Package Recommendation with Intra- and Inter-Package Attention Networks

Chen Li
Beijing University of Posts and Telecommunications
WeChat Search Application Department, Tencent
Beijing, China
chenlipuleya@163.com

Yuanfu Lu
WeChat Search Application Department, Tencent
Beijing, China
luyfroot@gmail.com

Wei Wang
WeChat Search Application Department, Tencent
Beijing, China
unoywang@tencent.com

Chuan Shi*
Beijing University of Posts and Telecommunications
Beijing, China
shichuan@bupt.edu.cn

Ruobing Xie
WeChat Search Application Department, Tencent
Beijing, China
xrbsnowing@163.com

Haili Yang
WeChat Search Application Department, Tencent
Beijing, China
hailiyang@tencent.com

Cheng Yang
Beijing University of Posts and Telecommunications
Beijing, China
yangcheng@bupt.edu.cn

Xu Zhang
WeChat Search Application Department, Tencent
Beijing, China
xuonezhang@tencent.com

Leyu Lin
WeChat Search Application Department, Tencent
Beijing, China
goshawklin@tencent.com

ABSTRACT

With the booming of online social networks in the mobile internet, an emerging recommendation scenario has played a vital role in information acquisition for user, where users are no longer recommended with a single item or item list, but a combination of heterogeneous and diverse objects (called a package, e.g., a package including news, publisher, and friends viewing the news). Different from the conventional recommendation where users are recommended with the item itself, in package recommendation, users would show great interests on the explicitly displayed objects that could have a significant influence on the user behaviors. However, to the best of our knowledge, few effort has been made for package recommendation and existing approaches can hardly model the complex interactions of diverse objects in a package. Thus, in this paper, we make a first study on package recommendation and propose an Intra- and inter-package attention network for Package Recommendation (IPRec). Specifically, for package modeling, an intra-package attention network is put forward to capture the object-level intention of user interacting with the package, while an inter-package attention network acts as a package-level information encoder that captures collaborative features of neighboring packages. In addition, to capture users preference representation, we present a user preference learner equipped with a fine-grained feature aggregation network and coarse-grained package aggregation network. Extensive experiments on three real-world datasets demonstrate that IPRec significantly outperforms the state of the arts. Moreover, the model analysis demonstrates the interpretability of our IPRec and the characteristics of user behaviors. Codes and datasets can be obtained at https://github.com/LeeChenChen/IPRec.

*Corresponding author

SIGIR '21, July 11–15, 2021, Virtual Event, Canada

ACM ISBN 978-1-4503-8037-9/21/07…$15.00
https://doi.org/10.1145/3404835.3462841

CCS CONCEPTS

• **Information systems** → **Collaborative filtering**; **Social networks**;

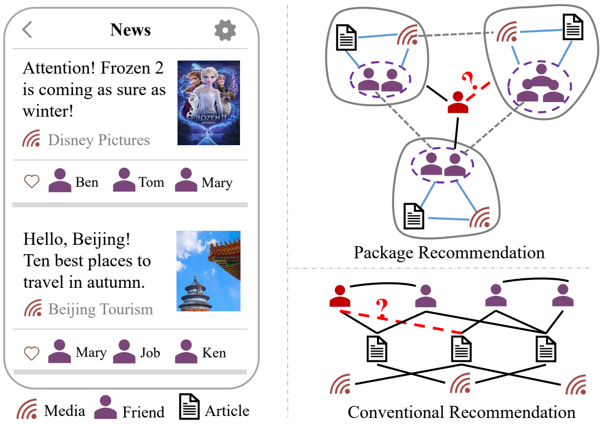
KEYWORDS

recommendation systems; neural networks; social influence

ACM Reference Format:
Chen Li, Yuanfu Lu, Wei Wang, Chuan Shi, Ruobing Xie, Haili Yang, Cheng Yang, Xu Zhang, and Leyu Lin. 2021. Package Recommendation with Intra- and Inter-Package Attention Networks. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '21), July 11–15, 2021, Virtual Event, Canada.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3404835.3462841

1 INTRODUCTION

The last decades have witnessed the booming of online social networks in the mobile internet, where people actively share opinions and obtain information. The prevalence of online social networks promotes the development of recommendation systems in various social platforms, such as *Following Feed* in *YouTube* and *Top Stories* in *WeChat*. Towards the personalized recommendation sticking to users' potential preference, recommendation systems considering



(a) Example of package recommendation (b) Comparison of recommendation paradigm

Figure 1: (a) A typical example of package recommendation in a real-world social platform. The publisher of the news and friends clicked the news are also explicitly displayed, which will have a certain influence to user. (b) Comparison between package recommendation and conventional recommendation.

the social factor (e.g., social relations or social influence) recently have also received a tremendous amount of research attention [22].

Impressed by the great successes of social platforms and recommendation systems, we present a novel social recommendation scenario in this work, named **Package Recommendation**, in which users are recommended with a combination of heterogeneous and diverse objects (e.g., news & publisher & friends or items & store & brand), rather than a single item or item list. Moreover, unlike traditional recommendation scenario that only displays the candidate items to the user, in package recommendation, objects in the package (e.g., the publisher of the news and the friends who click the news) are explicitly shown to the user, which could have a certain influence on user's behaviors and significantly change the conventional recommendation paradigm.

Package recommendation systems are blooming and have been widely used by hundreds of millions of users. Figure 1 illustrates a typical scenario of package recommendation and its difference from conventional recommendation. As shown in Figure 1(a), the user is recommended with two packages and each package comprises of a news article, the publishing media of the news, and who interacted with (e.g., shared/liked/reviewed) the news. In contrast to the traditional recommendation that only displays items to the user, the diverse objects, e.g., the publisher and friends, are also explicitly shown in package recommendation, which will highlight the influence of various objects on the user's behaviors. Still taking Figure 1(a) as an example, both recommended packages are clicked by the user, but the reasons are different. As a big fan of Disney, the user clicks the first package obviously due to his own interest since he will not miss any news published by Disney Inc. While for the second recommended package, the main reason why the user click it may be his friends (e.g., his spouse Mary) have read it, rather than attracted by the article itself. The intuitive examples clearly illustrate the differences between package recommendation and conventional recommendation, and heterogeneous objects in a package have multifaceted influence on user behaviors, which is also verified in our experiments.

While package recommendation shows a promising future on various social platforms, it is non-trivial to cope with the critical and unique packages, presenting us with two key challenges: *(1) How to fuse the heterogeneous and diverse objects in a package to capture the complex and multifaceted influence?* Each package is a combination of heterogeneous and diverse objects, and different users may be diverse in the preference of different objects. In particular, the friends within the package will have a complex social influence on people. Existing methods either ignore the influence or model the single social influence via social networks [4, 6], thus it is crucial to take the complex and multifaceted influence of objects in a package into consideration for the package recommendation. *(2) How to fuse the neighboring packages to capture the collaborative information?* Besides the intra-connections among the objects within the package, there are also inter-connections among packages. Existing recommendation algorithms are hardly to cope with critical packages [13, 17, 21]. Thus it is also vital to consider the inter-connection relationships for package modeling to capture the collaborative information.

To tackle the above challenges, we propose a novel **I**ntra- and inter-package attention network for **P**ackage **Rec**ommendation (**IPRec**). More specifically, for the first challenge, an intra-package attention network is put forward with a social influence encoder to disentangle multifaceted influence from social relations, and an interaction layer to derive the package representation encoding the complex and heterogeneous influence from diverse objects in a package. To address the second challenge, we propose an inter-package attention network that acts as a package-level information encoder to fuse collaborative features of neighboring packages. In addition, to capture the preference of user, we present a user preference learner equipped with two aggregation networks at different granularities. The fine-grained feature aggregator fuses the heterogeneous information with node- and type-level aggregations, while the coarse-grained package aggregation network aggregates the historical interacted packages. Finally, with the learned package representation and user representation, we input them to a multi-layer perceptron and predict the probability of the interaction. We summarize the contributions as follows:

- To the best of our knowledge, this work is the first attempt to study package recommendation, an emerging recommendation scenario in various online social platforms, which recommends to users a package including heterogeneous objects, rather than a single item or item list.
- To model package recommendation, we propose a novel IPRec method. To capture influence of heterogeneous objects in a package, a delicately designed package modeling module in IPRec consists of an intra-package attention network to capture the object-level attributes and multifaceted influence, and an inter-package attention network to fuse collaborative features of neighboring packages. In addition, two different levels of aggregation networks are designed to capture the preference of users.
- Extensive experiments on three real-world datasets show that IPRec significantly outperforms the state of the arts. Moreover, the model analysis also reveals the ability of IPRec to generate

recommendation explaination and discover the characteristics of user behaviours.

2 RELATED WORK

In this section, we review the related works on three areas: collaborative filtering methods, item-set recommendation methods and social recommendation methods.

2.1 Collaborative Filtering

The main basis of collaborative filtering (CF) is that users with similar preferences may be interested with similar items [20]. Matrix factorization [12] is one of the most successful methods that factorizes the rating matrix into two low-rank user-specific and item-specific latent representations then models the user preference as the inner product of the learned latent representations. After that, with the development of deep learning techniques, neural CF-based models learn a non-linear function to model user-item interactions with a multi-layer perceptron [6, 7, 29]. Recent Graph Neural Networks (GNNs) also have shed a light on CF-based recommendation systems and has shown its effectiveness and efficiency in many applications[8, 25, 27]. For example, GC-MC[1] uses a multi-link graph convolution layer to aggregate user and item features in bipartite graph. NGCF [26] explicitly incorporates collaborative signals by leveraging high-order connectivity in the user-item interaction graph. All the works mentioned above is only applicable to traditional recommendation scenarios that recommend a single item to users while not applicable to our package recommendation scenario.

2.2 Item-set Recommendation

There is a closely-related problem called item-set/bundle recommendation which recommends a set of items to users that could be purchased together [3, 13, 14]. Triple2vec [24] addresses the within-basket recommendation task by leveraging complementarity and compatibility holistically. BasConv [15] applies the Graph Convolution Network (GCN) [11] in User-Basket-Item (UBI) graph for next-basket recommendation. BGCN [2] unifies user-item interaction, user-bundle interaction and bundle-item affiliation into a heterogeneous graph and applies GCN for bundle recommendation. HFGN [14] construct a hierarchical structure upon user-outfit interactions and outfit-item mappings to recommend a set of well-matched fashion items. Different from our package recommendation where a package consists of heterogeneous objects, the aforementioned methods just recommend a single item or a set of items to users, making them hardly applicable to our scenario.

2.3 Social Recommendation

With the rise of online social platforms, social information are wildly utilized to enhance the recommendation performance. Ma et al. modeled social network information as regularization terms to constrain the matrix factorization framework [18]. TrustMF [30] factorizes social trust networks and maps users into truster space and trustee space. Recently, there emerges some works attempt to utilize GNNs to integrate social networks into recommendation systems. For example, DiffNet [28] designs a layer-wise influence propagation structure to model how users' latent embeddings evolve as the social diffusion process continues. Fan et al proposed GraphRec with a graph neural network to capture social relations and opinions in the user-item graph [4]. Although these works attempts to utilized social networks, but they all only focus on social influence and consider them as the side information to enhance the user or item representations. However, IPRec recommends a package that contains a set of heterogeneous objects and explicitly models their multiple influence to users.

3 PRELIMINARIES

In this section, we present the formal definition of a package and formalize the problem of package recommendation.

Definition 1. ***Package.*** A package is denoted as $\mathcal{P} = O^1 \cup O^2 \cup \cdots \cup O^{\mathcal{T}}$, where $O^t = \{o^t_1, o^t_2, \cdots, o^t_{|O^t|}\}$ is the object set of $t \in \mathcal{T}$ type, and $\mathcal{T}$ is the set of object types. In a package $\mathcal{P}$, there are connections between objects, referred as **intra-connection** of package $\mathcal{P}$. Two packages $\mathcal{P}_i$ and $\mathcal{P}_j$ can share one or more objects of type t, and the connection between two packages is denoted as an **inter-connection** via type t.

Example 1. As shown in Figure 1, $\mathcal{T}$ = {*Article, Media, Friend*} represents the object type set. {*Frozen2-related article*} ∪ {*DisneyPictures media*} ∪ {*Ben, Tom, Mary friends*} formulate a package $\mathcal{P}_1$. Similarly, we denote {*Beijing-related article*} ∪ {*BeijingTourism media*} ∪ {*Mary, Job, Ken friends*} as $\mathcal{P}_2$, thus $\mathcal{P}_1$ and $\mathcal{P}_2$ are inter-connected via the friend *Mary*.

Definition 2. ***Package Recommendation.*** Package recommendation aims to recommend a package, rather than an item or item list, to a user. There is a user set $U = (u_1, u_2, \cdots, u_m)$, a package set $P = (\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_n)$ and their interaction matrix $Y \in \mathbb{R}^{m \times n}$, where $y_{u,\mathcal{P}} = 1$ indicate the user u has interacted with the package $\mathcal{P}$, otherwise $y_{u,\mathcal{P}} = 0$. Given a user u w.r.t. a non-interacting package $\mathcal{P}$, the package recommendation aims to predict whether user u has a potential preference to the package $\mathcal{P}$.

4 METHODOLOGY

In this section, we first present an overview of the proposed IPRec, and further elaborate our key designs.

4.1 Overview of IPRec

To capture the multiple intentions of user interacting with a package, we propose a novel intra- and inter-package attention network for package recommendation called IPRec. Figure 2 shows the framework of IPRec with a toy example. In this example, the task is to predict the probability $\hat{y}_{u,\mathcal{P}}$ of user u interacting with $\mathcal{P}$, $\mathcal{P}_i$ and $\mathcal{P}_j$ are in inter-connections with $\mathcal{P}$, and $\{\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_N\}$ are the interacted packages of user u.

The proposed IPRec consists of three components: (1) A package modeling module including an intra-package attention network to capture the object-level intention of user interactions with the package and an inter-package attention network to capture the collaborative features of connected packages. (2) A user modeling module including two feature aggreators to indicate user interests from different perspectives. The fine-grained feature aggregation network that fuse objects features with node- and type-level aggregations, while the coarse-grained package aggregation network aggregating the historical interacted packages. (3) Finally, a predictor module inputs the learned package and user representations and outputs the probability $\hat{y}_{u,\mathcal{P}}$ of the interaction between u and $\mathcal{P}$. Note that we will illustrate our IPRec with this toy example in the following sections, while the method is suitable for general package recommendation.

4.2 Package Modeling

The first component of our IPRec is the package modeling module, which aims to learn latent representations of packages. As mentioned before, in package recommendation, there are intra-package connections linking objects in a package, and inter-package connections linking multiple packages. Thus, two attention networks, named intra-package and inter-package attention networks, are designed to respectively cope with intra-package and inter-package connections. The intra-package attention network fuses the object-level attributes, which is equipped with a social influence encoder to disentangle multifaceted influence from social relations and an interaction layer to derive the package representation encoding the complex and multifaceted influence from heterogeneous objects in a package. On the other hand, the inter-package attention network aggregates the neighboring packages by a gate attention mechanism to capture the collaborative features. In the following sections, we will present the intra-package attention network and inter-package attention network in detail.

4.2.1 Intra-Package Attention Network. As motivated, heterogeneous objects in a package contribute different to the package representation, especially friends in a package have different social influence on user behaviors. For example, a tech-expert friend may have a greater impact on users when the article in a package is related to technology while a close friend may contribute more on entertainment. Thus, we first disentangle the friend influence into different areas and then make all objects interact with each other in a package to model the complex intention of user interacting with the package.

In form, given a package $\mathcal{P} = \{O^{\tau} | \tau \in \mathcal{T}\}$ recommended to a user u, in our scenario, $\mathcal{T} = \{Article, Media, Friend\}$, where $O^{Article} = \{a\}$, $O^{Media} = \{m\}$, and $O^{Friend} = \{u_1, u_2, \cdots, u_x\}$[1]. Taking inspiration from [9, 19], we disentangle social influence of a friend u_i w.r.t. the article a as following:

$$\mathbf{u}_i^k = \mathbf{W}_f^k \mathbf{u}_i, \quad \mathbf{a}^k = \mathbf{W}_a^k \mathbf{a}, \tag{1}$$

where $\mathbf{u}_i^k$ and $\mathbf{a}^k$ respectively represent the k-th disentangled embeddings for friend u_i and article a in the package, and $k \in [1, K]$. Here $\mathbf{u}_i$ and $\mathbf{a}$ are the initial representations of u_i and a, $\mathbf{W}_f^k$ and $\mathbf{W}_a^k$ are the disentangled matrixes. Subsequently, with respect to the article a in the package $\mathcal{P}$, social influence of friends in $O^{Friend} \in \mathcal{P}$ is encoded in the k-th disentangled embedding $\mathbf{f}^k$ as:

$$\mathbf{f}^k = \sum_{u_i \in O^{Friend}} \alpha_i^k \mathbf{u}_i^k, \tag{2}$$

$$\alpha_i^k = \frac{\exp(\mathbf{z}^\top \cdot \tanh(\mathbf{w}[\mathbf{a}^k || \mathbf{u}_i^k]))}{\sum_{u_j \in O^{Friend}} \exp(\mathbf{z}^\top \cdot \tanh(\mathbf{w}[\mathbf{a}^k || \mathbf{u}_j^k]))}, \tag{3}$$

[1]Note that we truncate or padding the friend list with a max length of 20

where the operation $||$ denotes concatenation of two vectors, $\mathbf{z} \in \mathbb{R}^d$ and $\mathbf{w} \in \mathbb{R}^{d \times 2d}$ are trainable parameters for attention mechanism. In essence, the attention weight α_i^k captures the social influence of friend u_i w.r.t. the article a in the k-th disentangled space.

Intuitively, friends representations in different disentangled spaces contribute differently to social influence on the current user u, motivating us to further combine influence from K disentangled spaces with an attention mechanism:

$$\mathbf{f} = \sum_{k=1}^{K} \beta^k \mathbf{f}^k, \tag{4}$$

$$\beta^k = \frac{\exp(\mathbf{z}^\top \cdot \tanh(\mathbf{w}[\mathbf{u} || \mathbf{f}^k]))}{\sum_{k=1}^{K} \exp(\mathbf{z}^\top \cdot \tanh(\mathbf{w}[\mathbf{u} || \mathbf{f}^k]))}, \tag{5}$$

where $\mathbf{u}$ is the initial representation of user u, and $\mathbf{f}$ is the final friends embedding encoding the complex social influence on the current user u.

Next, since heterogeneous objects in a package have multifaceted influence on user behaviors, we make each object in a package interact with each other, so as to collaboratively fuse the heterogeneous and diverse information. Recall that we have the representations of article a, media m and friends $\{u_1, \cdots, u_x\}$, denoted as $\mathbf{a}$, $\mathbf{m}$ and $\mathbf{f}$, we define seven combinations as:

$$C = \{\mathbf{a}, \mathbf{m}, \mathbf{f}, \Gamma(\mathbf{a}, \mathbf{m}), \Gamma(\mathbf{a}, \mathbf{f}), \Gamma(\mathbf{m}, \mathbf{f}), \Gamma(\mathbf{a}, \mathbf{m}, \mathbf{f})\}, \tag{6}$$

where $\Gamma(\cdot)$ serves as a fusion function that can be concatenation, addition or element-wise product (we adopt element-wise product in this work). Then, we utilize an attention mechanism to distill different importance of multifaceted information for the current user u and fuse them as,

$$\mathbf{p} = \sum_{\mathbf{c} \in C} \gamma^c \mathbf{c}, \tag{7}$$

$$\gamma^c = \frac{\exp(\mathbf{z}^\top \cdot \tanh(\mathbf{w}[\mathbf{u} || \mathbf{c}]))}{\sum_{\mathbf{c}' \in C} \exp(\mathbf{z}^\top \cdot \tanh(\mathbf{w}[\mathbf{u} || \mathbf{c}']))}, \tag{8}$$

where $\mathbf{p}$ is the representation for the package $\mathcal{P}$ and $\mathbf{c} \in C$ is one combination of different objects. In fact, the information interactions between heterogeneous objects make full use of the different aspects and automatically mining the attractiveness of the user interacting with a package.

4.2.2 Inter-Package Attention Network. Besides the intra-connections among the objects of a package, there are inter-connections among the packages, which inject the collaborative information for package representations. Formally, given a package set $P = \{\mathcal{P}_1, \cdots, \mathcal{P}_{|P|}\}$, where each package $\mathcal{P}_i \in P$ has an inter-connection with the current package $\mathcal{P}$, thus we aggregate the package set with a gate filter to fuse the collaborative information as follows:

$$\mathbf{g}_i = \sigma(\mathbf{W}_1 \mathbf{p} + \mathbf{W}_2 \mathbf{p}_i + \mathbf{b}_{g_1}), \tag{9}$$

$$\tilde{\mathbf{p}} = \mathbf{p} + \sum_{\mathcal{P}_i \in P} \mathbf{g}_i \odot \mathbf{p}_i, \tag{10}$$

where $\mathbf{p}_i$ and $\mathbf{p}$ are the representations of package $\mathcal{P}_i$ and $\mathcal{P}$ learned with intra-package attention network. Here $\mathbf{W}_1$, $\mathbf{W}_2$ and $\mathbf{b}_{g_1}$ are learnable parameters of the gate filter, σ is the sigmoid function. The gate $\mathbf{g}_i$ filters the noise information and encode collaborative information into the final package representation $\tilde{\mathbf{p}}$.

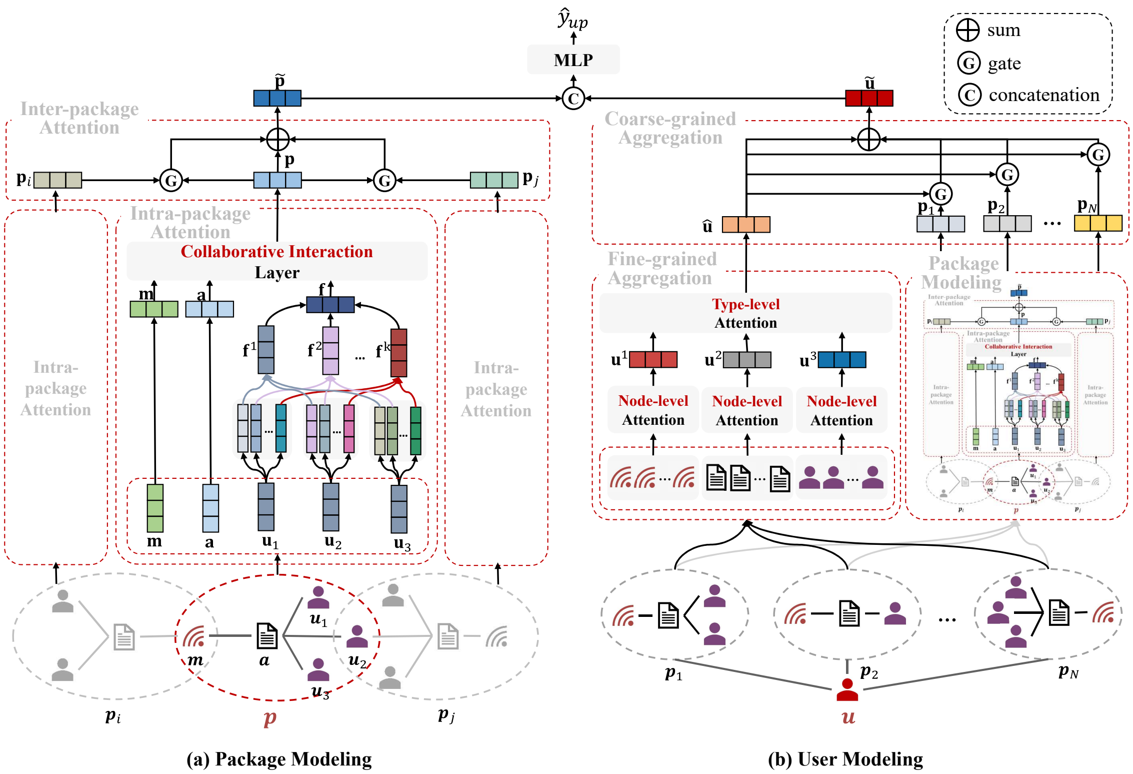


Figure 2: The framework of IPRec. It contains three components: (a) package modeling module, which learns the package embedding with intra- and inter-package attention networks to capture multifaceted influence and collaborative information; (b) user modeling module, which models the user preference at different granularities with fine-grained and coarse-grained aggregation networks; and rating prediction module, which predicts the rating with learned user and package embeddings.

4.3 User Modeling

The second component of our proposed IPRec is user modeling module, which is designed to capture the basic preference of users. Intuitively, both heterogeneous objects in a package and different packages provide multifaceted information for user preference, and they indicate user interests from different perspectives [10, 17]. Thus, we incorporate these multifaceted and heterogeneous information for user preference at two different granularities, including fusing the associated objects with a fine-grained feature aggregation network, and aggregating historical interacted packages by a coarse-grained package aggregation network.

4.3.1 Fine-Grained Aggregation Network. Considering that the multiple types of objects have different characteristics and their features may fall in different feature spaces, a well-designed hierarchical attentive aggregator is critical for objects feature aggregation to capture the heterogeneous information of different nodes and types.

Formally, given a user u, we denote different type objects associated with u as $\mathcal{A}_u = \mathcal{A}_u^1 \cup \cdots \cup \mathcal{A}_u^{\mathcal{T}}$, e.g., articles that u read or media that u subscribed. We extract all objects in the packages that the user has interacted with to capture the user preference at a fine-grained level. Firstly, we differentiate the contribution of multiple objectives with same type to the user preference aggregation. For all t-type objects in $\mathcal{A}_u^t$, we aggregate them in the t type space with a node-level attention as:

$$\mathbf{u}^t = \sigma(\sum_{v \in \mathcal{A}_u^t} \eta_{uv}\mathbf{v}), \tag{11}$$

$$\eta_{uv} = \frac{\exp(\sigma(\mathbf{w}_t^\top [\mathbf{u}||\mathbf{v}]))}{\sum_{v' \in \mathcal{A}_u^t} \exp(\sigma(\mathbf{w}_t^\top [\mathbf{u}||\mathbf{v}']))}, \tag{12}$$

where $\mathbf{u}^t$ is the t-induced representation of the user u, and $\mathbf{v}$ is the initial embeddings of object $v \in \mathcal{A}_u^t$. The term $\mathbf{w}_t \in \mathbb{R}^{2d}$ is the learnable parameter in t type space.

Further, given multiple embeddings for the user u in various type spaces $\{\mathbf{u}^1, \cdots, \mathbf{u}^{|\mathcal{T}|}\}$, we learn the attentive weights for multifaceted information in different type spaces, and then aggregate them with a type-level attention as:

$$\hat{\mathbf{u}} = \sigma(\sum_{t \in \mathcal{T}} \epsilon_u^t \mathbf{u}^t), \tag{13}$$

$$\epsilon_u^t = \frac{\exp(\mathbf{z}^\top \cdot \tanh(\mathbf{w}[\mathbf{u}||\mathbf{u}^t]))}{\sum_{t' \in \mathcal{T}} \exp(\mathbf{z}^\top \cdot \tanh(\mathbf{w}[\mathbf{u}||\mathbf{u}^{t'}]))}. \tag{14}$$

where $\hat{\mathbf{u}}$ is the fine-grained user representation that captures user preference at fine-grained level. Parameters $\mathbf{z}$ and $\mathbf{w}$ are learnable in model training.

4.3.2 Coarse-Grained Aggregation Network. To further capture the user preference at coarse-grained level, we aggregate historical interacted packages of the user u with a gate attention mechanism as follows:

$$\mathbf{g}_i = \sigma(\mathbf{W}_3\hat{\mathbf{u}} + \mathbf{W}_4\mathbf{p}_i + \mathbf{b}_{g_2}), \tag{15}$$

$$\tilde{\mathbf{u}} = \hat{\mathbf{u}} + \sum_{\mathcal{P}_i \in H} \mathbf{g}_i \odot \mathbf{p}_i, \tag{16}$$

where H is the set of packages that the user u interacted with, and $\mathbf{W}_3$, $\mathbf{W}_4$ and $\mathbf{b}_{g_2}$ are learnable parameters. The gate $\mathbf{g}_i$ filters the noise information and encode collaborative information into the final user representation $\tilde{\mathbf{u}}$.

4.4 Prediction and Optimization

Now we obtain the representations of user u and package $\mathcal{P}$, i.e., $\tilde{u}$ and $\tilde{p}$, we concatenate them and predict the probability score $\hat{y}_{u\mathcal{P}}$ of the interaction between u and $\mathcal{P}$ with a two-layer MLP:

$$\hat{y}_{u\mathcal{P}} = \sigma(\textsc{Mlp}([\tilde{\mathbf{u}}||\tilde{\mathbf{p}}])). \tag{17}$$

Finally, we optimize the following cross-entropy loss to estimate model parameters Θ:

$$\mathcal{L} = -\sum_{<u,\mathcal{P}>\in Y} (y_{u\mathcal{P}} \cdot \log(\hat{y}_{u\mathcal{P}}) + (1-y_{u\mathcal{P}}) \cdot \log(1-\hat{y}_{u\mathcal{P}})) + \lambda\,||\Theta||, \tag{18}$$

where $y_{u\mathcal{P}}$ is the ground truth and λ is the L2-regularization parameter for reducing overfitting.

4.5 Discussion

Here we give an analysis of the proposed IPRec with respect to model generality and time complexity.

Firstly, we show how IPRec generalizes traditional recommendation methods. Compared with traditional collaborative filtering methods or item-set recommendation methods that recommend a single item or item set, our IPRec recommends a package consists of heterogeneous objects, taking into consideration of multifaceted influence within and between package. If we neglect the multifaceted influence of objects, i.e, regarding the package as an item set where the heterogeneous objects are with same type, the IPRec degrades to the conventional item-set recommendation without heterogeneous information. On the other hand, if we regard the package as a single item, IPRec degrades to the traditional collaborative filtering model without social influence. Thus, package recommendation is a general recommendation scenario and the proposed IPRec is a flexible approach to capture the diverse heterogeneous information within and between packages.

Secondly, our IPRec is efficient and can be parallelized for large-scale datasets. Considering a single pair $< u, \mathcal{P} >$, the time complexity of the proposed IPRec is $O(Kxd^2+|P|d^2+N|\mathcal{P}|d^2+|\mathcal{T}|d^2+Nd^2)$, where N is the number of the interacted packages w.r.t. user u, $|\mathcal{P}|$ is the number of objects of a package. Since these quantities are all relatively small number (less than 50) and d is also set as a small number (e.g., 32 or 64), the overall complexity of IPRec is linear with the number of packages and users.

5 EXPERIMENTS

In this section, we first conduct experiments on three real-world datasets to evaluate our model performance, and then present an interpretability analysis. Lastly, we investigate the underlying mechanism of IPRec with ablation studies and parameter analysis.

Table 1: Statistics of three datasets.

	3-day	5-day	10-day
# users	554,222	778,996	1,743,824
# articles	344,231	555,328	1,136,893
# medias	166,416	222,618	1,321,274
# user-user	24,231,193	45,827,642	97,226,974
# user-package	616,454	1,129,310	2,280,469
Training set	1,990,174	3,718,736	7,916,428
Validation set	254,425	477,020	1,044,157
Test set	527,483	989,140	2,123,125
Density	4.0×10^{-7}	2.8×10^{-7}	1.2×10^{-7}

5.1 Experimental Setup

Datasets. We collect three real-world large-scale datasets from the user logs of Top Stories in WeChat, which is the biggest social platform serving more than one billion users in China and users can browse articles posted by official accounts of WeChat. To protect user privacy, we anonymize the data and conduct strict desensitization processing. The collected datasets ranges from 2020/03/01 to 2020/03/30 with different time scales and named **3-day**, **5-day** and **10-day** datasets respectively. Each dataset contains millions of interaction records generated by millions of users and packages, and each package contains a piece of article, the publishing media of the article and the friends who have interacted with the article. For each sample, the ground-truth is whether the user interacts with the article in the package. For each dataset, we split it into training, validation, and test set with a ratio of 7:1:2. To verify the robustness of our model, we vary the size of each training set from 40% to 100%. The detailed statistics of the three datasets are summarized in Table 1.

Baselines. We compare our IPRec with four categories of methods: (1) collaborative filtering methods (i.e., MF, DeepMF, NeuMF), (2) social recommendation methods (i.e., TrustMF, DiffNet), (3) graph neural network methods (i.e., GC-MC, NGCF) and (4) item-set recommendation (i.e., triple2vec, DAM). For item-set recommendation methods, we regard the package as a bundle/basket and all the heterogeneous objects in the package as the same type of items. For other methods, we regard the articles as the recommended items.

- **MF** [12] is a classic collaborative filtering method which predicts user preference by the inner product of user and item latent representations.
- **DeepMF** [29] is a deep matrix factorization model, where we apply a two layers MLP to process the interaction information of users and items.
- **NeuMF** [6] is a state-of-the-art neural CF model that uses multiple hidden layers to capture nonlinear feature interactions. Especially, we first randomly initialized the user and item embedding, and then apply NeuMF with a three-layer perceptron with latent factor sizes of 64, 32, 16 and MF for user and item embedding to capture their nonlinear and linear feature interactions.
- **TrustMF** [30] is a social recommendation method that factorizes social trust networks and maps users into truster space and

Table 2: **Experimental results on three datasets. The best method is bolded, and second best is underlined. The improvements of IPRec over the second best models are shown in the last column. '-' means that methods cannot obtain results due to out of memory.**

Methods	Ratio	MF	DeepMF	NeuMF	TrustMF	DiffNet	GC-MC	NGCF	triple2vec	DAM	IPRec	Improve.
3-day	40%	0.5655	0.6292	<u>0.6363</u>	0.6026	0.6148	0.5987	0.6090	0.5971	0.6078	**0.6538**	2.75%
	60%	0.5683	0.6310	<u>0.6438</u>	0.6078	0.6319	0.6021	0.6123	0.5976	0.6089	**0.6572**	2.08%
	80%	0.5710	0.6360	<u>0.6486</u>	0.6122	0.6432	0.6054	0.6158	0.5982	0.6103	**0.6620**	2.07%
	100%	0.5728	0.6393	0.6497	0.6171	<u>0.6520</u>	0.6092	0.6220	0.5988	0.6129	**0.6691**	2.62%
5-day	40%	0.5624	0.6189	<u>0.6350</u>	0.5987	0.6211	0.6020	0.6144	0.5644	0.5930	**0.6558**	3.28%
	60%	0.5673	0.6228	<u>0.6415</u>	0.6012	0.6355	0.6065	0.6189	0.5637	0.5967	**0.6634**	3.41%
	80%	0.5685	0.6269	<u>0.6474</u>	0.6078	0.6465	0.6112	0.6250	0.5642	0.6012	**0.6701**	3.51%
	100%	0.5712	0.6313	0.6498	0.6095	<u>0.6580</u>	0.6152	0.6310	0.5648	0.6089	**0.6754**	2.64%
10-day	40%	0.5572	0.6101	<u>0.6370</u>	-	0.6335	-	-	0.5681	0.5834	**0.6667**	4.66%
	60%	0.5623	0.6168	0.6432	-	<u>0.6522</u>	-	-	0.5683	0.5892	**0.6732**	3.22%
	80%	0.5642	0.6180	0.6478	-	<u>0.6645</u>	-	-	0.5684	0.5922	**0.6781**	2.05%
	100%	0.5658	0.6203	0.6502	-	<u>0.6700</u>	-	-	0.5687	0.5960	**0.6853**	2.29%

trustee space. Here we build the social network with the friend list in package.
- **DiffNet** [28] utilizes a graph convolution network [11] to model social influence propagation in the social network for recommendation. Here we apply a two layers of GCN as suggested in the original paper.
- **GC-MC** [1] adopts graph convolution network to learn user and item embeddings. Here we use one graph convolution layer as suggested in the original paper.
- **NGCF** [26] is a state-of-the-art graph neural network based method. Here a three layers GNN is adopted to learn the high-order collaborative embeddings.
- **triple2vec** [24] is one of the most recent works to address the within-basket recommendation task by leveraging complementarity and compatibility holistically. We consider a package as a basket within all objects are the same type and recommend the next basket.
- **DAM** [3] is specially designed for the bundle recommendation, which jointly models user-bundle interactions and user-item interactions in a multi-task manner.

We use the widely adopted metric to measure the performance of different methods, AUC [16]. A larger AUC value indicates a better performance.

5.2 Implementation Detail

We randomly initialize model parameters with Xavier initializer [5] and adopt RMSProp [23] to optimize our IPRec model. The number of disentangled space K in our IPRec is set to 4. To avoid over-fitting, we apply early stopping strategy and dropout (dropout rate is 0.4). For all methods, the learning rate, batch size, and regularization parameter are set to 0.0001, 512 and $1e-5$, respectively. Besides, we set the dimension of representation to $d = 64$ for all methods. For other parameters of baselines, we optimize them empirically under the guidance of literature. Specifically, for NeuMF and DiffNet, the learning rate is set to 0.001, for DAM, the learning rate is set to 0.0005 and the weight decay is set to 0.001 as suggested in the original papers. For NGCF and GC-MC, we apply a three graph convolution layers and one layer respectively.

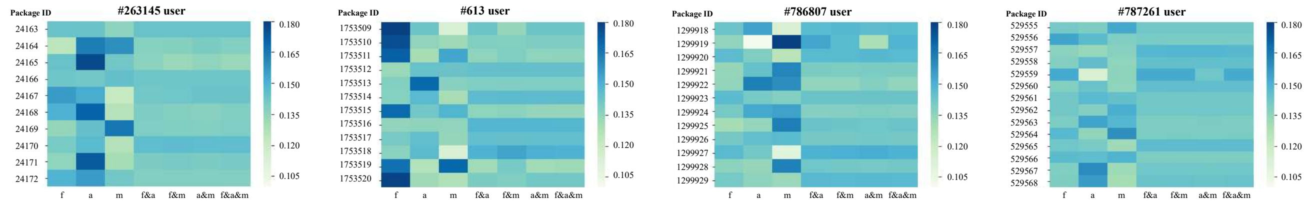
All experiments are conducted on a Linux server with four GPUs (NVIDIA GTX-1080 *4) and two CPUs (Intel Xeon E5-2690 * 2), and its operating system is Ubuntu 16.04.1. We implement the proposed IPRec with deep learning library Tensorflow[2]. The Python and Tensorflow versions are 3.6 and 1.12.0, respectively.

5.3 Performance Evaluation

We empirically compare IPRec to several state-of-the-art baselines, and vary the training ratio from 40% to 100% to draw robust results. Table 2 demonstrates the performance of all methods on three real-world datasets, and we draw the following conclusions.

Overall, our proposed IPRec consistently yields the best performance among all methods on three datasets, which brings an AUC improvement by 2.05%-4.66% compared to the best performed baseline. The significant improvement attributes to the intra- and inter-package modeling for heterogeneous and collaborative information. Among different baselines, traditional MF is least competitive since it is hardly to cope with multifaceted information in our package recommendation scenario. We notice that NGCF performs worse than NeuMF in our scenario. The reason is that an item is interacted by many users in our datasets while graph-based methods (i.e. NGCF) which aggregate all neighborhoods will cause over-smoothing problem and decrease the recommendation performance. Social recommendation methods (i.e., TrustMF and DiffNet) perform better due to the incorporation of social relations and influence, but sitll underperform our IPRec on all datasets. The reason might be that both of them simply utilize the social relations as side information, without exploring complex and multifaceted friend influence on user behaviors. In contrast, in IPRec, we design the intra-package attention network to learn the disentangled representations for friends with respect to the specific package, so as to carefully capture complex social influence and user intention in a package. Item-set recommendation methods utilize all the objects

[2] https://www.tensorflow.org



(a) Attention distribution of #263145 user. (b) Attention distribution of #613 user. (c) Attention distribution of #786807 user. (d) Attention distribution of #787261 user.

Figure 3: Micro interpretability analysis, where 'f', 'a' and 'm' mean friends, article and media respectively.

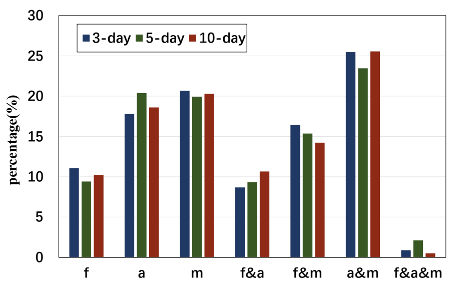


Figure 4: Macro interpretability analysis.

in a package and get a slight performance improvement but still underperform our proposed IPRec. This makes sense since these methods regard the heterogeneous objects as the same type of item, which neglects the multifaceted information.

From a vertical comparison, IPRec continues to perform best against different sizes of training data, which implies the stability and robustness of our method. Moreover, as the data scale becomes larger, the improvement of our IPRec is more obvious, indicating that IPRec is more suitable for large-scale data in real-world industrial applications. On the contrary, graph convolutional network based methods suffer from the computational efficiency for the large scale dataset, and even cannot be trained due to the out of memory issue.

5.4 Interpretability Analysis

Another major benefit of IPRec is that our recommendations are highly interpretable by capturing the actual reasons and intention of users interacting with the package. Recall that we learn the different importance of interacted information between heterogeneous objects in a package, i.e., γ^c in Eq. (8). For a user, this importance intuitively reveals the latent user preference. Thus, from micro and macro perspective, we next conduct analysis on the learned attention distribution γ^c to explore the reasons for user behaviors.

Micro Analysis. We select four users that pay attention to different factors in the package recommendation scenario, and then visualize their attention distributions of objects in their interacted packages in Figure 3. Intuitively, with respect to different users, the attention distributions of heterogeneous objects and combinations are significantly different, which indicates different reasons for users interacting with packages. More specifically, we drew the following interesting findings:

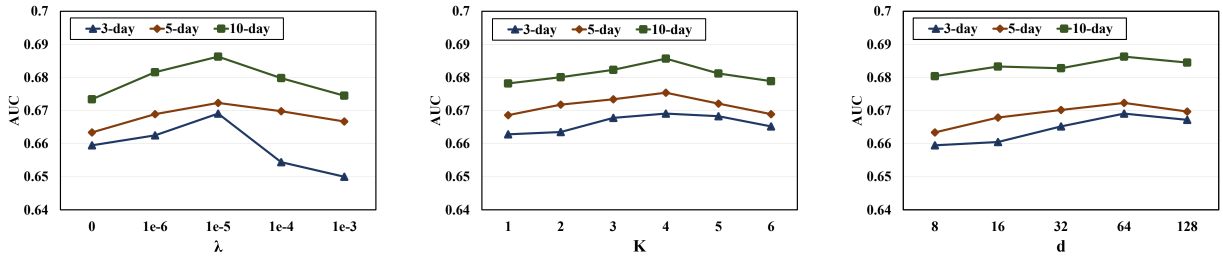
(1) #263145 user mainly focuses on the article in a package, since the attention weights of articles in most interacted packages are obviously higher than that of other objects or combinations. Actually, the attention distribution w.r.t. #263145 user also shows that he/she is a conventional user in a recommendation system, where user behaviors are significantly influenced by the article and users have their own personal preference on the news articles.

(2) However, the unique characteristic of package recommendation, i.e., recommending a combination of heterogeneous objects, makes user also influenced by other objects. As shown in Figure 3(b) and (c), #613 and #786807 users pay much more attention to friends and media in a package, respectively, rather than the article itself. In particular, #613 user is obviously affected by his/her friends who have interacted with the packages, which implies that the #613 user is a person who values social relationships and would like to read the articles recommended and filtered by friends.

(3) In addition, there is a group of people who do not have particular interests. Figure 3(d) shows an example of attention distribution of such people. The attention distribution of the #787261 user tends to be uniform which means there is no particular preference when he/she is browsing information. At one time, the #787261 user clicks the package can be attracted by articles' content, but at another time, he/she interacts with the package may be caused by the influence of friends. Because of the diversity of people's latent intents, it is critical for package recommendation to explore the underlying reasons for user behaviors and encode them into user preference.

Macro Analysis. From a global perspective, we further analyze the proportion of different reasons that cause users to interact with packages. Specifically, if the object or combination has the highest attention weight (i.e., γ^c in Eq. (8)) in most packages interacted by the user u, we take it as the main reason for u's behaviors. Thus, we can calculate the proportion of different reasons in the entire dataset, as shown in Figure 4.

Visibly , there are similar trends on the three datasets. Quite a lot of users are only concerned with single object in a package, achieving almost 50%, which is in line with the above individual analysis. The combinations of heterogeneous objects also have a great impact on users. For instance, the collaborative influence of article and media is the main reason that users interact with packages, which achieves a percentage of about 25%, much larger than that of single article or media. This makes sense, since the endorsement of the media can increase the credibility of the article, thereby having a greater impact on user behaviors.



(a) Effect of the regularization weight λ. (b) Effect of the number of disentangle spaces K. (c) Effect of the embedding dimension d.

Figure 5: Parameter analysis.

Table 3: Ablated models analysis.

Model	3-day	5-day	10-day
IPRec$_{w/o\ att}$	0.6598	0.6638	0.6745
IPRec$_{w/o\ dual}$	0.6638	0.6710	0.6800
IPRec$_{w/o\ inter}$	0.6645	0.6690	0.6814
IPRec	**0.6691**	**0.6754**	**0.6853**

5.5 Ablation Analysis

Next, we investigate the underlying mechanism of our IPRec with three ablated models: **IPRec$_{\mathbf{w/o\ att}}$** that removes the attention network and uses mean pooling for objects in a package for intra-package embeddings, **IPRec$_{\mathbf{w/o\ dual}}$** that replaces the attention layer in user modeling with mean pooling, and **IPRec$_{\mathbf{w/o\ inter}}$** that removes the inter-packages attention network and coarse-grained aggregation network.

As presented in Table 3, overall, our proposed IPRec is consistently superior to all variant models. IPRec$_{\mathrm{w/o\ att}}$ is least competitive because the mean pooling operation lose the different information in heterogeneous objects in a package, leading to a worse representation for the package. IPRec$_{\mathrm{w/o\ dual}}$ and IPRec$_{\mathrm{w/o\ inter}}$ achieves better performance on three datasets, which demonstrates the necessity to capture the different influence of heterogeneous objects. However, they still underperform the IPRec, illustrating the limitation in just modeling heterogeneous objects in a package. It is also vital for package recommendation to capture the inter-connections among packages.

5.6 Parameter Analysis

Lastly, we investigate the impact of parameters on the performance of our IPRec, including the number of spaces of disentangled embeddings for friends in a package (k), the scale of regularization weight (λ), and the embedding dimension d.

Figure 5(a) shows the impact of the number of disentangle spaces on the recommendation performance. We vary the value of K in the range of $\{1, 2, 3, 4, 5, 6\}$, while keeping the other parameters the same. We observe that our IPRec achieves optimal performance when the number of space is set to 4, and IPRec is generally stable around the optimal setting, indicating that IPRec is robust w.r.t. the number of disentangle spaces.

Figure 5(b) summarizes the impact of the L2-regularization parameter λ, with the setting of $\{0, 1e-6, 1e-5, 1e-4, 1e-3\}$. Obviously, the recommendation performance of IPRec first increases, reaching the best at $\lambda = 1e-5$, and then begins to decrease. It is reasonable that a too small or too large value of λ will cause overfitting or underfitting that decrease the performance.

Figure 5(c) displays the impact of the embedding dimension d. Visibly, our IPRec achieves the best performance at $d = 64$ setting, indicating it well express the semantic information of users and packages space. We also find that the performance of our model first increases with the growth of d and then drops. We analysis that the reason may be that too small dimension has insufficient capability of capturing the necessary information, and too large dimension introduces unnecessary noise and reduces the model generalization ability.

6 CONCLUSION

In this paper, we first studied a widely adopted recommendation scenario, named package recommendation, where users are recommended to a combination of heterogeneous objects. To model package recommendation, we proposed a novel intra- and inter-package attention network called IPRec, which captures the multifaceted influence of heterogeneous objects in a package and the collaborative information among packages. To capture the preference of users, we designed a fine-grained feature aggregation network and coarse-grained package aggregation network to model the user embeddings from different perspectives. Extensive experimental results on three real-world datasets demonstrated that our proposed IPRec consistently outperformed the state-of-the-art methods, and revealed interesting interpretabilty.

7 ACKNOWLEDGEMENT

This work is supported by the National Natural Science Foundation of China (No. U20B2045, 61772082, 61702296, 62002029). Chen Li is also supported by 2020 Tencent Rhino-Bird Elite Training Program.

REFERENCES

[1] Rianne van den Berg, Thomas N Kipf, and Max Welling. 2017. Graph convolutional matrix completion. *arXiv preprint arXiv:1706.02263* (2017).

[2] Jianxin Chang, Chen Gao, Xiangnan He, Depeng Jin, and Yong Li. 2020. Bundle recommendation with graph convolutional networks. In *Proceedings of the 43rd international ACM SIGIR conference on Research and development in Information Retrieval*. 1673–1676.

[3] Liang Chen, Yang Liu, Xiangnan He, Lianli Gao, and Zibin Zheng. 2019. Matching User with Item Set: Collaborative Bundle Recommendation with Deep Attention Network.. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. 2095–2101.

[4] Wenqi Fan, Yao Ma, Qing Li, Yuan He, Eric Zhao, Jiliang Tang, and Dawei Yin. 2019. Graph neural networks for social recommendation. In *Proceedings of the World Wide Web Conference*. 417–426.

[5] Xavier Glorot and Yoshua Bengio. 2010. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. 249–256.

[6] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural collaborative filtering. In *Proceedings of the 26th international conference on world wide web*. 173–182.

[7] Cheng-Kang Hsieh, Longqi Yang, Yin Cui, Tsung-Yi Lin, Serge Belongie, and Deborah Estrin. 2017. Collaborative metric learning. In *Proceedings of the 26th international conference on world wide web*. 193–201.

[8] Linmei Hu, Chen Li, Chuan Shi, Cheng Yang, and Chao Shao. 2020. Graph neural news recommendation with long-term and short-term interest modeling. *Information Processing & Management* 57, 2 (2020), 102142.

[9] Linmei Hu, Siyong Xu, Chen Li, Cheng Yang, Chuan Shi, Nan Duan, Xing Xie, and Ming Zhou. 2020. Graph Neural News Recommendation with Unsupervised Preference Disentanglement. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 4255–4264.

[10] Linmei Hu, Tianchi Yang, Chuan Shi, Houye Ji, and Xiaoli Li. 2019. Heterogeneous graph attention networks for semi-supervised short text classification. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*. 4823–4832.

[11] Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907* (2016).

[12] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *Computer* 42, 8 (2009), 30–37.

[13] Duc-Trong Le, Hady W Lauw, and Yuan Fang. 2019. Correlation-sensitive next-basket recommendation. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. 2808–2814.

[14] Xingchen Li, Xiang Wang, Xiangnan He, Long Chen, Jun Xiao, and Tat-Seng Chua. 2020. Hierarchical Fashion Graph Network for Personalized Outfit Recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2020.*

[15] Zhiwei Liu, Mengting Wan, Stephen Guo, Kannan Achan, and Philip S Yu. 2020. BasConv: Aggregating Heterogeneous Interactions for Basket Recommendation with Graph Convolutional Neural Network. In *Proceedings of the 2020 SIAM International Conference on Data Mining*. SIAM, 64–72.

[16] Jorge M Lobo, Alberto Jiménez-Valverde, and Raimundo Real. 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global ecology and Biogeography* 17, 2 (2008), 145–151.

[17] Yuanfu Lu, Ruobing Xie, Chuan Shi, Yuan Fang, Wei Wang, Xu Zhang, and Leyu Lin. 2020. Social influence attentive neural network for friend-enhanced recommendation. In *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*.

[18] Hao Ma, Haixuan Yang, Michael R Lyu, and Irwin King. 2008. Sorec: social recommendation using probabilistic matrix factorization. In *Proceedings of the 17th ACM conference on Information and knowledge management*. 931–940.

[19] Jianxin Ma, Peng Cui, Kun Kuang, Xin Wang, and Wenwu Zhu. 2019. Disentangled graph convolutional networks. In *Proceedings of International Conference on Machine Learning*. 4212–4221.

[20] Yue Shi, Martha Larson, and Alan Hanjalic. 2014. Collaborative filtering beyond the user-item matrix: A survey of the state of the art and future challenges. *Comput. Surveys* 47, 1 (2014), 1–45.

[21] Weiping Song, Zhiping Xiao, Yifan Wang, Laurent Charlin, Ming Zhang, and Jian Tang. 2019. Session-based social recommendation via dynamic graph attention networks. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*. 555–563.

[22] Jiliang Tang, Xia Hu, and Huan Liu. 2013. Social recommendation: a review. *Social Network Analysis and Mining* 3, 4 (2013), 1113–1133.

[23] Tijmen Tieleman and Geoffrey Hinton. 2012. Lecture 6.5-rmsprop, coursera: Neural networks for machine learning. *University of Toronto, Technical Report* (2012).

[24] Mengting Wan, Di Wang, Jie Liu, Paul Bennett, and Julian McAuley. 2018. Representing and recommending shopping baskets with complementarity, compatibility and loyalty. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*. 1133–1142.

[25] Xiang Wang, Xiangnan He, Yixin Cao, Meng Liu, and Tat-Seng Chua. 2019. Kgat: Knowledge graph attention network for recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 950–958.

[26] Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, and Tat-Seng Chua. 2019. Neural graph collaborative filtering. In *Proceedings of the 42nd international ACM SIGIR conference on Research and development in Information Retrieval*. 165–174.

[27] Xiao Wang, Ruijia Wang, Chuan Shi, Guojie Song, and Qingyong Li. 2020. Multi-component graph convolutional collaborative filtering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 6267–6274.

[28] Le Wu, Peijie Sun, Yanjie Fu, Richang Hong, Xiting Wang, and Meng Wang. 2019. A neural influence diffusion model for social recommendation. In *Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval*. 235–244.

[29] Hong-Jian Xue, Xinyu Dai, Jianbing Zhang, Shujian Huang, and Jiajun Chen. 2017. Deep Matrix Factorization Models for Recommender Systems.. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*. 3203–3209.

[30] Bo Yang, Yu Lei, Jiming Liu, and Wenjie Li. 2016. Social collaborative filtering by trust. *IEEE transactions on pattern analysis and machine intelligence* 39, 8 (2016), 1633–1647.