou, et al. A knowledge-aware attentional reasoning network for recommendation. In *AAAI*, pages 6999–7006, 2020.