Knowledge-aware Coupled Graph Neural Network for Social Recommendation

Chao Huang[1*], **Huance Xu**[2*], **Yong Xu**[2,3,4†], **Peng Dai**[1], **Lianghao Xia**[2], **Mengyin Lu** [1], **Liefeng Bo**[1], **Hao Xing**[5], **Xiaoping Lai**[5], **Yanfang Ye**[6]
[1]JD Finance America Corporation, USA
[2]South China University of Technology, China, [3]Peng Cheng Laboratory, China
[4]Communication and Computer Network Laboratory of Guangdong, China
[5]VIPS Research, China, [6]Case Western Reserve University, USA
chaohuang75@gmail.com, {cshuance.xu, cslianghao.xia}@mail.scut.edu.cn, yxu@scut.edu.cn,
{peng.dai,mengyin.lu,liefeng.bo}, {hao.xing,tom.lai}vipshop.com, yanfang.ye@case.edu

Abstract

Social recommendation task aims to predict users' preferences over items with the incorporation of social connections among users, so as to alleviate the sparse issue of collaborative filtering. While many recent efforts show the effectiveness of neural network-based social recommender systems, several important challenges have not been well addressed yet: (i) The majority of models only consider users' social connections, while ignoring the inter-dependent knowledge across items; (ii) Most of existing solutions are designed for singular type of user-item interactions, making them infeasible to capture the interaction heterogeneity; (iii) The dynamic nature of user-item interactions has been less explored in many social-aware recommendation techniques. To tackle the above challenges, this work proposes a **K**nowledge-aware **C**oupled **G**raph **N**eural Network (KCGN) that jointly injects the inter-dependent knowledge across items and users into the recommendation framework. KCGN enables the high-order user- and item-wise relation encoding by exploiting the mutual information for global graph structure awareness. Additionally, we further augment KCGN with the capability of capturing dynamic multi-typed user-item interactive patterns. Experimental studies on real-world datasets show the effectiveness of our method against many strong baselines in a variety of settings. Source codes are available at: *https://github.com/xhcdream/KCGN*.

Introduction

In recent years, social recommendation which aims to exploit users' social information for modeling users' preferences in recommendations, has attracted significant attention (Liu et al. 2019). As has been stated in many social-aware recommendation literature (Wu et al. 2019a; Chen et al. 2019b), social influences between users have high impacts on users' interactive behavior over items in various recommender scenarios, such as e-commence (Lin, Gao, and Li 2019) and online review platforms (Chen et al. 2020a). Hence, researchers propose to incorporate social ties into the collaborative filtering architecture as side information to characterize connectivity information across users.

The most common paradigm for state-of-the-art social recommender systems is to learn an embedding function, which unifies user-user and user-item relations into latent representations. To tackle this problem, many studies have developed various neural network techniques to integrate social information with the user-item interaction encoding as constraints. For example, attention-based mechanism has been utilized to aggregate correlations among different users (Chen et al. 2019a,b). Furthermore, inspired by the recent advance of graph neural architectures, several attempts are built upon the message passing frameworks over the user-user social graph. For example, social influence is simulated with layer-wise diffusion scheme for information fusion (Wu et al. 2019a). GraphRec (Fan et al. 2019) employs the graph attention network to model the relational structures between users. To enable the modeling context-aware social effects, DANSER (Wu et al. 2019b) stacks two-stage of graph attention layer for distinguishing the multi-faceted social homophily and influence.

While these solutions have provided encouraging results, several key aspects have not been well addressed yet. In particular, *First*, in real-life scenarios, there typically exist relations between items which characterize item-wise fruitful semantics relatedness, and are helpful to understand user-item interactive patterns (Wang et al. 2019a). For instance, in online retailing systems, products of the same categories (*e.g.*, food & grocery, clothing & shoes) or complement with each other, could be correlated to enrich the knowledge representation of items (Xin et al. 2019). For online review platforms, the exploiting of dependencies among the venues with the same functionality, is able to provide external knowledge in assisting user preference learning (Yu et al. 2019). However, the majority of existing social recommender systems fail to capture item-wise relational structures, which can hardly distill the knowledge-aware collaborative signals from the co-interactive behaviors of users.

Second, to simplify the model design, most of current social recommendation methods have thus far focused on modeling singular type of interactive relations between user and

[*]Both authors contribute equally to this work
[†]Corresponding author: Yong Xu

item. Yet, many practical recommendation scenarios may involve the diversity of users' interaction over items (Cen et al. 2019; Xia et al. 2020). Take the e-commerce site as an example, the effective encoding of multi-typed user-item interactive patterns (*e.g.*, page view, add-to-favorite and purchase) and their underlying inter-dependencies (*e.g.*, add-to-favorite activities may serve as useful indicators for making purchase decisions), is crucial to more accurately inference of user's complex interest in social recommendation tasks.

Third, the time dimension of the social recommendation deserves more investigation, so as to capture behavior dynamics. Most of recent approaches ignore the dynamic nature of user-item interactions and assume that the factor influencing the interactive behavior is only the identity of items (Song et al. 2019). While there exist a handful of recent work that consider the sequential information in social recommendation (Song et al. 2019; Sun, Wu, and Wang 2018), their are limited in their intrinsic design for singular type of user-item relations. This makes them insufficient to yield satisfactory embeddings with the preservation of different interaction signals in a dynamic manner for more complex scenarios.

While intuitively useful to integrate the above dimensions into social recommendation frameworks, two unique technical challenges arise in achieving this goal. Specifically, graph-structured neural network can be applied to naturally model the topological information of social node instances, such as the graph-based convolutional network (Wu et al. 2019a) or attention mechanism (Wu et al. 2019b; Fan et al. 2019). However, their non-linear aggregation functions can only learn the local proximity between users and are incapable of capturing the broader context of the graph structure (*e.g.*, users with the isomorphic social structures) (You, Ying, and Leskovec 2019). Hence, how to jointly capture knowledge-aware user-user and item-item local relations, as well as retain the high-order social influence and item dependencies under global context, remains a significant challenge. Additionally, it is also very challenging to handle the dynamic multi-typed user-item interactions, so as to capture the dynamic relation-aware structural dependencies across users and items with arbitrary duration.

The Present Work. In light of the aforementioned motivations and challenges, we study the social recommendation problem by proposing the **K**nowledge-aware **C**oupled **G**raph **N**eural Network (KCGN). To jointly deal with the user-user and item-item local and global relational structure awareness, we incorporate the mutual information estimation schema into the coupled graph neural architecture. This design enables the collaboration between neural mutual information estimator and graph-structured representation learning paradigm, which preserves the node-level unique characteristics and graph-level substructure knowledge across users and items. In addition, to capture the dynamic multi-typed interactive patterns, we integrate a relation-aware message passing framework with the relative temporal encoding strategy, which endows KCGN with the capability of incorporating the temporal information into the multi-typed user-item interaction graph learning.

Our contributions can be highlighted as follows:

- We propose to capture both user-user and item-item relations with the developed coupled graph neural network. Through the joint modeling of user- and item-wise dependent structures, our KCGN can enhance the social-aware user embeddings with the preservation of knowledge-aware cross-item relations in a more thorough way.
- We propose a relation-aware graph neural module to encode the multi-typed user-item interactive patterns, and further incorporate the temporal information into the message passing kernel to enhance the learning of collaborative relations for recommendation.
- We conduct extensive experiments on three real-world datasets to show the superiority of our KCGN when competing with several baselines from various research lines. Further studies on scalability evaluation validate the model efficiency of KCGN over state-of-the-art social recommender systems. We also show that our model maintains strong performance in the cold-start scenarios when user-item interactions are sparse.

Problem Definition

We first introduce key definitions of social recommendation with item relational knowledge and different types of user-item interactions. We consider a typical recommendation scenario, in which we have I users $U = \{u_1, ..., u_i, ..., u_I\}$ and J items $V = \{v_1, ..., v_j, ..., v_J\}$. To capture the multi-typed user-item interaction signals, we define a multi-typed interaction tensor as below:

Definition 1 ***Multi-typed Interaction Tensor X***. *We define a three-way tensor $\boldsymbol{X} \in \mathbb{R}^{I\times J\times K}$ to represent the different types of interactions between user and item, where K (indexed by k) denotes the number of interaction types (page view, purchase, or like, dislike). In **X**, the element $x_{i,j}^k = 1$ if user u_i interacts with item v_j with the interaction type of k and $x_{i,j}^k = 0$ otherwise. To deal with the interaction dynamics, we also define a temporal tensor $\boldsymbol{T} \in \mathbb{R}^{I\times J\times K}$ with the same size of **X** to record the timestamp information ($t_{i,j}^k$) of each corresponding interaction $x_{i,j}^k$.*

Definition 2 ***User Social Graph*** G_u. *$G_u = \{U, E_u\}$ represents the social relationships (edges E_u) among users (nodes U), where there exists an edge $e_{i,i'}$ between user u_i and $u_{i'}$ given they are socially connected.*

Definition 3 ***Item Inter-Dependency Graph*** G_v. *We further define $G_v = \{V, E_v\}$ to represent the inter-dependence of items. In particular, we characterize the item-wise relations with a triple $\{v_j, e_{j,j'}, v_{j'} | v_j, v_{j'} \in V\}$, where edge $e_{j,j'}$ describes the relationship between item v_j and $v_{j'}$, e.g., v_j and $v_{j'}$ belong to the same product categories and have similar functionality, or are interacted by the same user under the same interaction type of k.*

Task Formulation. We formulate the studied recommendation task in this paper as: **Input**: multi-typed interaction tensor $\mathbf{X} \in \mathbb{R}^{I\times J\times K}$, user social graph G_u and item inter-dependence graph G_v. **Output**: a predictive function that effectively forecasts the future user-item interaction.

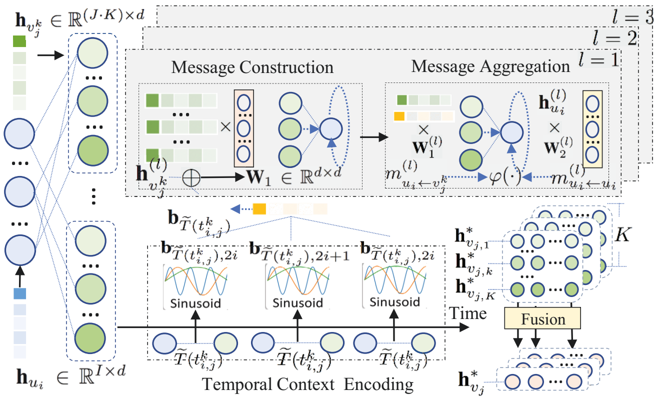


Figure 1: The architecture of the multi-typed interactive pattern modeling. $\oplus$ denotes the element-wise addition.

Methodology

Multi-typed Interactive Pattern Modeling

To encode the multi-typed collaborative relations, we propose a relation-aware graph neural architecture, which is built upon the message passing paradigm (as shown in Figure 1), to empower KCGN to capture the dedicated patterns of different types of user-item interactions. Specifically, given the multi-typed interaction tensor $\mathbf{X}$, we first construct a multi-typed relation graph G_m by representing the interaction heterogeneity with type-specific item sub-vertices $v_j \rightarrow (v_j^1, ..., v_j^k, ..., v_j^K)$, where K denotes the number of interaction types. Each edge between u_i and v_j^k represents the corresponding interaction with the k-th type. After that, there are $(I + J \cdot K)$ vertices in our multi-typed graph $G_m = (V_m, E_m)$, where $V_m = U \cup V'$ and $v_j^k \in V'$. Here, V' is the new type-aware item set.

Message Construction Phase. We first generate the message between user vertex u_i and his/her interacted type-specific item vertex v_j^k as follows:

$$m_{u_i \leftarrow v_j^k} = \gamma(\mathbf{h}_{v_j^k}, \rho_{i,j}^k); \quad m_{v_j^k \leftarrow u_i} = \gamma(\mathbf{h}_{u_i}, \rho_{i,j}^k) \tag{1}$$

where $\gamma(\cdot)$ denotes the information encoding function over the input feature embeddings $\mathbf{h}_{v_j^k} \in \mathbb{R}^{(J \cdot K) \times d}$, $\mathbf{h}_{u_i} \in \mathbb{R}^{I \times d}$. $\rho_{i,j}^k$ is the decay factor to normalize the propagated influence with node degrees (Chen et al. 2020b), *i.e.*, $\rho = \frac{1}{\sqrt{|N_i||N_j^k|}}$, where N_i denotes the number of neighboring nodes of user u_i and N_j^k represents the number of connected user nodes of item v_j under the relation type of k. Hence, the constructed message can be unfolded as:

$$m_{u_i \leftarrow v_j^k} = \frac{1}{\sqrt{|N_i||N_j^k|}} (\mathbf{h}_{v_j^k} \cdot \mathbf{W}_1) \tag{2}$$

where $\mathbf{W}_1 \in \mathbb{R}^{d \times d}$ is the weight matrix. We apply the similar operation for the message propagation from u_i to type-specific item v_j^k.

Temporal Context Encoding Scheme. Inspired by the recommendation techniques with the modeling of temporal information (Sun et al. 2019; Huang et al. 2019), in our framework, we allow the user-item interactions happening at different timestamps interweave with each other, by introducing a temporal context encoding scheme to model the dynamic dependencies across different types of users' interactions. Motivated by the positional encoding algorithm in Transformer architecture (Vaswani et al. 2017; Sun et al. 2019; Wu et al. 2020), we map the timestamp t_{i,j^k} of individual interaction $x_{i,j}^k$ into separate time slot as: $\widetilde{T}(t_{i,j}^k)$. We employ the sinusoid functions to generate the relative time embedding for edge $e_{i,j}^k \in E_m$ in G_m as:

$$\begin{aligned}
\mathbf{b}_{\widetilde{T}(t_{i,j}^k),2i} &= sin(\widetilde{T}(t_{i,j}^k)/10000^{\frac{2i}{d}}) \\
\mathbf{b}_{\widetilde{T}(t_{i,j}^k),2i+1} &= cos(\widetilde{T}(t_{i,j}^k)/10000^{\frac{2i+1}{d}})
\end{aligned} \tag{3}$$

where $(2i)$ and $(2i+1)$ denotes the element index with the even and odd position in embedding $\mathbf{b}_{\widetilde{T}(t_{i,j}^k)}$, respectively.

High-Order Message Aggregation Phase. We incorporate the propagated message between user u_i and item $v_{i,j}^k$, as well as temporal context $\mathbf{b}_{\widetilde{T}(t_{i,j}^k)}$ on their interaction edge $e_{i,j}^k$, into our information propagation paradigm as below:

$$\begin{aligned}
\mathbf{h}_{u_i}^{(l+1)} &= \varphi\Big(m_{u_i \leftarrow u_i}^{(l)} + \sum_{(j,k) \in N_{u_i}} m_{u_i \leftarrow v_j^k}^{(l)}\Big) = \varphi\Big(\frac{1}{|N_{u_i}|}\mathbf{h}_{u_i}^{(l)}\mathbf{W}_2^{(l)} \\
&+ \sum_{(j,k) \in N_{u_i}} \frac{1}{|N_{v_j^k}|}(\mathbf{h}_{v_j^k}^{(l)} \oplus \mathbf{b}_{\widetilde{T}(t_{i,j}^k)})\mathbf{W}_1^{(l)}\Big)
\end{aligned} \tag{4}$$

where $\varphi(\cdot)$ denotes the LeakyReLU function to perform the transformation. $m_{u_i \leftarrow u_i}^{(l)}$ is the self-propagated message with the weight matrix $\mathbf{W}_2^{(l)} \in \mathbb{R}^{d \times d}$. $\oplus$ denotes the element-wise addition. l is the index of L graph layers. We finally generate the user/item embeddings (*i.e.*, $\mathbf{h}_{u_i}^*$, $\mathbf{h}_{v_{i,j}^k}^*$) with the following concatenation operation $\|$ as follows:

$$\begin{aligned}
\mathbf{h}_{u_i}^* &= (\mathbf{h}_{u_i}^{(0)} \parallel \mathbf{h}_{u_i}^{(1)} \parallel \cdots \parallel \mathbf{h}_{u_i}^{(L)}) \\
\mathbf{h}_{v_{j,k}}^* &= (\mathbf{h}_{v_{j,k}}^{(0)} \parallel \mathbf{h}_{v_{j,k}}^{(1)} \parallel \cdots \parallel \mathbf{h}_{v_{j,k}}^{(L)})
\end{aligned} \tag{5}$$

We generate the summarized representation $\mathbf{h}_{v_j}^*$ over all item sub-vertex embeddings $\mathbf{h}_{v_{j,k}}^*$ $(k \in [1, ..., K])$ with a gating mechanism (Ma, Kang, and Liu 2019), to differentiate the importance of type-specific interaction patterns.

Knowledge-aware Coupled Graph Neural Module

To jointly inject the user- and item-wise inter-dependent knowledge into our user preference modeling, we develop a knowledge-aware coupled graph neural network which enables the collaboration between the mutual information learning and graph representation paradigm. While many efforts have been devoted to modeling graph structural information, they are limited in their ability in capturing both local and global graph substructure awareness (Velickovic et al. 2019), such as the user- and item-specific social/knowledge dependent information and high-order relationships across users/items. KCGN is equipped with a dual-stage graph learning paradigm (As shown in Figure 2).

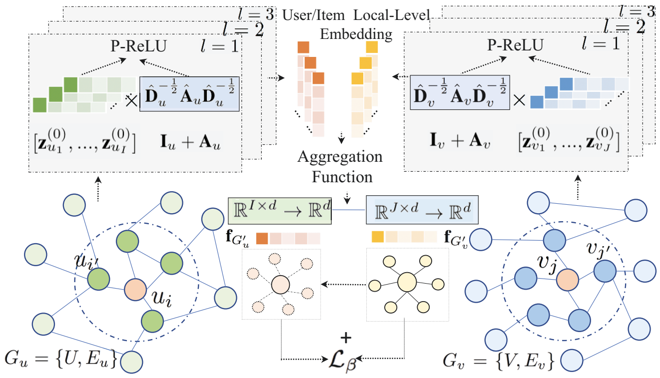


Figure 2: The architecture of joint encoding of user-user and item-item inter-dependent relational structures.

Local Relational Structure Modeling. We first learn the user- and item-specific specific embeddings ($\mathbf{z}_{u_i}, \mathbf{z}_{v_j}$) which preserves the local connection information over user social graph G_u and item inter-dependent graph G_v with the following graph-based update functions ($\mathbf{z}^0_{u_i}$=$\mathbf{h}^*_{u_i}$, $\mathbf{z}^0_{v_j}$ = $\mathbf{h}^*_{v_j}$):

$$[\mathbf{z}^{(l+1)}_{u_1}, ..., \mathbf{z}^{(l+1)}_{u_I}] = \varphi\Big([\mathbf{z}^{(l)}_{u_1}, ..., \mathbf{z}^{(l)}_{u_I}] \cdot \eta(G_u)\Big)$$
$$[\mathbf{z}^{(l+1)}_{v_1}, ..., \mathbf{z}^{(l+1)}_{v_J}] = \varphi\Big([\mathbf{z}^{(l)}_{v_1}, ..., \mathbf{z}^{(l)}_{v_J}] \cdot \eta(G_v)\Big) \quad (6)$$

where $\eta(\cdot)$ denotes the adjacent relations of G_u and G_v with the symmetric normalization strategy in the information aggregation across the neighboring users/items, *e.g.*, $\eta(G_v) = \hat{\mathbf{D}}_v^{-\frac{1}{2}} \hat{\mathbf{A}}_v \hat{\mathbf{D}}_v^{-\frac{1}{2}}$. Hence, $\hat{\mathbf{A}}_v$ is the addition of identity matrix $\mathbf{I}_v$ and adjacent matrix $\mathbf{A}_v$, so as to incorporate the information self-propagation (Chen et al. 2020b).

In this graph learning paradigm, we aim to inject both local- and global-level relational structures over the user social graph and item relation graph into our learned user/item representations. Different from the existing graph neural network approaches (Velickovic et al. 2019; Xu et al. 2020) which model the mutual relations between local feature embeddings and a single global representation, we enrich the global semantics with the consideration of connected graph substructures (*e.g.*, the entire social relations of all users may consist of different connected sub-graphs G'_u). In particular, we first generate a fused graph-level representation $\mathbf{f}_{G'_u}, \mathbf{f}_{G'_v} \in \mathbb{R}^d$ by applying the mean pooling over node-specific embeddings.

We design our neural mutual information estimator based on a discriminator $\mathcal{D}(x, y)$ for node-graph pairwise relationships, to provide probability scores for sampled pairs. To be specific, we generate positive samples as $(\mathbf{z}_{u_i}, \mathbf{f}_{G'_u})$, $(\mathbf{z}_{v_j}, \mathbf{f}_{G'_v})$, and negative samples as $(\widetilde{\mathbf{z}}_{u_i}, \mathbf{f}_{G'_u})$, $(\widetilde{\mathbf{z}}_{v_j}, \mathbf{f}_{G'_v})$. Here, $\widetilde{\mathbf{z}}_{u_i}$ and $\widetilde{\mathbf{z}}_{u_i}$ are randomly picked with node shuffling to generate the misplaced node-graph pairwise relations.

Due to the rationality of cross-entropy in mutual information maximization (Wang et al. 2020), we define our noise-contrastive knowledge-aware loss function $\mathcal{L}_\beta$ as follows:

$$\mathcal{L}_\beta = -\frac{\lambda_1}{N^u_{pos} + N^u_{neg}} \Big(\sum_{i=1}^{N^u_{pos}} \tau(\mathbf{z}_{u_i}, \mathbf{f}_{G'_u}) \cdot log\sigma(\mathbf{z}_{u_i} \cdot \mathbf{f}_{G'_u})$$
$$+ \sum_{i=1}^{N^u_{neg}} \tau(\widetilde{\mathbf{z}}_{u_i}, \mathbf{f}_{G'_u}) \cdot log[1 - \sigma(\widetilde{\mathbf{z}}_{u_i} \cdot \mathbf{f}_{G'_u})] \Big)$$
$$-\frac{\lambda_2}{N^v_{pos} + N^v_{neg}} \Big(\sum_{i=1}^{N^v_{pos}} \tau(\mathbf{z}_{v_j}, \mathbf{f}_{G'_v}) \cdot log\sigma(\mathbf{z}_{v_j} \cdot \mathbf{f}_{G'_v})$$
$$+ \sum_{i=1}^{N^v_{neg}} \tau(\widetilde{\mathbf{z}}_{v_j}, \mathbf{f}_{G'_u}) \cdot log[1 - \sigma(\widetilde{\mathbf{z}}_{v_j} \cdot \mathbf{f}_{G'_v})] \Big) \quad (7)$$

where N^u_{pos}/N^v_{pos} and N^v_{neg}/N^v_{neg} denotes the number of positive and negative instances sampled over sub-graph G'_u and G'_v. $\tau(\cdot)$ is an indicator function, *e.g.*, $\tau(\mathbf{z}_{v_j}, \mathbf{f}_{G'_v}) = 1$ and $\tau(\widetilde{\mathbf{z}}_{v_j}, \mathbf{f}_{G'_v}) = 1$ corresponds to the positive and negative pair instances. λ_1 and λ_2 are balance parameters. We aim to minimize $\mathcal{L}_\beta$ which is equivalent to maximize the mutual information, to jointly preserve the node-specific user/item characteristics and global graph-level dependencies.

Model Optimization

We define our loss $\mathcal{L}$ which includes (i) multi-typed user-item interaction encoding; (ii) knowledge-aware user-user and item-item inter-dependent relation learning. Particularly, $\mathcal{L}$ integrates the pairwise BPR loss, which is widely adopted in recommendation tasks (Wang et al. 2019c), with the mutual information maximization paradigm as:

$$\mathcal{L} = \sum_{(i,j^+,j^-)\in O} - \text{In}\, \sigma(\widehat{x}_{i,j^+} - \widehat{x}_{i,j^-}) + \lambda \parallel \Theta \parallel^2 + \mathcal{L}_\beta \quad (8)$$

the pairwise training data is denoted as $O = \{(u, j^+, j^-)|(u, j^+) \in \mathcal{R}^+, (u, j^-) \in \mathcal{R}^-\}$ ($\mathcal{R}^+$, $\mathcal{R}^-$ denotes the observed and unobserved interactions, respectively). $\widehat{x}_{i,j}$ is the calculated score with the inner-product between the embedding of u_i and v_j. Θ are trainable parameters, $\sigma(\cdot)$–sigmoid. λ controls the strength of L_2 regularization for overfitting alleviation.

Time Complexity Analysis. KCGN takes $O(|E| \times d)$ for the message passing in handling the user-user, user-item and item-item relations, where $|E|$ denotes the number of edges. Also, $O((I + J \cdot K) \times d^2)$ computation is spent by the transformations. Typically, the first term is dominant due to information compression. In conclusion, KCGN is comparable in time efficiency compared with current GNN-based recommendation methods. Our model only utilizes moderate memory to store node embeddings ($O((I + J \cdot K) \times d)$), which is similar to the existing methods.

Evaluation

In this section, we conduct experiments on different real-world datasets to evaluate the performance of our method from the following aspects:

- **RQ1**: Does *KCGN* consistently outperform other baseline in terms of recommendation accuracy?

Dataset	Epinions	Yelp	E-commerce
# of Users	18,081	43,043	334,042
# of Items	251,722	66,576	195,940
# of Interactions	715,821	283,512	1,930,466

Table 1: Statistics of Experimented Datasets.

- **RQ2**: How is the performance of *KCGN*'s variants with the combination of different relation encoders?
- **RQ3**: How is forecasting performance of compared methods *w.r.t* different interaction density degrees?
- **RQ4**: How do the representations benefit from the collectively encoding of global knowledge-aware cross-interactive patterns in social recommendation?
- **RQ5**: How do different hyper-parameter settings impact the performance of our *KCGN* framework?
- **RQ6**: How is the model efficiency of the *KCGN*?

Experimental Settings

Dataset. Table 1 lists the statistics of three datasets. We describe the details of those datasets as follows:
Epinions. This data records the user's feedback over different items from a social network-based review system Epinions (Fan et al. 2019). Each explicit rating score (ranging from 1 to 5) is regarded as an individual type of interaction: negative, below average, neutral, above average, positive.
Yelp. This data is collected from the Yelp platform, in which user-item interactions are differentiated with the same split rubric in Epinions. Furthermore, user's social connections (with common interests) are contained in this data.
E-Commerce. It is collected from a commercial e-commerce platform with different types of interactions, *i.e.*, *page view*, *add-to-cart*, *add-to-favorite* and *purchase*. User's relations are constructed with their co-interact patterns.

The item inter-dependency graph G_v on the above datasets are constructed based on item categories.

Evaluation Protocols. We adopt two widely used evaluation metrics for social recommendation tasks (Chen et al. 2019a): *Hit Ratio (HR@N)* and *Normalized Discounted Cumulative Gain (NDCG@N)*. We follow the evaluation settings in (Chen et al. 2019b; Wu et al. 2019a) and employ the leave-one-out method for generating training and test data instances. To be consistent with (Sun et al. 2019), we associate each positive instance with 99 negative samples.

Baselines. In our experiments, we perform the performance comparison by considering the following baselines:
Probabilistic Matrix Factorization Method.

- **PMF** (Mnih and *et al* 2008): it is a probabilistic approach with the matrix factorization for user/item factorization.

Conventional Social Recommendation Methods.

- **TrustMF** (Yang et al. 2016): this method incorporates the truth relationships between users into the matrix factorization architecture for user interaction embedding.

Attentive Social Recommendation Techniques.

- **SAMN** (Chen et al. 2019a): this model is a dual-stage attention network which learns the influences between the target user and his/her neighboring nodes.
- **EATNN** (Chen et al. 2019b): This transfer learning model is also on the basis of attention mechanism to jointly fuse information from user's interactions and social signals.

Graph Neural Networks Social Recommender Systems.

- **DiffNet** (Wu et al. 2019a): it is a deep influence propagation framework to model the social diffusion process.
- **GraphRec** (Fan et al. 2019): it aggregates the social relations between users via a graph neural architecture.
- **NGCF+S** (Wang et al. 2019c): we incorporate the social ties into the state-of-the-art graph-structured neural collaborative filtering model for joint message propagation.
- **Danser** (Wu et al. 2019b): it is composed of two graph attention layers for capturing the social influence and homophily, respectively from both users and items.
- **LR-GCCF** (Chen et al. 2020b): it is a new graph-based collaborative filtering model based on graph convolutional network by removing non-linear transformations.

Social Recommendation with Sequential Pattern.

- **DGRec** (Song et al. 2019): it jointly models the dynamic user's preference and the underlying social relations.

Knowledge Graph-enhanced Recommendation.

- **KGAT** (Wang et al. 2019b): it is a graph attentive message passing framework which utilizes the knowledge graph to enhance the recommendation with item side information.

Implementation Details. The *KCGN* is implemented with Pytorch and Adam optimizer is adopted for hyperparameter estimation. The training process is performed with the learning rate range of [0.001, 0.005, 0.01], and the batch size selected from $[1024, 2048, 4096, 8192]$. The embedding size is tuned from the range of $[8, 16, 32, 64]$. In our evaluations, we employ the early stopping for training termination when the performance degrades for 5 continuous epochs on the validation data.

Overall Model Performance Comparison (RQ1)

Table 2 reports the results of *KCGN* and many baselines in predicting the overall interactions in terms of HR@10 and NDCG@10. It can be seen that *KCGN* consistently obtains the best performance across different recommendation scenarios in terms of two metrics, which justifies the effectiveness of our method in integrating user-user and item-item relations, with the multi-typed user-item interactive patterns.

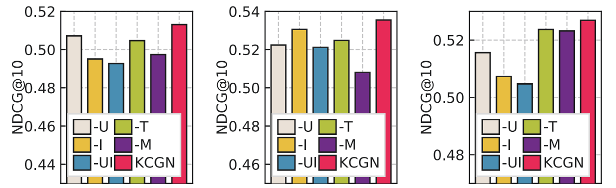
Compared with traditional approaches, neural network based models usually achieve better performance, due to the modeling of high-level non-linearities during the feature interaction learning phase. Among various compared approaches, the GNN-based models outperforms the attentive social recommender systems, which ascertains the rationality of applying graph neural networks for high-order relations across users/items in a recursive way. Different from those GNN techniques, our framework integrates the social and knowledge-aware relations from global context via a mutual information encoding paradigm, and also captures

Data	Metrics	PMF	TrustMF	DiffNet	SAMN	DGRec	EATNN	NGCF+S	KGAT	GraphRec	Danser	LR-GCCF	*KCGN*
Epinions	HR	0.619	0.635	0.632	0.639	0.626	0.642	0.707	0.675	0.686	0.669	0.677	**0.742**
	NDCG	0.410	0.417	0.416	0.425	0.412	0.448	0.498	0.470	0.478	0.462	0.478	**0.513**
Yelp	HR	0.698	0.756	0.785	0.751	0.766	0.771	0.781	0.772	0.7605	0.774	0.769	**0.8026**
	NDCG	0.460	0.495	0.512	0.486	0.495	0.506	0.523	0.511	0.494	0.508	0.518	**0.530**
E-Com	HR	0.654	0.674	0.722	0.676	0.672	0.683	0.694	0.689	0.668	0.670	0.690	**0.735**
	NDCG	0.431	0.452	0.519	0.461	0.441	0.456	0.476	0.473	0.439	0.443	0.485	**0.529**

Table 2: Performance comparison of all methods in interaction prediction.

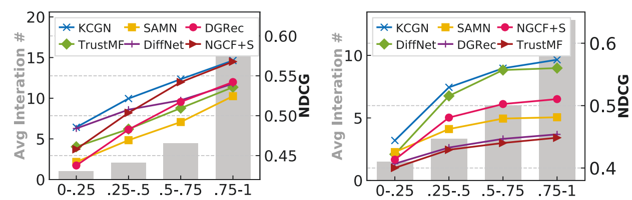
Data	Metrics	DiffNet	DGRec	KGAT	GraphRec	Danser	*KCGN*
Epinions	HR	0.628	0.625	0.685	0.678	0.653	**0.745**
	NDCG	0.411	0.409	0.480	0.465	0.444	**0.519**
Yelp	HR	0.809	0.808	0.791	0.781	0.790	**0.839**
	NDCG	0.542	0.534	0.530	0.520	0.533	**0.573**
E-Com	HR	0.894	0.900	0.886	0.849	0.872	**0.911**
	NDCG	0.673	0.659	0.653	0.627	0.649	**0.710**

Table 3: Prediction results for like/purchase activities on three datasets in terms of *HR@10* and *NDCG@10*.



(a) Epinions (b) Yelp (c) E-commerce

Figure 3: Ablation studies for different sub-modules of KCGN framework, in terms of *HR@10* and *NDCG@10*.



(a) Yelp (b) E-Commerce

Figure 4: Performance of KCGN and baselines over users with different sparsity from Yelp and E-Commerce data.

interaction dynamics, which results in better performance.

We further investigate the performance of our *KCGN* in making recommendations on the target type of interactions (*e.g.*, positive feedback on Epinions and Yelp or user's purchase on E-commerce). The results are shown in Table 3. We can observe that *KCGN* still achieves significant improvement, with the careful consideration of different types of user-item interaction signals. While the baseline KGAT proposes to incorporate the auxiliary knowledge graph, it fails to explicitly differentiate type-specific interaction patterns.

Impact of Different Relation Encoders (RQ2)

We next perform experiments to evaluate the impact of the incorporation of multi-typed user-item interactions, user-wise relations, item-wise dependencies, and the temporal context, with the following five contrast variants of *KCGN*.

- *KCGN*-M: *KCGN* without modeling multi-typed interaction patterns and only with singular-type interactions.
- *KCGN*-U: *KCGN* without the social relation encoder for capturing the social signals in the recommendation.
- *KCGN*-I: *KCGN* without the external knowledge to characterize the item dependency.
- *KCGN*-UI: *KCGN* without both the user- and item-wise relation encoders and remove the coupled mutual information paradigms in the joint learning framework.
- *KCGN*-T: *KCGN* without the temporal context encoding.

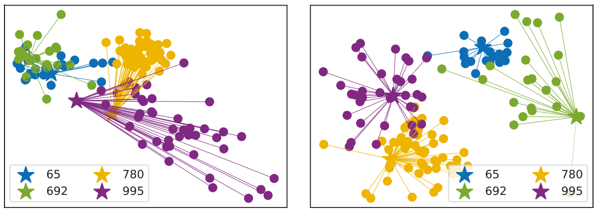
Figure 3 shows the comparison results of different variants. We can see that the joint model *KCGN* achieves the best performance. As such, it is necessary to build a joint framework to simultaneously capture social dimension (users' social influence), item dimension (knowledge-aware inter-item relations), multi-typed interactions, and time-aware user's interest, for making recommendations. In addition, *KCGN*-UI performs worse than *KCGN*-U and *KCGN*-I, which again confirms the efficacy of our designed relation aggregation functions.

Performance over Sparsity Degrees (RQ3)

One key motivation to exploit social- and knowledge-aware side information is to alleviate the sparsity issue, which limits the model robustness. Hence, we further evaluate our *KCGN* for both inactive and active users. In particular, we partition the target users into four sparsity levels in terms of their interaction densities. Figure 4 presents the evaluation results on different user groups on Yelp and E-Commerce data in terms of *NDCG@10*. We can observe that *KCGN* outperforms representative baselines in most cases, especially on sparest user groups. This suggests that incorporating both user and item side knowledge as their external relations, empowers the representations of inactive users through our recursive information aggregation architecture.

Qualitative Analyses of *KCGN* (RQ4)

We illustrate how our side knowledge-aware multi-typed relation encoding schema benefit the ability of embedding user's preference into the latent learning space. In particular, we sample several users and their four- and five-star rated items from Yelp dataset, and further visualize the corresponding user/item embeddings learned by NGCF+S and our KCGN (as shown in Figure 5). From the results, we can



(a) NGCF+S (b) KCGN

Figure 5: Visualized embeddings for users (stars) and their 4- or 5-rated item (circles), learned by KCGN and NGCF+S.

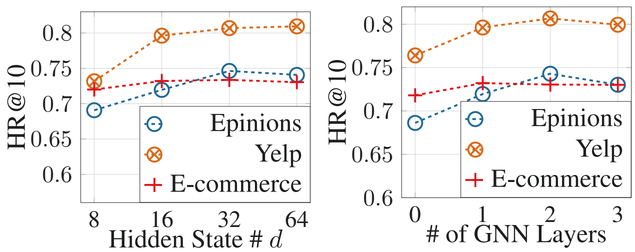


Figure 6: Hyper-parameter study of *KCGN*

notice that: i) the visualized embeddings could well preserve the relationships between users and their interacted items with a clustering phenomenon (represented with the same color); ii) *KCGN* could provide a better separation for different users and their interacted items. Hence, the above observations verify the superior representation learning ability of KCGN through the encoding function which maps the side knowledge and interaction units into effective latent space.

Parameter Sensitivity Study (RQ5)

Impact of # Recursive Graph Layers. Figure 6 shows the experimental results with different number of embedding propagation layers over user-item interaction graph. We can observe that increasing the depth of *KCGN* could boost the performance, *i.e.*, *KCGN*-2 performs better than *KCGN*-0 (without the graph structure) and *KCGN*-1 (only consider 1-hop neighbors). The performance improvement lies in the effective modeling of high-order collaborative effects across users and items. *KCGN* with 3 graph layers performs worse than *KCGN*-2, suggests that exploring higher-level relations may involve noise.

Impact of Embedding Dimension. We notice that the accuracy is initially improved with larger embedding size due to the stronger representation ability. However, the performance degrades with the further increase of dimensionality, which indicates the overfitting phenomenon.

Model Efficiency Study (RQ6)

We finally investigate the computation cost of our *KCGN* when competing with state-of-the-art baselines. As shown in Table 4, we can observe that *KCGN* achieves competitive time efficiency (measured by running time of each epoch) when compared with neural social recommendation methods. It is worthwhile pointing out that methods with stacking multiple graph attention layers is time-consuming, due to their pairwise attentive weights calculations for social or knowledge graph information aggregation.

Data	DGRec	SAMN	EATNN	KGAT	GraphRec	*KCGN*
Epinions	4.4	4.7	10.7	60.5	40.5	17.5
Yelp	2.6	8.9	13.5	20.9	7.3	3.7
E-Com	82.5	78.3	152.7	342.8	497.3	70.2

Table 4: Model computational cost with running time (s).

Related Work

Social-aware Recommender Systems. Deep learning has been revolutionizing recommender systems and many neural network models have been proposed for social recommendation scenario (Yin et al. 2019; Chen et al. 2020a). For example, attention mechanisms are introduced to learn the influences between users, such as SAMN (Chen et al. 2019a) and EATNN (Chen et al. 2019b). It is worth mentioning that several recent efforts explore the GNNs for incorporating social relations into the user-item interaction encoding (Wu et al. 2019b; Fan et al. 2019; Wu et al. 2019a; Xu et al. 2020). Different from these methods, KCGN focus on fusing the heterogeneous relations from different aspects (social, item knowledge and temporal), to boost the performance.

Graph Methods for Recommendation. Many recent efforts have been devoted to exploring insights from GNNs for modeling collaborative signals in recommender systems. For example, inspired by the graph convolutional operations, PinSage (Ying et al. 2018) and NGCF (Wang et al. 2019c) aim to aggregate high-hop neighboring feature information over the user-item interaction graph. Several subsequent extensions have been developed to revisit the graph-based CF effects, such as LightGCN (He et al. 2020), LR-GCCF (Chen et al. 2020b) and KHGT (Xia et al. 2021). Motivated by these works, we propose a new knowledge-aware graph neural architecture for social recommendation.

Conclusion

In this paper, we propose KCGN, an end-to-end framework that naturally incorporates knowledge-aware item dependency into the social recommender systems. KCGN unifies the user-user and item-item relation structure learning with a coupled graph neural network under a mutual information-based neural estimator. To handle the dynamic user-item interaction heterogeneity, we design a relation-aware graph encoder to empower KCGN to maintain dedicated representations of multi-typed interaction signals with the incorporation of temporal information. Through extensive experiments on real-world datasets, we demonstrate that KCGN achieves substantial gains over state-of-the-art baselines.

Acknowledgments

We thank the anonymous reviewers for their constructive feedback and comments. This work is supported by National Nature Science Foundation of China (62072188, 61672241), Natural Science Foundation of Guangdong Province (2016A030308013), Science and Technology Program of Guangdong Province (2019A050510010).

References

Cen, Y.; Zou, X.; Zhang, J.; Yang, H.; Zhou, J.; and Tang, J. 2019. Representation learning for attributed multiplex heterogeneous network. In *KDD*, 1358–1368.

Chen, C.; Zhang, M.; Liu, Y.; and Ma, S. 2019a. Social attentional memory network: Modeling aspect-and friend-level differences in recommendation. In *WSDM*, 177–185.

Chen, C.; Zhang, M.; Wang, C.; Ma, W.; Li, M.; Liu, Y.; and Ma, S. 2019b. An efficient adaptive transfer neural network for social-aware recommendation. In *SIGIR*, 225–234.

Chen, H.; Yin, H.; Chen, T.; Wang, W.; Li, X.; and Hu, X. 2020a. Social Boosted Recommendation with Folded Bipartite Network Embedding. *TKDE* .

Chen, L.; Wu, L.; Hong, R.; Zhang, K.; and Wang, M. 2020b. Revisiting Graph Based Collaborative Filtering: A Linear Residual Graph Convolutional Network Approach. In *AAAI*, 27–34.

Fan, W.; Ma, Y.; Li, Q.; He, Y.; Zhao, E.; Tang, J.; and Yin, D. 2019. Graph Neural Networks for Social Recommendation. In *WWW*, 417–426. ACM.

He, X.; Deng, K.; Wang, X.; Li, Y.; Zhang, Y.; and Wang, M. 2020. LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation .

Huang, C.; Wu, X.; Zhang, X.; Zhang, C.; Zhao, J.; Yin, D.; and Chawla, N. V. 2019. Online purchase prediction via multi-scale modeling of behavior dynamics. In *KDD*, 2613–2622.

Lin, T.-H.; Gao, C.; and Li, Y. 2019. Cross: Cross-platform recommendation for social e-commerce. In *SIGIR*, 515–524.

Liu, C.; Wang, X.; Lu, T.; Zhu, W.; Sun, J.; and Hoi, S. 2019. Discrete social recommendation. In *AAAI*, volume 33, 208–215.

Ma, C.; Kang, P.; and Liu, X. 2019. Hierarchical gating networks for sequential recommendation. In *KDD*, 825–833.

Mnih, A.; and *et al*. 2008. Probabilistic matrix factorization. In *NIPS*, 1257–1264.

Song, W.; Xiao, Z.; Wang, Y.; Charlin, L.; Zhang, M.; and Tang, J. 2019. Session-based social recommendation via dynamic graph attention networks. In *WSDM*, 555–563.

Sun, F.; Liu, J.; Wu, J.; Pei, C.; Lin, X.; Ou, W.; and Jiang, P. 2019. BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer. In *CIKM*, 1441–1450.

Sun, P.; Wu, L.; and Wang, M. 2018. Attentive recurrent social recommendation. In *SIGIR*, 185–194.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; and Polosukhin, I. 2017. Attention is all you need. In *NIPS*, 5998–6008.

Velickovic, P.; Fedus, W.; Hamilton, W. L.; Liò, P.; Bengio, Y.; and Hjelm, R. D. 2019. Deep Graph Infomax. In *ICLR*.

Wang, H.; Zhang, F.; Zhang, M.; Leskovec, J.; Zhao, M.; Li, W.; and Wang, Z. 2019a. Knowledge-aware graph neural networks with label smoothness regularization for recommender systems. In *KDD*, 968–977.

Wang, P.; Fu, Y.; Zhou, Y.; Liu, K.; Li, X.; and Hua, K. 2020. Exploiting Mutual Information for Substructure-aware Graph Representation Learning. In *IJCAI*.

Wang, X.; He, X.; Cao, Y.; Liu, M.; and Chua, T.-S. 2019b. Kgat: Knowledge graph attention network for recommendation. In *KDD*, 950–958.

Wang, X.; He, X.; Wang, M.; Feng, F.; and Chua, T.-S. 2019c. Neural graph collaborative filtering. In *SIGIR*, 165–174.

Wu, L.; Sun, P.; Fu, Y.; Hong, R.; Wang, X.; and Wang, M. 2019a. A neural influence diffusion model for social recommendation. In *SIGIR*, 235–244.

Wu, Q.; Zhang, H.; Gao, X.; He, P.; Weng, P.; Gao, H.; and Chen, G. 2019b. Dual graph attention networks for deep latent representation of multifaceted social effects in recommender systems. In *WWW*, 2091–2102.

Wu, X.; Huang, C.; Zhang, C.; et al. 2020. Hierarchically Structured Transformer Networks for Fine-Grained Spatial Event Forecasting. In *WWW*, 2320–2330.

Xia, L.; Huang, C.; Xu, Y.; Dai, P.; Zhang, B.; and Bo, L. 2020. Multiplex Behavioral Relation Learning for Recommendation via Memory Augmented Transformer Network. In *SIGIR*, 2397–2406.

Xia, L.; Xu, Y.; Huang, C.; Dai, P.; Zhang, X.; Yang, H.; Pei, J.; and Bo, L. 2021. Knowledge-Enhanced Hierarchical Graph Transformer Network for Multi-Behavior Recommendation. In *AAAI*.

Xin, X.; He, X.; Zhang, Y.; Zhang, Y.; and Jose, J. 2019. Relational collaborative filtering: Modeling multiple item relations for recommendation. In *SIGIR*, 125–134.

Xu, H.; Huang, C.; Xu, Y.; Xia, L.; Xing, H.; et al. 2020. Global context enhanced social recommendation with hierarchical graph neural networks. In *ICDM*.

Yang, B.; Lei, Y.; Liu, J.; and Li, W. 2016. Social collaborative filtering by trust. *TPAMI* 39(8): 1633–1647.

Yin, H.; Wang, Q.; Zheng, K.; Li, Z.; Yang, J.; and Zhou, X. 2019. Social influence-based group representation learning for group recommendation. In *ICDE*, 566–577. IEEE.

Ying, R.; He, R.; Chen, K.; Eksombatchai, P.; Hamilton, W. L.; and Leskovec, J. 2018. Graph convolutional neural networks for web-scale recommender systems. In *KDD*, 974–983.

You, J.; Ying, R.; and Leskovec, J. 2019. Position-aware graph neural networks. In *ICML*.

Yu, L.; Zhang, C.; Liang, S.; and Zhang, X. 2019. Multi-order attentive ranking model for sequential recommendation. In *AAAI*, volume 33, 5709–5716.