Bi-preference Learning Heterogeneous Hypergraph Networks for Session-based Recommendation

XIAOKUN ZHANG and BO XU, Dalian University of Technology, China
FENGLONG MA, Pennsylvania State University, USA
CHENLIANG LI, Wuhan University, China
YUAN LIN and HONGFEI LIN, Dalian University of Technology, China

Session-based recommendation intends to predict next purchased items based on anonymous behavior sequences. Numerous economic studies have revealed that item price is a key factor influencing user purchase decisions. Unfortunately, existing methods for session-based recommendation only aim at capturing user interest preference, while ignoring user price preference. Actually, there are primarily two challenges preventing us from accessing price preference. First, the price preference is highly associated to various item features (i.e., category and brand), which asks us to mine price preference from heterogeneous information. Second, price preference and interest preference are interdependent and collectively determine user choice, necessitating that we jointly consider both price and interest preference for intent modeling. To handle above challenges, we propose a novel approach Bi-Preference Learning Heterogeneous Hypergraph Networks (BiPNet) for session-based recommendation. Specifically, the customized heterogeneous hypergraph networks with a triple-level convolution are devised to capture user price and interest preference from heterogeneous features of items. Besides, we develop a Bi-Preference Learning schema to explore mutual relations between price and interest preference and collectively learn these two preferences under the multi-task learning architecture. Extensive experiments on multiple public datasets confirm the superiority of BiPNet over competitive baselines. Additional research also supports the notion that the price is crucial for the task.

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

Additional Key Words and Phrases: Session-based recommendation, price and interest preference, heterogeneous hypergraph, multi-task learning

This work is supported by the Natural Science Foundation of China (No.62376051, No.62076046 and No.62006034).

This article is extended from Reference [61] appearing in SIGIR 2022. In this extension, we (1) further introduce item brand into user behaviors modeling; (2) reconstruct the heterogeneous hypergraph to encode rich information; (3) propose a novel triple-level convolution to aggregate information on the heterogeneous hypergraph, where the co-occurrence, intra-type and inter-type relations among various nodes are explored simultaneously; (4) explore the mutual relations between price and interest preference with the help of multi-task learning; (5) reveal user intent via predicting price and interest preference collectively; (6) conduct data analysis on real-world datasets to support our motivation; (7) add new datasets and up-to-date baselines to verify the effectiveness of the proposed model; and (8) conduct a more detailed analysis about the approach, obtain more insights, and review more related literature.

Authors' addresses: X. Zhang, B. Xu, Y. Lin, and H. Lin (corresponding author), Dalian University of Technology, No. 2 Linggong Road, Dalian, China, 116024; e-mails: dawnkun1993@gmail.com, xubo@dlut.edu.cn, zhlin@dlut.edu.cn, hflin@dlut.edu.cn; F. Ma, Pennsylvania State University, University Park, PA 16801; e-mail: fenglong@psu.edu; C. Li, Wuhan University, Bayi Rd., Wuhan, China, 430075; e-mail: cllee@whu.edu.cn.

1046-8188/2023/12-ART68 $15.00
https://doi.org/10.1145/3631940

ACM Reference format:
Xiaokun Zhang, Bo Xu, Fenglong Ma, Chenliang Li, Yuan Lin, and Hongfei Lin. 2023. Bi-preference Learning Heterogeneous Hypergraph Networks for Session-based Recommendation. *ACM Trans. Inf. Syst.* 42, 3, Article 68 (December 2023), 28 pages.
https://doi.org/10.1145/3631940

1 INTRODUCTION

Recommender system (RS), which is ubiquitous in the modern information age, becomes a vital tool to alleviate information overload. RS facilitates online consumption by providing personalized services for individuals, especially in the e-commerce websites. Conventional RS [23, 35] utilizes user profiles and long-term behaviors to predict their future actions by assuming that the user identification is available. However, with increasingly strict privacy policy and widespread unregistered users, what the RS can access is merely the short behavior sequence of an anonymous user (i.e., session), where these existing conventional RS models are no longer applicable. To improve user experience in this situation, **session-based recommendation (SBR)** is proposed to provide efficient recommendation for anonymous users within a limited number of interactions [16, 25, 52].

Due to its highly practical value, SBR has received widespread attention since it appeared. Relying on the powerful learning capacity of neural networks, recent methods for SBR model user behaviors via various kinds of neural structure, such as **Recurrent Neural networks (RNN)** [16, 25], **Convolutional Neural Networks (CNN)** [42, 43], Attention Networks [22, 27, 40] and **Graph Neural Networks (GNN)** [4, 52]. Although remarkable performance has been achieved, the existing SBR methods are only committed to modeling the user *interest preference* in terms of how much a user likes an item. Another important factor, i.e., the user *price preference* in terms of how much money a user would like to pay for the item, is completely ignored by existing methods. Different from other item features (e.g., size, color, and style) indirectly influencing user bahaviors, it has been validated that the price can significantly determine whether the user would make a purchase [7, 44]. Moreover, it is purchased items instead of favorite ones that RS should predict, because that only purchase behaviors can bring profits to businesses and make users further interact with the system (like reviews). Therefore, it is urgent that considering price factor in SBR for accurately predicting the user's behaviors. However, there are mainly two challenges to be solved when we aim to incorporate price into SBR.

First, user price preference is diverse and heavily depends on some features of items, i.e., category and brand [44, 63]. Taking a common example, a user may buy *expensive camera* for work, while she could also purchase *cheap sellotape* to use in daily life. This example suggests that user price preference can dramatically change based on categories of items. Besides, some users are willing to spend extra money on items of famous brands, which is called *brand effect* in economics [7]. That means item brands can also influence user price preference. Obviously, the relevant categories and brands should be considered when we model user price preference in SBR. Under such situation, various kinds of information will be necessary to characterize user behaviors, such as *a series of items*, *item prices*, *item categories*, and *item brands*. Those *heterogeneous information* brings intractable challenges for the existing methods such as References [4, 16, 25, 40, 52–54], because that all of them are developed to model the single type information (i.e., item id). Thus, to incorporate price into SBR, the first challenge is how to copy with such heterogeneous information.

Second, user price preference and interest preference are interdependent and collectively determine user choice. It is quite common that users make price-interest tradeoffs when shopping. Also, the *price elasticity* is used by economists to describe the phenomenon that the money a user would like to spend on an item fluctuates with her interest in it [1]. It indicates that price and interest preference influence each other and contribute to user choices together. Taking an example, it is possible that a user buys an expensive item because of strong interest, even though the item price is higher than what she anticipated. In this instance, the user alters her price preference because of the influence of interest. Likewise, a user will also buy items with little interest out of their low price (it is just what frequently happened during the promotions). Obviously, to predict user purchase decisions, we have to explore the *mutual relations* between price and interest preference and *jointly model these two preferences*. However, the existing SBR methods [4, 16, 25, 40, 52–54] only use item ID as the supervised signal to model user interest preference, while can neither explicitly capture user price preference nor adequately handle rich dependencies between price preference and interest preference in determining user choice. Therefore, another challenge facing us is how to accurately model purchase intent of users under these two preferences.

For the first challenge, it is intuitive that using general heterogeneous graph to handle *heterogeneous information* [18, 30]. However, the existing heterogeneous graph-based approaches can only model pairwise (dyadic) associations [41], e.g., the relations between (*category*, *price*) or (*brand*, *price*). In our settings, there are complex high-order relations among different nodes, e.g., the triadic associations in <*price*, *category*, *brand*> jointly affect an item price semantics. Thus, the current heterogeneous graph-based methods are ineffective in our settings. Moreover, constrained by over-smoothing problem, heterogeneous graph-based methods are unable to model the dependencies between distant nodes within a session [4]. Fortunately, the hypergraph whose hyperedge can connect multiple nodes is devised recently to capture high-order data relations with arbitrary distance [54, 64]. Adopting the merits of heterogeneous graph and hypergraph, we propose *heterogeneous hypergraph networks*, which is able to capture complex high-order dependencies among various nodes, to copy with heterogeneous information in SBR.

Specifically, the heterogeneous information we consider in the heterogeneous hypergraph includes item ID, item price, item category and item brand that are directly relevant to user price and interest preference. Besides, three types of hyperedges, i.e., feature hyperedge, price hyperedge and session hyperedge, are designed to facilitate price and interest preference learning in the heterogeneous hypergraph. Based on the constructed heterogeneous hypergraph, we devise a triple-level convolution to learn node embeddings, where three kinds of relations among nodes are emphasized, i.e., co-occurrence, intra-type and inter-type relations. After that, the price and interest preferences of users are obtained by attention layers.

As to the second challenge, we should explore the *mutual relations* between price and interest preference and *take both them into account* to predict user behaviors. The multi-task learning schema is proposed to handle multiple relevant tasks jointly, where the mutual relations among these tasks can be leveraged to improve model capability on all tasks at once [62]. Inspired by this, we propose a *bi-preference learning schema*. To be specific, besides using item ID to guide the model to learn user interest preference, the bi-preference learning schema also utilizes the price of target item as another supervised signal to formulate user price preference. Under the schema of multi-task learning, it can explore the mutual relations between price and interest preference and predict both of them simultaneously. Eventually, we reveal user purchase intent based on the learned price preference and interest preference.

All in all, to incorporate price into SBR, we propose a novel approach termed **Bi-Preference Learning Heterogeneous Hypergraph Networks (BiPNet)**, where the main contributions can be summarized as follows:

- We identify and emphasize the importance of the price in influencing user behaviors and consider both price and interest preference in SBR to offer personalized services. As far as we know, this is the pioneer work to incorporate price into SBR.

— We propose a novel BiPNet to improve SBR. In BiPNet, the customized heterogeneous hypergraph networks are devised to handle heterogeneous information for price and interest preference learning. We also develop a bi-preference learning schema to predict price and interest preference simultaneously via exploring rich mutual relations between them.
— We conduct extensive experiments over multiple real-world datasets and the results show that our proposed BiPNet outperforms existing state-of-the-art SBR approaches. Further analysis also validates the significance of the price for SBR.

The rest of the article is organized as follows. Section 2 briefly reviews the literature most relevant to our work. We then conduct data analysis on a real-world dataset to examine the factors influencing user price preference and formulate the task in Section 3. In Section 4, we elaborate the proposed model BiPNet. The detailed experimental settings are introduced in Section 5. Section 6 presents the experimental results. Last, in Section 7, we conclude the article and look forward to the future work.

2 RELATED LITERATURE

Considering that this work aims to improve SBR by introducing price, proposing a customized heterogeneous hypergraph and drawing support from multi-task learning, we briefly review the related work from following four aspects: session-based recommendation, price-aware recommendation, heterogeneous graph/hypergraph, and multi-task learning.

2.1 Session-based Recommendation

With the ability to predict user actions via limited information, session-based recommendation has become a research hot-spot recently [29, 48]. Traditional methods use Markov Chain [38] or Matrix Factorization [34] to mine sequential dependencies in user behavior sequences. Enlightened by the Nearest-neighbor-based approaches [23], some methods calculate the item similarity (Item-KNN [35]) or session similarity (SKNN [19]) to determine the recommended items.

As the pioneer work introducing neural networks into SBR, GRU4Rec [16] utilizes RNN to tackle the task. Afterwards, many works improve SBR by applying RNN or its variants such as References [15, 33]. NARM [25] introduces attention mechanism into the GRU4Rec to emphasize the user main purpose. CSRM [47] further boosts the performance of NARM by exploring collaborative information from neighboring sessions. As another typical neural structure, CNN are used in SBR to capture user intent like Reference [43] with three-dimensional (3D) CNN, and Caser [42] with the standard 2D CNN. Attention networks are also popular in SBR because of its capacity for filtering out irrelevant items [22, 58, 59]. Liu et al. [27] proposed a short-term attention/memory priority model to capture both user general interest and current intents. Adopting the cloze objective from **Natural Language Processing (NLP)**, BERT4Rec [40] applies deep bi-directional self-attention to mine sequential patterns in the task. With the ability to capture pairwise patterns of item transitions, GNN have been widely used in SBR recently [30, 32, 49]. As the first method applying GNN in SBR, SR-GNN [52] views each session as a graph and adopts gated graph convolution to obtain item embeddings. LESSR [4] handles two information loss issues of GNN for SBR, i.e., lossy session encoding and ineffective long-range dependency capturing. PosRec [31] exploits positional information to improve GNN-based SBR. Relying on the ability of hypergraph in capturing complex high-order relations among nodes, Wang et al. [46] and Xia et al. [54] improved SBR by adopting hypergraph architecture. To alleviate data sparsity problem, data augmentation technique is also used to enhance model performance such as co-training [53], coupled framework [50], random walks [6], and graph augmentation [3, 39, 56]. Moreover, some methods aim to capture the fine-grained preference evolution of users via continuum model [13] or learning multi-granularity user intent [12, 57]. Considering that existing architectures of neural network perform differently on different context, some works automatically ensemble distinct networks to adapt to various session scenarios [2, 5]. More recently, some methods try to incorporate auxiliary features of item to model user intent from multiple perspectives like item category [24] and item text [17, 60]. Our earlier work CoHHN [61] emphasizes the importance of price factor on influencing user purchase behaviors. In this work, we further extend CoHHN by incorporating item brand, reconstructing the customized heterogeneous hypergraph and proposing a new triple-level convolution to aggregate rich heterogeneous information. Besides, BiPNet mines mutual relations between price and interest preference with the help of multi-task learning, where both of item ID and price are used as supervised signals to model user behaviors.

2.2 Price-aware Recommendation

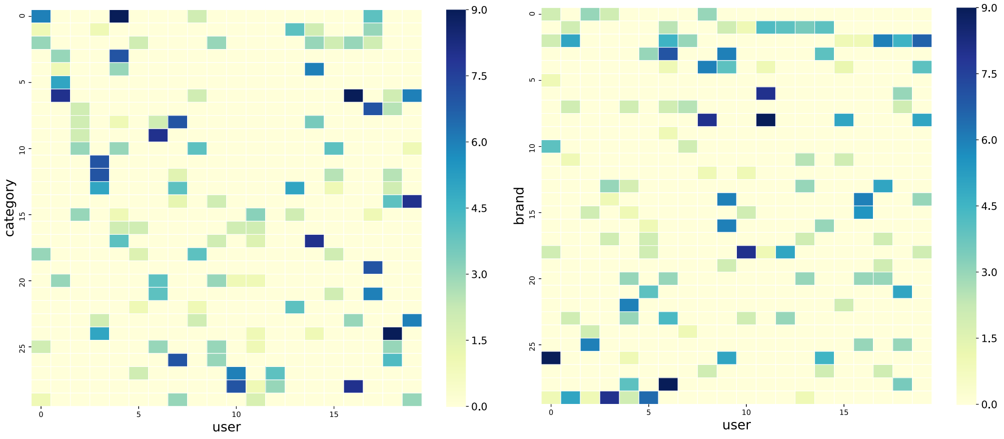
It is obvious that price is a key factor considered by users when they are shopping. Surprisingly, little effort has been made to incorporate it into the recommender system for performance improvements. As pioneers considering price in recommender system, Schafer et al. [36] explored user price sensitivity to promote online consumption. Kamishima et al. [21] integrated price personalization into recommendation algorithms to improve its commercial viability. Wan et al. [45] introduced consumer theories from economics into the lager-scale recommender systems for improving the user satisfaction. Guo et al. [11] studied the impact of multi-category inter purchase time and price on user purchase behaviors. Zheng et al. [63] modeled user price awareness by designing a unified heterogeneous graph where item price and categories are considered. Wu et al. [51] emphasized the item price competitiveness, which is used to measure the advantage of the item's price over its comparison prices, in predicting user purchase behaviors. Unfortunately, simply regarding the price factor as an auxiliary information, most of above methods ignore the vital role of price in user purchase decision. Besides, all of them use the item ID to guide the model learning, while failing to explicitly extract user price preference. In addition, there is no efforts bridging price and SBR, and we are the first to fill this gap.

2.3 Heterogeneous Graph and Hypergraph

Heterogeneous graph, also known as heterogeneous information networks, can effectively encode heterogeneous information among multi-type nodes. Many researchers formulate available data as heterogeneous graph and solve corresponding tasks via extracting relevant semantics on the heterogeneous graph. For example, Hu et al. [18] aggregated node-level and type-level information in heterogeneous graph that consists of short text, topics and entities for short text classification. Pang et al. [30] built a heterogeneous global graph containing information of historical sessions, item transitions and co-occurrence to capture user preference in SBR. Fan et al. [8] offered intent recommendation via propagating heterogeneous information with a metapath-guided method.

Hypergraph is a graph structure where a hyperedge can connect with more than two nodes [64]. Since its hyperedge can contain arbitrary number of nodes, the hypergraph is able to capture high-order relations among distinct nodes [41, 46, 54]. With this special capacity, the hypergraph has obtained increasing attention recently. For instance, Feng et al. [9] presented a hypergraph neural networks framework for learning data embeddings. Jiang et al. [20] utilized clustering methods to search optimal hypergraph structures and conducted convolution operation for representation learning. Also, some methods introduce the hypergraph structure into the recommendation task to mine beyond-pairwise relations among items [46, 54].

Considering that there are rich heterogeneous information and complex high-order relations in SBR, we combine the merits of both heterogeneous graph and hypergraph to build a novel



(a) price distribution for users on different categories (b) price distribution for users on different brands

Fig. 1. Price heatmaps for users on different categories or brands. The depth of the color means how much money a user is will to spend on the corresponding categories/brands.

heterogeneous hypergraph architecture. The customized heterogeneous hypergraph endows the proposed BiPNet with the capacity to extract price and interest preference of users.

2.4 Multi-task Learning

Multi-task learning aims to handle multiple relevant tasks jointly via exploring mutual relations among these tasks [62]. It treats all tasks equally and leverages the knowledge learned from each task to improve the model performance on different tasks. Recently, there a few works attempting to combine recommendation task with other auxiliary tasks to improve the accuracy for user modeling. In light of similar items with distinct IDs may reflect same user intent, Liu et al. [28] view item text as additional signals to learn shared intent within similar items. Shalaby et al. [37] introduced the task of predicting item category to improve the model's predictive performance. In this work, considering that price and interest preference influence each other and collectively determine user choice, we draw support from multi-task learning to deduce these two preferences simultaneously and reveal user purchase intent.

3 MOTIVATION AND PRELIMINARIES

3.1 Motivation Study

In this part, we conduct data analysis on the "Grocery and Gourmet Food" dataset from Amazon to examine the relations between user price preference and categories/brands (Section 5.2 for details of datasets). To clearly display characteristics of data distribution, we randomly select 20 users who purchased at least 20 items and interacted with more than four categories/brands from the dataset. The average price levels on each category/brand for a user is used to indicate her price preference in the category/brand (refer to Section 3.3 for price discretization). We present the price heatmap for users on different categories/brands at Figure 1, where the darker the color is, the more money the user is willing to spend on corresponding categories/brands.

As shown in Figure 1(a) and (b), a user acceptable price level is extremely diverse among different categories/brands. That is, a user can be willing to buy expensive items in some categories/brands, while buying cheap items in others. It indicates that the item category/brand does influence the user price preference. As a result, it is rational to consider these two factors when modeling user

Table 1. Main Notations Used in This Article

Notation	Description
$\mathcal{I}$, $n/\lvert\mathcal{I}\rvert$	item set, the total number of items
x_i	an item
$\mathcal{S} = [x_1, x_2, \ldots, x_m]$	a session with m items generated by an anonymous user
min, max	min/max item price in a category
$\Phi(*)$	cumulative distribution function of logistic distribution
$round(*)$	rounding operation
V and E	heterogeneous node set and hyperedge set
V^{id}, V^p, V^c and V^b	node sets with type ID, price, category and brand
p_j, c_k and b_l	a node with type price, category and brand
$\mathbf{v}_i^{id}$, $\mathbf{v}_j^p$, $\mathbf{v}_k^c$ and $\mathbf{v}_l^b$	node embeddings with type ID, price, category and brand
$\mathbf{h}_i^{id}$, $\mathbf{h}_j^p$, $\mathbf{h}_k^c$ and $\mathbf{h}_l^b$	updated embeddings for ID, price, category and brand nodes
$avg(*)$	average pooling operation
$f_a(*)$ and $f_b(*)$	intra-type and inter-type convolution operation
$\mathbf{u}_{\mathrm{p}}$ and $\mathbf{u}_{\mathrm{I}}$	the user price preference and interest preference
ρ	the number of price levels
r	the repeating number of triple-level convolution
h	the number of self-attention heads

price preference. Moreover, different users show different price preferences for a category/brand. As we can see from Figure 1(a), for a certain category, some users would like to pay high prices, while others would not. It suggests that there is divergence in user price preference, which motivates us to incorporate price into SBR for offering personalized services.

3.2 Problem Statement

This work aims to improve SBR by exploring user price and interest preference from various item features simultaneously. $\mathcal{I}$ is the item set, where $n = |\mathcal{I}|$ is the total number of items users have purchased. In general, some features are available for an item $x_i \in \mathcal{I}$, such as item ID (x_i^{id}), item price(x_i^p), item category(x_i^c), and item brand(x_i^b). In SBR, an anonymous user have purchased m items during a certain period, generating a session $\mathcal{S} = [x_1, x_2, \ldots, x_m]$, where items are chronologically ordered. Our goal is to predict the next item x_{m+1} the user would like to buy based on $\mathcal{S}$. The main notations used in this work are detailed in Table 1.

3.3 Price Discretization

As indicated in Reference [63], we cannot estimate whether an item is expensive or cheap by its absolute price. Taking an example, \$300 is cheap for a computer, but it could be extremely expensive for a T-shirt. Apparently, to intuitively compare the item price from distinct categories, we need to divide absolute price into price levels based on item category. Asnat et al. [10] found that the logistic distribution is more compatible with price distribution on a category than commonly used uniform distribution. The logistic probability density function is plotted in Figure 2, where the curve is high in the middle and low on both sides. In the middle, there are more alternatives with similar price for users to select, leading to higher user sensitivity for price, which requires us to discretize price levels more finely. For both sides, the reverse applies. Moreover, neural models generally achieve good performance under the situation of balanced data. Thus, the size of each price level should be similar. In light of the aforementioned factors, as shown in Figure 2, we divide

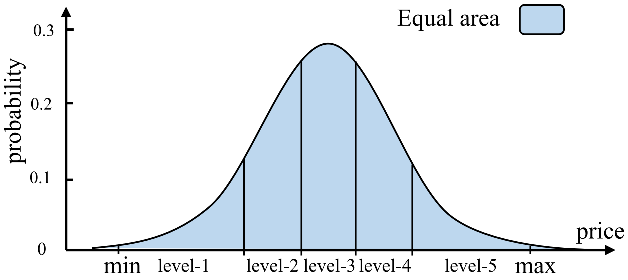


Fig. 2. Logistic distribution. We discretize absolute price into different price levels by equally partitioning price probability distribution in a category.

price into ρ levels (e.g., $\rho = 5$) by making the probability of every interval equal. More specifically, for an item whose absolute price is x_p and $[min, max]$ is the min/max price among its category, its price level p_i is set as follows:

$$p_i = round(\frac{\Phi(x_p) - \Phi(min)}{\Phi(max) - \Phi(min)} \times \rho), \tag{1}$$

where $\Phi(x)$ denotes the cumulative distribution function for logistic distribution, and we can formally defined $\Phi(x)$ as follows:

$$\Phi(x) = P(X \leq x) = \frac{1}{1 + e^{-\pi \frac{x-\mu}{\sqrt{3}\delta}}}, \tag{2}$$

where μ is expected value and δ is standard deviation both of which are estimated by maximum likelihood in this work.

3.4 Heterogeneous Hypergraph Construction

Let $\mathcal{G} = (V, E)$ be the proposed heterogeneous hypergraph that consists of node set V and hyperedge set E. A node $x_i^\tau \in V$ has a type τ and nodes belonging to the same type τ constitute a homogeneous node set V^τ. We consider four types of information in this article, i.e., item ID (V^{id}), item price (V^p), item category (V^c), and item brand (V^b), which are directly relevant to user price and interest preference. That is, $V = V^{id} \bigcup V^p \bigcup V^c \bigcup V^b$. A hyperedge $\epsilon \in E$ can contain any number of nodes with any types. Three types of hyperedges are designed to encode rich correlations among nodes: (1) a *feature hyperedge* contains four features of the purchased item, (2) a *price hyperedge* contains the price nodes appeared in a session, and (3) a *session hyperedge* contains the ID nodes within a session. Note that feature hyperedges are used to aggregate heterogeneous information in the heterogeneous hypergraph. We utilize the price/session hyperedges to mine user price/interest preference. If two nodes appear in any hyperedge simultaneously, then they are defined as having adjacent relations. The leftmost part of Figure 3 shows a toy example that illustrates how we construct the heterogeneous hypergraph.

4 THE PROPOSED APPROACH

4.1 Overview of the BiPNet

The schematic illustration of the proposed BiPNet is presented in Figure 3. On the basis of the customized heterogeneous hypergraph, a novel triple-level convolution is designed to aggregate heterogeneous information from three levels, i.e., co-occurrence level, intra-type level, and inter-type level, for node embedding learning. We then apply attention layers with temporal information to obtain price and interest preference of users. After that, a bi-preference learning schema is

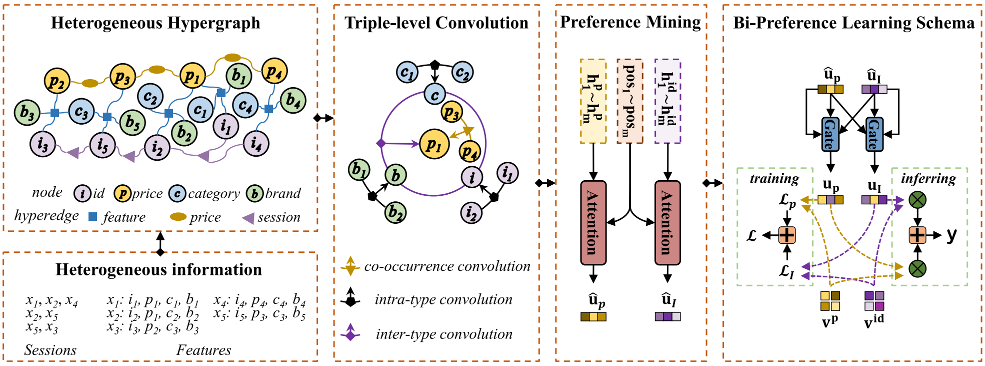

Heterogeneous Hypergraph
p_2 p_3 p_1 b_1 p_4 b_3 c_3 c_2 c_1 c_4 b_4 b_5 b_2 i_1 i_3 i_5 i_2 i_4
node id price category brand
hyperedge feature price session
Heterogeneous information
x_1, x_2, x_4
x_2, x_5
x_5, x_3
Sessions
x_1: i_1, p_1, c_1, b_1
x_2: i_2, p_1, c_2, b_2
x_3: i_3, p_2, c_3, b_3
x_4: i_4, p_4, c_4, b_4
x_5: i_5, p_3, c_3, b_5
Features
Triple-level Convolution
c_1 c_2 c p_3 p_1 p_4 b_1 b b_2 i i_1 i_2
co-occurrence convolution
intra-type convolution
inter-type convolution
Preference Mining
$\mathbf{h}_1^p \sim \mathbf{h}_m^p$
$\mathbf{pos}_1 \sim \mathbf{pos}_m$
$\mathbf{h}_1^{id} \sim \mathbf{h}_m^{id}$
Attention
Attention
$\hat{\mathbf{u}}_p$
$\hat{\mathbf{u}}_I$
Bi-Preference Learning Schema
$\hat{\mathbf{u}}_p$ $\hat{\mathbf{u}}_I$
Gate
Gate
training
inferring
$\mathbf{u}_p$ $\mathbf{u}_I$
$\mathcal{L}_p$
$\mathcal{L}$
$\mathcal{L}_I$
$\mathbf{y}$
$\mathbf{v}^p$
$\mathbf{v}^{id}$

Fig. 3. Schematic illustration of the proposed BiPNet. Based on the customized heterogeneous hypergraph, the triple-level convolution is proposed to learn node embeddings via mining co-occurrence, intra-type and inter-type relations among various nodes. Two attention layers are then utilized to mine user price and interest preference. Finally, the bi-preference learning schema is devised to jointly predict user price and interest preference by exploring multual relations between these two preferences.

devised to reveal user purchase intent via predicting price and interest preference simultaneously. In following parts, we will describe each component in detail.

4.2 Triple-level Convolution

For a target node in the heterogeneous hypergraph, there are multiple relations we should consider to build its embedding. On the one hand, as demonstrated in most recommendation algorithms [16, 23, 34, 52], items purchased together possess similar characteristics, i.e., there are *co-occurrence* relations in these items. In our settings, for a target node (e.g., p_1 in Figure 3), the nodes with same type appeared in a session (p_3, p_4) have similar semantics and can help model understand its meaning. On the other hand, for the target node (p_1), the adjacent nodes with other types (e.g., $\{c_1, c_2\} \rightarrow p_1$, $\{b_1, b_2\} \rightarrow p_1$ and $\{i_1, i_2\} \rightarrow p_1$) can also provide useful information [18]. More concretely, the nodes with a certain type (e.g., category) contain homogeneous information, where different nodes $(\{c_1, c_2\})$ may have different importance to the target node, i.e., *intra-type* relations. Also, different types furnish heterogeneous information to the target node (e.g., category → price, brand → price or ID → price), i.e., *inter-type* relations. Therefore, we propose a novel triple-level convolution to learn node embeddings via collectively investigating these three types of relations on the heterogeneous hypergraph, i.e., *co-occurrence convolution*, *intra-type convolution*, and *inter-type convolution*.

Co-occurrence convolution. The goal of the co-occurrence convolution is to enhance the node embeddings via exploring co-occurrence relations. Let $\mathbf{v}_i^t \in \mathbb{R}^d$ be the embedding of the target node x_i^t with type t. The N_i^t is the set of nodes whose type is t and having co-occurrence relations with x_i^t. For the simplicity and efficiency, the co-occurrence convolution applies the average pooling operation to extract co-occurrence information as follows:

$$\mathbf{c}_i = avg(N_i^t). \tag{3}$$

Intra-type convolution. The intra-type convolution attempts to distinguish the significance of nodes within a certain type, so as to extract relevant information of this type about the target node. For a target node $\mathbf{v}_i^t$, the N_i^τ is the set of its adjacent nodes with the type τ. The intra-type convolution prioritizes the nodes within N_i^τ to learn the specific type embedding about $\mathbf{v}_i^t$ for the type τ as follows:

$$\mathbf{e}_i^{\tau} = \sum_k \alpha_k \mathbf{v}_k^{\tau}, \tag{4}$$

$$\alpha_k = \frac{exp(\mathbf{v}_i^{t\top}\mathbf{W}\mathbf{v}_k^{\tau})}{\sum_{\mathbf{v}_j^{\tau} \in N_i^{\tau}} \exp(\mathbf{v}_i^{t\top}\mathbf{W}\mathbf{v}_j^{\tau})}, \tag{5}$$

where $\mathbf{W} \in \mathbb{R}^{d\times d}$ is the learnable parameter used to evaluate the similarity between the target node $\mathbf{v}_i^t$ and $\mathbf{v}_k^{\tau} \in N_i^{\tau}$. The $\mathbf{e}_i^{\tau} \in \mathbb{R}^d$ represents the information that is needed to propagate from type τ to $\mathbf{v}_i^t$. Considering that the nodes belonging to the same type possess homogeneous information, the intra-type convolution aggregates the information via linearly weighted addition as in Equation (4). For the convenience of the following descriptions, we simplify the intra-type convolution formulated by Equations (4) and (5) as follows:

$$\mathbf{e}_i^{\tau} = f_a(N_i^{\tau}). \tag{6}$$

Inter-type convolution. The inter-type convolution aims to integrate relevant heterogeneous information from different types into the target node. Considering that different types offer distinct semantics to the target node, the inter-type convolution fuses the type embeddings via gating operations as follows:

$$\mathbf{h}_i = \mathbf{v}_i^t + \mathbf{g}_1 \odot \mathbf{e}_i^{\tau 1} + \mathbf{g}_2 \odot \mathbf{e}_i^{\tau 2} + \mathbf{g}_3 \odot \mathbf{e}_i^{\tau 3}, \tag{7}$$

$$\mathbf{g}_1 = tanh(\mathbf{e}_i + \mathbf{W}_1\mathbf{e}_i^{\tau 1}), \tag{8}$$

$$\mathbf{g}_2 = tanh(\mathbf{e}_i + \mathbf{W}_2\mathbf{e}_i^{\tau 2}), \tag{9}$$

$$\mathbf{g}_3 = tanh(\mathbf{e}_i + \mathbf{W}_3\mathbf{e}_i^{\tau 3}), \tag{10}$$

$$\mathbf{e}_i = \mathbf{W}_a[\mathbf{v}_i^t; \mathbf{e}_i^{\tau 1}; \mathbf{e}_i^{\tau 2}; \mathbf{e}_i^{\tau 3}], \tag{11}$$

where $\mathbf{W}_a \in \mathbb{R}^{d\times 4d}$, $\mathbf{W}_1$, $\mathbf{W}_2$, and $\mathbf{W}_3 \in \mathbb{R}^{d\times d}$ are learnable parameters, $\odot$ is element-wise product, $tanh(*)$ is activate function, and [;] denotes concatenation. The $\mathbf{W}_a$ merges various heterogeneous information into $\mathbf{e}_i$, which is used to guide the information fusing next. We then utilize gating mechanism to obtain the embedding of target node $\mathbf{h}_i$. Noting that the semantics of $\mathbf{h}_i$ are enriched by rich heterogeneous information, which enables BiPNet to perceive various preferences of users. We can simplify the inter-type convolution defined by Equations (7)–(11) as

$$\mathbf{h}_i = f_b(\mathbf{v}_i^t, \mathbf{e}_i^{\tau 1}, \mathbf{e}_i^{\tau 2}, \mathbf{e}_i^{\tau 3}). \tag{12}$$

As detailed above, the triple-level convolution conducts information aggregating on the heterogeneous hypergraph by exploring co-occurrence relations, intra-type relations, and inter-type relations existing within heterogeneous nodes. Formally, the triple-level convolution updates node embeddings with four types, i.e., $\mathbf{v}_i^{id} \in V^{id}$, $\mathbf{v}_j^p \in V^p$, $\mathbf{v}_k^c \in V^c$ and $\mathbf{v}_l^b \in V^b$, as follows:

$$\mathbf{h}_i^{id} = f_b(\mathbf{v}_i^{id}, f_a(N_i^p), f_a(N_i^c), f_a(N_i^b)) + \mathbf{c}_i, \tag{13}$$

$$\mathbf{h}_j^{p} = f_b(\mathbf{v}_j^{p}, f_a(N_j^{id}), f_a(N_j^c), f_a(N_i^b)) + \mathbf{c}_j, \tag{14}$$

$$\mathbf{h}_k^{c} = f_b(\mathbf{v}_k^{c}, f_a(N_k^{id}), f_a(N_k^p), f_a(N_k^b)) + \mathbf{c}_k, \tag{15}$$

$$\mathbf{h}_l^{b} = f_b(\mathbf{v}_l^{b}, f_a(N_l^{id}), f_a(N_l^p), f_a(N_l^c)) + \mathbf{c}_l, \tag{16}$$

where $\mathbf{h}_i^{id}$, $\mathbf{h}_j^p$, $\mathbf{h}_k^c$, and $\mathbf{h}_l^b \in \mathbb{R}^d$ are updated embeddings with type of ID, price, category, and brand, respectively. With repeating the triple-level convolution r times, we can handle heterogeneous information to model user price and interest preference.

4.3 Preference Mining

After obtaining node embeddings, we further mine a user price and interest preference exposed in a session to predict her future behaviors.

Price preference mining. Intuitively, the price preference of a user is hidden in the item price within the session. Therefore, we mine user price preference based on price hyperedge, i.e., price sequence $[\mathbf{h}_1^p, \mathbf{h}_2^p, \mathbf{h}_3^p, \dots, \mathbf{h}_m^p]$. Given that the user price preference may change alone with time, we first introduce temporal information into the price embedding to encode the dynamics. Specifically, the reversed position embeddings [49, 54] is used to represent the temporal information, i.e., $\mathbf{pos}_i \in \mathbb{R}^d$. We then integrate the price embedding with position embedding as follows:

$$\mathbf{h}_i^* = tanh(\mathbf{W}_p[\mathbf{h}_i^p; \mathbf{pos}_i] + \mathbf{b}_p), \tag{17}$$

where $\mathbf{W}_p \in \mathbb{R}^{d\times 2d}$ and $\mathbf{b}_p \in \mathbb{R}^d$ are learnable parameters. $\mathbf{h}_i^* \in \mathbb{R}^d$ is the modified embedding of price node whose semantics are enriched by temporal information. Afterwards, considering that the self-attention mechanism is good at modeling transition relations within a sequence for semantics capturing [22, 40, 55], we employ multi-head self-attention to extract user price preference as follows:

$$\mathbf{E}_p = [\mathbf{h}_1^*; \mathbf{h}_2^*; \dots; \mathbf{h}_m^*], \tag{18}$$

$$head_i = Attention(\mathbf{W}_i^Q \mathbf{E}_p, \mathbf{W}_i^K \mathbf{E}_p, \mathbf{W}_i^V \mathbf{E}_p), \tag{19}$$

$$\mathbf{S}_p = [head_1; head_2; \dots; head_h], \tag{20}$$

where $\mathbf{W}_i^Q, \mathbf{W}_i^K$, and $\mathbf{W}_i^V \in \mathbb{R}^{\frac{d}{h}\times d}$ are learnable parameters that are used to convert the input to query, key, and value space, respectively, and h is the number of self-attention heads. After processing the price sequence through multi-head self-attention, we use the hidden state S_p to represent a user price preference, i.e., $\hat{\mathbf{u}}_\mathrm{p} \in \mathbb{R}^d$, as follows:

$$\hat{\mathbf{u}}_\mathrm{p} = \mathbf{S}_p^{(m)}. \tag{21}$$

Interest preference mining. We rely on the session hyperedge, i.e., item ID sequence $[\mathbf{h}_1^{id}, \mathbf{h}_2^{id}, \mathbf{h}_3^{id}, \dots \mathbf{h}_m^{id}]$, to mine user interest preference, since it is the items a user has bought that express her interest [4, 25, 54]. The user interest preference always dynamically changes according to time [48, 54, 58], so we also use the position embeddings ($\mathbf{pos}_i$) to enhance item ID embeddings as follows:

$$\mathbf{v}_i^* = tanh(\mathbf{W}_f[\mathbf{h}_i^{id}; \mathbf{pos}_i] + \mathbf{b}_f), \tag{22}$$

where $\mathbf{W}_f \in \mathbb{R}^{d\times 2d}$ and $\mathbf{b}_f \in \mathbb{R}^d$ are learnable parameters. As defined by Equation (22), the ith item in a session is represented by $\mathbf{v}_i^* \in \mathbb{R}^d$. Following the common method to formulate user interest preference [27, 52], we define the interest preference of a user as follows:

$$\hat{\mathbf{u}}_\mathrm{I} = \sum_{i=1}^{m} \beta_i \mathbf{h}_i^{id}, \tag{23}$$

$$\beta_i = \mathbf{z}\sigma(\mathbf{A}_1\mathbf{v}_i^* + \mathbf{A}_2\bar{\mathbf{v}}^* + \mathbf{b}), \tag{24}$$

where $\mathbf{A}_1, \mathbf{A}_2 \in \mathbb{R}^{d\times d}$, and $\mathbf{b} \in \mathbb{R}^{\mathbf{d}}$ are learnable parameters, $\mathbf{z}^T \in \mathbb{R}^d$ is the attention matrix used to determine the importance of an item within the session. The $\bar{\mathbf{v}}^* = \frac{1}{m}\sum_{i=1}^m \mathbf{v}_i^*$ is the average embeddings of all item ID within the session, which is utilized to evaluate the contribution of various items to user interest preference. Consequently, the user interest preference is represented as $\hat{\mathbf{u}}_\mathrm{I} \in \mathbb{R}^d$.

4.4 Bi-preference Learning Schema

As discussed before, both price preference and interest preference are indispensable in SBR. There exists rich mutual relations between these two preferences, leading to their joint determination for user purchase decisions. The multi-task learning mechanism can enhance the model capacity by leveraging the mutual information among multiple relevant tasks [62]. Inspired by this, we develop a bi-preference learning schema under the multi-task learning architecture to collectively predict price and interest preference. We first enrich the semantics of price preference $\hat{\mathbf{u}}_\mathrm{p}$ and interest preference $\hat{\mathbf{u}}_\mathrm{I}$ as follows:

$$\mathbf{u}_\mathrm{p} = \mathbf{r}_\mathrm{p} * \hat{\mathbf{u}}_\mathrm{p} + (1 - \mathbf{r}_\mathrm{p}) * \hat{\mathbf{u}}_\mathrm{I}, \tag{25}$$

$$\mathbf{u}_\mathrm{I} = \mathbf{r}_\mathrm{I} * \hat{\mathbf{u}}_\mathrm{I} + (1 - \mathbf{r}_\mathrm{I}) * \hat{\mathbf{u}}_\mathrm{p}, \tag{26}$$

$$\mathbf{r}_\mathrm{p} = \sigma(\mathbf{W}_1^p(\hat{\mathbf{u}}_\mathrm{p}) + \mathbf{W}_2^p(\mathbf{m})), \tag{27}$$

$$\mathbf{r}_\mathrm{I} = \sigma(\mathbf{W}_1^I(\hat{\mathbf{u}}_\mathrm{I}) + \mathbf{W}_2^I(\mathbf{m})), \tag{28}$$

$$\mathbf{m} = tanh(\mathbf{W}_1^{pi}(\hat{\mathbf{u}}_\mathrm{p}) + \mathbf{W}_2^{pi}(\hat{\mathbf{u}}_\mathrm{I}) + \mathbf{b}_{pi}), \tag{29}$$

where $\mathbf{W}_*^{pi}$, $\mathbf{W}_*^P$, $\mathbf{W}_*^I \in \mathbb{R}^{d\times d}$, and $\mathbf{b}_{pi} \in \mathbb{R}^d$ are learnable parameters. The $\mathbf{m} \in \mathbb{R}^d$ mergers the price and interest preferences to guide the fusion of these two kinds of information next. Referring to the gating mechanism [16], we set the "remember gate," i.e., $\mathbf{r}_\mathrm{p}/\mathbf{r}_\mathrm{I} \in \mathbb{R}^d$, to determine the degree of the price/interest information needed to retain when merging these two preferences. After that, we further enhance the representations of these two preferences ($\mathbf{u}_\mathrm{p}$ and $\mathbf{u}_\mathrm{I}$) by exploring mutual relations between them with the help of multi-task learning where user price preference and interest preference are deduced simultaneously.

Interest preference prediction. As common practices in previous works [4, 12, 16, 25, 40, 52, 53], we can formulate the task of user interest preference prediction based on user interest embedding $\mathbf{u}_\mathrm{I}$ and item ID embedding $\mathbf{v}_i^{id}$. Formally, we predict user interest preference based on $\mathbf{u}_\mathrm{I}$ and $\mathbf{v}_i^{id}$ as follows:

$$y_i^I = softmax(\mathbf{u}_\mathrm{I}^\top \mathbf{v}_i^{id}), \tag{30}$$

$$\mathcal{L}_I = -\sum_{j=1}^{n} p_j^I \log(y_j^I), \tag{31}$$

where y_i^I is predicted probability of item x_i to be liked by the user, p^I is ground-truth distribution of item ID, and $\mathcal{L}_I$ is the loss of interest preference prediction.

Price preference prediction. Both price and interest preferences play an important role in determining user choice. Thus, different from existing works [4, 12, 25, 52] that only take user interest preference into consideration, BiPNet emphasizes the significance of user price preference. To be specific, besides predicting user interest, BiPNet also introduces the item price as extra signals to formulate the task of price preference prediction. Given the learned representation of user price preference $\mathbf{u}_\mathrm{p}$ and item price $\mathbf{v}_i^p$, we can utilize the price level of next purchased item to formulate the price preference prediction task as follows:

$$y_i^p = softmax(\mathbf{u}_\mathrm{p}^\top \mathbf{v}_i^p), \tag{32}$$

$$\mathcal{L}_p = -\sum_{j=1}^{n} p_j^p \log(y_j^p), \tag{33}$$

where y_i^p is predicted probability of item price level, p^p is ground-truth price level distribution, and $\mathcal{L}_p$ is the loss of price preference prediction.

4.5 Model Training and Inferring

Training phase. At training phase, we explore rich mutual relations between price and interest preference by deducing these two preferences simultaneously. Concretely, we jointly consider the tasks of price preference prediction and interest preference prediction to reveal user purchase intent as follows:

$$\mathcal{L} = \mathcal{L}_p + \mathcal{L}_I. \tag{34}$$

Note that, different from common paradigm of multi-task learning that balances the importance of different tasks with a constant, we directly add losses of these two tasks as final loss. It implies that we view the price and interest preference to be equally crucial in affecting user behaviors.

Inferring phase. As discussed in former sections, a user purchase behaviors is determined by both her price preference and interest preference on an item. Thus, at inferring phase, we should consider these preferences collectively. Based on the learned user preferences ($\mathbf{u}_\mathrm{p}$ and $\mathbf{u}_\mathrm{I}$) and item features ($\mathbf{v}_i^p$ and $\mathbf{v}_i^{id}$), the purchased probability of an item can be calculated as follows:

$$y_i = softmax(\mathbf{u}_\mathrm{p}^\top \mathbf{v}_i^p + \mathbf{u}_\mathrm{I}^\top \mathbf{v}_i^{id}), \tag{35}$$

where the y_i represents the probability the user buys item x_i.

5 EXPERIMENTAL SETTINGS

5.1 Research Questions

To examine BiPNet's performance in SBR, we conduct a number of experiments on three public real-world datasets. Specifically, we aim to answer following research questions:

- **RQ1:** Does the proposed BiPNet achieve state-of-the-art performance compared with competitive SBR baselines? (Section 6.1)
- **RQ2:** Does each design choice contribute positively to the performance of the proposed BiPNet? (Sections 6.2–6.6)
- **RQ3:** How does the proposed BiPNet perform under different price levels? (Section 6.7)
- **RQ4:** What's the performance of SBR under various session length? (Section 6.8)
- **RQ5:** How does the key hyper-parameters influence BiPNet? (Section 6.9)
- **RQ6:** How does various information influence BiPNet's complexity? (Section 6.10)

5.2 Datasets and Preprocessing

We use following three public real-world datasets in this work:

- *Cosmetics*,[1] as a public dataset in kaggle platform, traces user actions in an online cosmetics shop. The records of a month (October 2019) are used in this work, where we choose the interactions with type "purchase" and "add_to_cart" to formulate user purchase behaviors.
- Amazon[2] records rich characteristics of users and items from the popular e-commerce website Amazon [14]. We select two representative sub-datasets "Grocery and Gourmet Food" (*Grocery*) and "Toys and Games" (*Toys*) in this work. To simulate the scenario of SBR, the user behaviors happened within one day are formulated as a session.

As the common settings in SBR [4, 25, 52–54], we delete the sessions that only contain one item in all datasets. To guarantee data quality [63], we apply 10-core settings to filter out data, where each item must appear at least 10 times and each price level/category/brand must contain at least

[1] https://www.kaggle.com/mkechinov/ecommerce-events-history-in-cosmetics-shop
[2] http://jmcauley.ucsd.edu/data/amazon/

Table 2. Statistics of Datasets

Datasets	Cosmetics	Grocery	Toys
#item	13,026	6,230	18,979
#price level	10	5	5
#category	226	550	430
#brand	148	1,306	1,429
#interaction	654,947	365,415	714,770
#session	109,251	153,383	287,733
avg.length	5.99	2.38	2.48

10 items. That is, BiPNet jointly explores item ID, price, category and brand to reveal user intents. We define the last item in a session as the ground truth and use the remaining actions to model user behaviors. The earliest 70% sessions formulate the training set. And the next 20% of sessions are used as validation set to determine the hyper-parameters of BiPNet and baselines. The remaining 10% are used as test set to evaluate the model performance. Table 2 presents the detailed statistics of all datasets.

5.3 Evaluation Metrics

To evaluate the performance of BiPNet and all baselines, we employ two widely used metrics Prec@k and MRR@k:

- Prec@k: Precision evaluates whether the ground-truth item is in the recommendation list. More specifically, Prec@k measures the proportion of cases where the ground-truth item is within the top-*k* recommendation list.
- MRR@k: Mean Reciprocal Rank considers the position of the ground-truth item in the recommendation list. Formally, MRR@k is the mean of the reciprocal rankings of ground-truth item in the list. If the rank is more than k, then the reciprocal rank is manually set to zero.

Note that, Prec@k does not take the item's actual rank into account as long as it is among the recommendation list. In contrast, the item rank is taken into consideration by MRR@k, which is significant when the recommended order matters. In addition, for both Prec@k and MRR@k, larger numbers indicate better model performance. According to conventional settings [4, 25, 40, 53], the k is set as 10 and 20 in this work.

5.4 Comparison Methods

Following competitive methods are adopted as baselines to examine the performance of the proposed BiPNet:

- *S-POP* recommends the most popular items within the current session.
- *SKNN* evaluates the similarity between current session and past sessions to determine the recommended items.
- *GRU4Rec* [16] captures sequential relations within the session via GRU.
- *NARM* [25] improves GRU4Rec by introducing attention mechanism to mine user main intention.
- *BERT4Rec* [40] uses a bidirectional self-attention to model user sequential behaviors.
- *SR-GNN* [52] views each session as a graph and applies graph neural networks on the graph to model pairwise relations between items.
- *SR-GNN+* modifies SR-GNN by adding the embeddings of item price, category and brand as inputs to examine the effect of side information on SBR.

- *LESSR* [4] tackles lossy session encoding and ineffective long range dependency of GNN for SBR.
- *S^2-DHCN* [54] introduces hypergraph architecture into SBR and further improve performance by contrastive learning.
- *COTREC* [53] explores internal and external connectivity of sessions with self-supervised learning and co-training.
- *MSGIFSR* [12] captures user fine-grained preferences evolution by modeling multi-granularity consecutive user intent.
- *MGS* [24] exploits item attribute information (e.g., categories and brands) for accurate preferences learning.
- *DGNN* [26] mines explicit and implicit item relations with GNN to improve SBR.
- *CoHHN* [61] is a preliminary version of this article, which extracts user price preference from item price and category.

5.5 Implementation Details

For fair comparison, we fix embedding size at 128 for all neural methods. We determine other hyper-parameters of BiPNet and baselines by grid search based on their performance at Prec@20 on validation set.[3] As to the main hyper-parameters of BiPNet, we investigate the number of price levels ρ in $\{2, 5, 10, 50, 100, 200\}$, the repeating number of triple-level convolution r in $\{1, 2, 3, 4\}$ and the number of self-attention heads h in $\{1, 4, 8, 16, 32\}$. Besides, we set the mini-batch size at 100 and optimize the model via Adam where the initial learning rate is 0.001. To simulate the actual scenarios of SBR as in References [4, 25, 52], the length of session is cut off at 19. We run every model five times with random initial seeds and use the average performance on text sets as final results for all approaches. The source codes are available online.[4]

6 EXPERIMENTS

6.1 Overall Performance (RQ1)

Table 3 presents the overall performance of all methods on three datasets, where we can get following insights:

First, baselines' performance exhibits considerable discrepancy across distinct datasets. For example, MGS achieves best results among baselines in Cosmetics, while its performance in Grocery and Toys is unsatisfactory. We argue that the inconsistency of baselines' performance comes from that all baselines only focus on modeling user interest preference. All of them ignore that the price preference depending on various information is also critical when users make purchase decisions. In fact, users focus on different preferences on different datasets, e.g., users possess different price sensitivity on different context. Thus, constrained by considering only single kind user preference, baselines are unable to achieve good performance across various context.

Second, drawing support from attention mechanism to distinguish the importance of items within sessions, NARM and BERT4Rec perform much better than GRU4Rec. Nevertheless, the structure of BERT4Rec does not achieve amazing performance in the task of SBR as it has done in the field of NLP. We speculate the reason for this is that the BERT structure is good at capturing distant transition patterns among tokens (i.e., items in SBR) in a sequence. As suggested in Table 2, the session usually contains a few items, i.e., the average number of items existing

[3] The settings of neural baseline methods are as follows: NARM (1 GRU layer with 128 hidden units), BERT4Rec (four heads and two layers), LESSR (four layers), S^2-DHCN (three layers), COTREC (three layers), MSGIFSR (two layers and three intent granularity levels), MGS (three samples in mirror graph), and DGNN (two layers).
[4] https://github.com/Zhang-xiaokun/BiPNet

Table 3. Overall Performance on Three Datasets

Method	Cosmetics		Grocery		Toys	
	Prec@10	MRR@10	Prec@10	MRR@10	Prec@10	MRR@10
S-POP	38.05	28.09	38.19	33.31	25.09	24.83
SKNN	42.52	31.22	67.87	47.42	45.49	32.16
GRU4Rec	21.29 ± 0.18	15.61 ± 0.13	59.38 ± 0.19	53.78 ± 0.14	35.99 ± 0.13	28.44 ± 0.10
NARM	43.99 ± 0.09	34.18 ± 0.11	69.68 ± 0.13	62.54 ± 0.08	45.92 ± 0.11	37.43 ± 0.09
BERT4Rec	40.95 ± 0.08	24.43 ± 0.07	68.60 ± 0.18	61.29 ± 0.14	43.76 ± 0.17	35.40 ± 0.14
SR-GNN	45.87 ± 0.10	34.52 ± 0.05	69.89 ± 0.14	62.78 ± 0.11	46.76 ± 0.15	37.76 ± 0.09
SR-GNN+	46.03 ± 0.12	34.49 ± 0.11	69.97 ± 0.12	62.67 ± 0.15	46.93 ± 0.11	37.81 ± 0.13
LESSR	41.98 ± 0.09	26.07 ± 0.11	68.22 ± 0.12	60.99 ± 0.16	43.80 ± 0.09	35.45 ± 0.11
S^2-DHCN	42.87 ± 0.08	33.19 ± 0.12	61.41 ± 0.13	50.58 ± 0.11	38.67 ± 0.12	30.33 ± 0.08
COTREC	46.30 ± 0.06	34.72 ± 0.08	70.03 ± 0.08	51.29 ± 0.16	47.49 ± 0.10	33.59 ± 0.13
MSGIFSR	43.70 ± 0.11	26.51 ± 0.06	69.38 ± 0.13	61.64 ± 0.09	45.80 ± 0.16	36.44 ± 0.09
MGS	47.01 ± 0.15	34.84 ± 0.09	69.92 ± 0.12	62.28 ± 0.18	47.00 ± 0.08	37.65 ± 0.11
DGNN	44.47 ± 0.18	33.86 ± 0.12	69.55 ± 0.15	62.63 ± 0.17	45.30 ± 0.13	37.59 ± 0.14
CoHHN	47.95 ± 0.12	35.02 ± 0.08	70.54 ± 0.11	63.16 ± 0.10	47.89 ± 0.14	38.54 ± 0.08
BiPNet	**48.64*** ± 0.14	**35.48*** ± 0.11	**71.12*** ± 0.10	**63.51*** ± 0.08	**48.48*** ± 0.09	**38.71*** ± 0.07

Method	Cosmetics		Grocery		Toys	
	Prec@20	MRR@20	Prec@20	MRR@20	Prec@20	MRR@20
S-POP	40.75	30.22	41.30	33.52	25.80	24.88
SKNN	49.83	32.36	70.07	47.55	48.63	32.38
GRU4Rec	23.88 ± 0.16	15.79 ± 0.17	62.53 ± 0.15	54.03 ± 0.14	38.24 ± 0.18	29.01 ± 0.13
NARM	48.52 ± 0.17	34.50 ± 0.09	71.54 ± 0.12	62.81 ± 0.10	48.42 ± 0.13	37.81 ± 0.08
BERT4Rec	49.40 ± 0.07	25.01 ± 0.06	71.01 ± 0.16	61.45 ± 0.14	47.44 ± 0.14	35.66 ± 0.10
SR-GNN	50.75 ± 0.16	34.78 ± 0.13	71.68 ± 0.14	63.03 ± 0.11	49.48 ± 0.11	38.05 ± 0.08
SR-GNN+	50.98 ± 0.12	34.90 ± 0.17	71.93 ± 0.16	63.11 ± 0.13	49.71 ± 0.09	37.92 ± 0.11
LESSR	50.18 ± 0.07	26.64 ± 0.11	70.24 ± 0.09	61.14 ± 0.08	47.10 ± 0.12	35.67 ± 0.07
S^2-DHCN	49.79 ± 0.15	33.33 ± 0.09	63.28 ± 0.17	50.87 ± 0.12	41.18 ± 0.09	30.61 ± 0.10
COTREC	51.41 ± 0.13	34.93 ± 0.07	72.17 ± 0.12	51.43 ± 0.15	50.17 ± 0.12	33.89 ± 0.08
MSGIFSR	50.15 ± 0.12	27.10 ± 0.08	71.71 ± 0.09	61.80 ± 0.07	49.69 ± 0.14	36.71 ± 0.11
MGS	52.28 ± 0.17	35.16 ± 0.11	71.87 ± 0.14	62.53 ± 0.11	50.02 ± 0.12	37.99 ± 0.09
DGNN	49.88 ± 0.14	34.17 ± 0.09	71.29 ± 0.13	62.85 ± 0.14	48.45 ± 0.09	38.10 ± 0.10
CoHHN	54.50 ± 0.16	35.42 ± 0.12	72.87 ± 0.14	63.35 ± 0.12	51.43 ± 0.15	38.81 ± 0.11
BiPNet	**55.28*** ± 0.12	**35.94*** ± 0.08	**73.66*** ± 0.13	**63.69*** ± 0.09	**52.66*** ± 0.11	**39.02*** ± 0.07

The results (%) produced by the best baseline and the best performer in each column are underlined and boldfaced respectively. Statistical significance of pairwise differences of BiPNet against the best baseline (*) is determined by the t-test ($p < 0.01$).

in a session is 5.99, 2.38, and 2.48 on Cosmetics, Grocery, and Toys, respectively. Therefore, BERT4Rec cannot unleash its strength in such short sequences, leading to its unsurprising performance.

Third, SR-GNN and LESSR obtain remarkable results by utilizing GNN to capture pairwise relations between adjacent items. Especially in the terms of MRR on Grocery and Toys, SR-GNN achieves strong performance among baselines. However, unable to model high-order relations existing among items, the overall performance of SR-GNN and LESSR is not optimal. Exploring internal and external connectivity of sessions, COTREC achieves competitive performance in terms of Precision metric. It builds graphs from distinct views, i.e., item view and session view, to augment session data. The introduction of two views enables COTREC to model beyond pairwise patterns within sessions, which enhances its performance. Similar as CORTEC, DGNN explores explicit and implicit item correlations by building distinct item graphs for complex relation modeling. In contrast, S^2-DHCN captures high-order relations among items via hypergraph convolutional networks. MGS performs admirably among baselines, especially in the Cosmetics. We argue that the introduction of auxiliary attribute information, i.e., categories and brands, endows MGS with the ability to finely-grained mine user interest, leading to its impressive performance. In addition, SR-GNN+ generally outperforms SR-GNN, which suggests that considering side information can improve SBR. Besides, the modified SR-GNN+ is still defeated by the proposed BiPNet, which customizes model structures for user price preference understanding. It verifies the unique effectiveness of the proposed BiPNet on modeling user price preference in SBR.

Finally, our proposed BiPNet achieves consistent best performance on all metrics in all datasets. Concretely, compared with the best baselines, the relative improvement of BiPNet at Prec@20 and MRR@20 is around 5.74% and 2.22% on Cosmetics, 2.06% and 0.92% on Grocery, and 4.96% and 2.41% on Toys. We believe that the superiority of BiPNet mainly comes from (1) it considers the price preference when predicting user purchase behaviors and (2) it collectively emphasizes the importance of both price and interest preference on determining user choices. In BiPNet, a customized heterogeneous hypergraph with a novel triple-level convolution is devised to capture price and interest preference of users from heterogeneous information. Besides, a bi-preference learning schema is proposed to model user intent via deducing price and interest preference simultaneously. Based on above factors, our model is able to offer satisfactory personalized services. In addition, BiPNet outperforms our previous method CoHHN. It demonstrates that (1) it is rationale to incorporate item brands to model user price preference, (2) the triple-level convolution contributes to obtaining expressive embeddings, and (3) the bi-preference learning schema is able to accurately reveal user purchase intent.

6.2 The Effect of Price Information (RQ2)

The main contribution of our proposed BiPNet is that it considers price information in SBR and regards the price preference as important as interest preference in determining user purchase behaviors. We remove the price information in the BiPNet to build a variant named BiPNet-p. Moreover, we only retain the price information to learn the node embeddings in the triple-level convolution, while not deliberately mining user price preference subsequently. We name this variant as BiPNet-pp. In addition, to further investigate the utility of price information in SBR, we incorporate item price into two representative baselines NARM and SR-GNN to formulate NARM+p and SR-GNN+p. Specifically, NARM+p and SR-GNN+p view the addition of item ID and price embeddings as inputs.

From Table 4, we can observe that (1) the BiPNet-pp achieves better performance than BiPNet-p in all cases. By utilizing item price information for updating item embeddings, the BiPNet-pp is able to perceive user price sensitivity, which makes it perform better than BiPNet-p. It suggests that considering price contributes to understanding user intent, thus increasing the recommendation accuracy. It is also in accord with reality that the price is an important factor for users to consider when shopping. (2) The BiPNet outperforms BiPNet-pp in a large margin, which indicates that we should view the price as a critical factor instead of accessory one when modeling user behaviors. It also verifies our point of view that the user price preference plays an significant role in determining user choices and we should view the price preference as important as interest one to provide satisfactory personalized services. (3) Generally, the baseline methods considering item price outperform its counterparts without, i.e., NARM+p vs. NARM and SR-GNN+p vs. SR-GNN. It highlights once more that taking the price information into account can improve SBR.

Table 4. The Effect of Price Information

Methods	Cosmetics		Grocery		Toys	
	Prec@20	MRR@20	Prec@20	MRR@20	Prec@20	MRR@20
NARM	48.52	34.50	71.54	62.81	48.42	37.81
NARM+p	48.86	34.69	71.79	62.95	48.28	37.93
SR-GNN	50.75	34.78	71.68	63.03	49.48	38.05
SR-GNN+p	50.86	34.97	71.84	62.97	49.63	38.12
MGS	52.28	35.16	71.87	62.53	50.02	37.99
CoHHN	54.50	35.42	72.87	63.35	51.43	38.81
BiPNet-p	54.38	35.12	72.89	62.61	51.79	37.65
BiPNet-pp	54.85	35.68	73.36	62.80	52.25	38.03
BiPNet	**55.28***	**35.94***	**73.66***	**63.69***	**52.66***	**39.02***

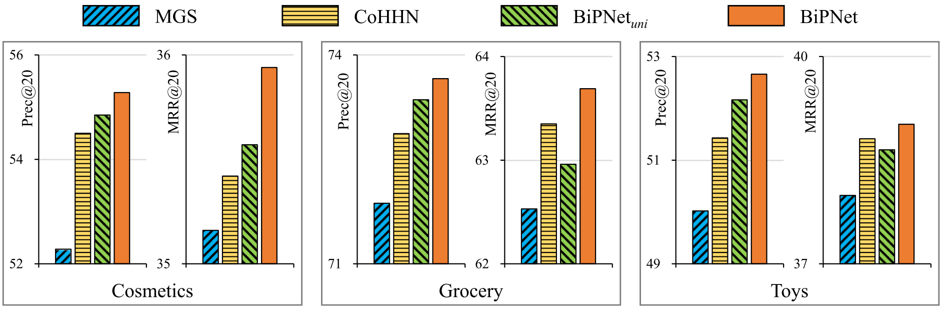


Fig. 4. The effect of the price discretization.

6.3 The Effect of Price Discretization (RQ2)

As discussed in Section 3.3, to compare the degree of price across various categories, we divide absolute price into price levels. Different from common practice where item price levels are simply defined by uniform quantization, we find that the price distribution tends to be logistic distribution and perform price discretization via equally partitioning the logistic probability distribution. The BiPNet$_{uni}$ obtains the price levels by uniform quantization like in Reference [63], where $p_i = \lfloor \frac{x_p - min}{max - min} \times \rho \rfloor$.

Figure 4 presents the results of MGS, CoHHN, BiPNet$_{uni}$, and BiPNet in terms of Prec@20 and MRR@20 on all datasets. BiPNet achieves better performance than BiPNet$_{uni}$, which indicates the superiority of our proposed method for price discretization. We conjecture the reasons for the good performance are (1) the price distribution is more compatible with logistic distribution rather than uniform one, and (2) equal probability interval makes the price reasonably arranged in each price level and enables the model to finely perceive price preference of users. In addition, BiPNet$_{uni}$ outperforms MGS by a large margin, which proves the benefits of introducing price into SBR again.

6.4 The Influence of Heterogeneous Information (RQ2)

The user price preference varies greatly according to item features, e.g., category and brand. Therefore, besides item price, we introduce extra information into our BiPNet to facilitate price preference modeling. In this part, we zoom into each type information to examine its influence on BiPNet's performance. The ablation results are shown in Table 5, where "p," "c," and "b" mean price, category, and brand information, respectively. We can get following insights from Table 5:

Table 5. The Influence of Heterogeneous Information

methods	Cosmetics		Grocery		Toys	
	Prec@20	MRR@20	Prec@20	MRR@20	Prec@20	MRR@20
(a) w/o p	54.38	35.12	72.89	62.61	51.79	37.65
(b) w/o c	54.43	35.45	73.21	63.05	52.18	38.72
(c) w/o b	54.74	35.67	72.93	63.07	51.90	38.80
(d) w/o pc	53.01	35.09	73.16	62.68	52.19	37.66
(e) w/o pb	54.54	35.56	73.03	62.55	51.96	37.45
(f) w/o cb	54.27	35.41	72.18	62.81	51.11	38.63
(g) w/o pcb	52.15	32.67	71.87	61.63	50.94	36.84
BiPNet	**55.28***	**35.94***	**73.66***	**63.69***	**52.66***	**39.02***

(1) The variant (g) without incorporating any extra information obtains the worst performance. It proves that considering extra features in SBR does contribute to revealing user intent, increasing the prediction accuracy. (2) Except the BiPNet, there is no variants consistently performing well on all cases. It indicates that the user behaviors present complexity and are influenced by multiple item features. Going further, each type of information we consider in the model has unique impact on user decisions in different context. (3) In general, the variants incorporating price information perform better than the variants do not, e.g., variants ((b) and (c)) against (a). Especially, variant (f) that only considers item price achieves overwhelming superiority over variant (g) that does not incorporate any side information. It demonstrates again that price is indeed a crucial factor in determining user behaviors and taking it into account brings significant improvements for SBR. (4) The variant only containing price while ignoring other factors performs worse than variants considering price and related features, e.g., variant (f) is defeated by variant ((b) and (c)). It verifies our hypothesis that the user price preference is related to different item features. Thus, only incorporating price information without considering related features cannot accurately extract price preference of users. (5) BiPNet achieves best performance in all cases, which demonstrates that the proposed model can handle the heterogeneous information in SBR, leading to good understanding of user behaviors.

6.5 The Effect of Triple-level Convolution (RQ2)

Considering that there are mainly three kinds of relations among heterogeneous nodes, i.e., co-occurrence relations, intra-type relations, and inter-type relations, we propose a novel triple-level convolution to aggregate rich information in the heterogeneous hypergraph for node representation learning. To demonstrate the effectiveness of this design choice, we replace the triple-level convolution with commonly used GCN like in Reference [63] to formulate the variant BiPNet$_{gcn}$. That is, for a target node, BiPNet$_{gcn}$ obtains its embedding by linearly aggregating embeddings of its adjacent nodes without regard to distinct relations among them. Moreover, compared with CoHHN, BiPNet further enhances the node representation learning by performing co-occurrence convolution. To examine its utility, we discard co-occurrence convolution from BiPNet and obtain a variant BiPNet-co.

As presented in Figure 5, BiPNet performs much better than BiPNet$_{gcn}$ in both metrics on all datasets, which shows the superiority of the proposed triple-level convolution in representation learning under such complex heterogeneous situation. We believe its superiority comes from that the triple-level convolution finely distinguishes distinct relations existing among various nodes and designs customized method to extract meaningful information for every relation. It endows the BiPNet with the capacity to obtain accurate node embeddings from rich heterogeneous

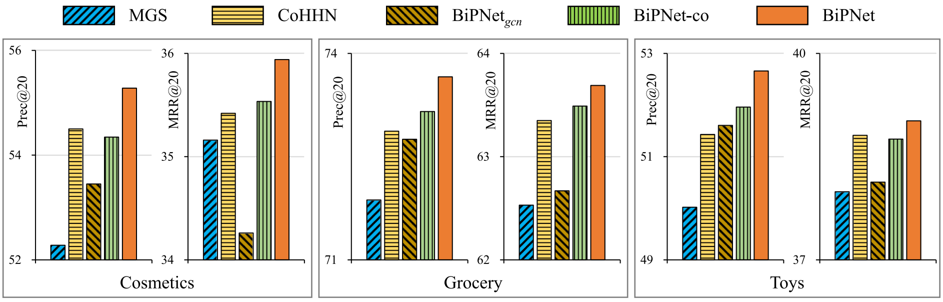


Fig. 5. The effect of triple-level convolution.

Table 6. The Effect of Bi-preference Learning Schema

methods	Cosmetics		Grocery		Toys	
	Prec@20	MRR@20	Prec@20	MRR@20	Prec@20	MRR@20
MGS	52.28	35.16	71.87	62.53	50.02	37.99
CoHHN	54.50	35.42	72.87	63.35	51.43	38.81
BiPNet-BiP	54.76	35.43	73.21	62.80	51.86	38.41
BiPNet_{CE}	54.99	35.74	73.41	62.52	52.38	38.53
BiPNet	**55.28***	**35.94***	**73.66***	**63.69***	**52.66***	**39.02***

information, thus clearly revealing user intent. Besides, BiPNet_{gcn} considering item price, categories, and brands generally performs better than MGS incorporating item categories and brands (except on MRR@20 in Cosmetics), which demonstrates the significance of price in influencing user behaviors in SBR again. In addition, BiPNet achieves better performance than BiPNet-co, which verifies the effectiveness of co-occurrence convolution in BiPNet. It is intuitive, since nodes with co-occurrence relations possess similar semantics that can contribute to learning meaningful node embeddings.

6.6 The Effect of Bi-preference Learning Schema (RQ2)

As stated in former sections, a user often makes price-interest tradeoff when she shops online. Obviously, there are complex mutual relations between user price preference and interest preference. Both price and interest preference are indispensable in SBR and they collectively determine user choices. To accurately capture these two preferences, we propose a bi-preference learning schema under the multi-task learning architecture. To validate the effectiveness of the proposed schema, we remove it from the model to build a variant BiPNet-BiP. The BiPNet-BiP directly applies the price $\hat{\mathbf{u}}_{\mathrm{p}}$ and interest preference $\hat{\mathbf{u}}_{\mathrm{I}}$ into Equation (35) to formulate the recommendation list. Besides, we replace the multi-task learning paradigm with commonly used cross-entropy loss [24, 25, 52, 61] to formulate BiPNet_{CE}. That is, BiPNet_{CE} applies single loss based on Equation (35) for model training while failing to incorporate extra signals to collectively deduce user price and interest preference.

As shown in Table 6, BiPNet outperforms BiPNet-BiP in terms of Prec@20 and MRR@20 in all datasets, which indicates the effectiveness of the proposed bi-preference learning schema for the task. It also suggests that there exists complex mutual relations in price and interest preference. The bi-preference learning schema can explore the mutual relations between price and interest preference, leading to BiPNet's good performance. Moreover, BiPNet_{CE} is defeated by BiPNet, which demonstrates the rationale and necessity of the multi-task learning paradigm in our settings.

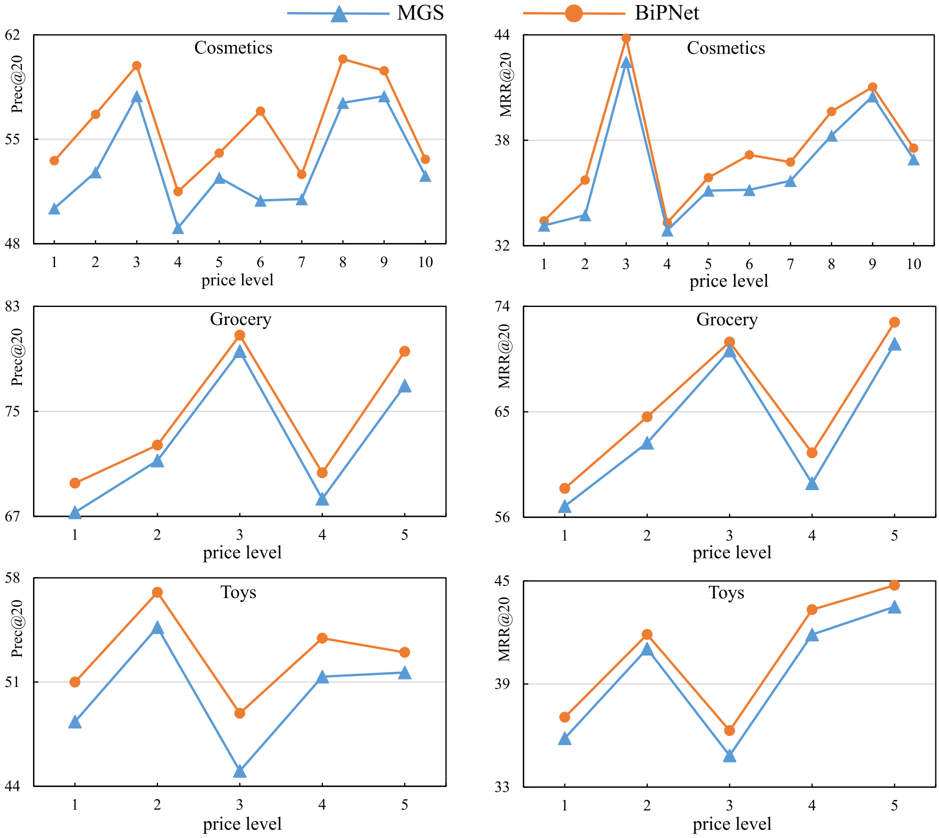


Fig. 6. Performance under different price levels.

With the help of the multi-task learning architecture, the proposed BiPNet is able to accurately deduce user price and interest preference, contributing to its prediction for user behaviors. In addition, both BiPNet-BiP and BiPNet$_{CE}$ achieve much better performance than MGS, especially in terms of Prec@20. We believe that the good performance comes from their modeling for price preference, which proves the significance of taking item price into account once more in SBR.

6.7 Performance under Different Price Levels (RQ3)

The former sections have indicated that both price and interest preference contribute to predicting user purchase behaviors. Zooming in each preference, nevertheless, users present great behavioral discrepancy. With possessing distinct tastes, there are high personalization in interest preference for different users. For example, some users like T-shirts with round neck, while some others may prefer polo shirt. As to price preference, in contrast, most users tend to prefer items with low price. This is straightforward to understand, since people always want to spend as little money as possible to get desired items. Thus, we could improve the overall performance by blindly providing users with cheap items. However, such a move will cause decline of business revenue given that expensive items always offer huge profits. Therefore, we conduct experiments under different price levels to examine the performance of the proposed BiPNet and competitive baseline MGS. Figure 6 plots their performance patterns on all datasets in terms of Prec@20 and MRR@20, where we can get following observations:

On the one hand, BiPNet outperforms MGS on all cases in all datasets. It indicates that our BiPNet can understand user needs in different price levels and provide accurate recommendation accordingly. In other words, BiPNet can not only improve user satisfaction but also increase the income of business. We believe that this attribute of the BiPNet can promote the benign development of e-commerce. On the other hand, models' performance patterns show different trends in distinct datasets. More specifically, in Cosmetics and Toys, models perform better in low (e.g., 3 in Cosmetics and 2 in Toys) and high (e.g., 8, 9 in Cosmetics and 4, 5 in Toys) price than in medium one (e.g., 4, 5, 6, 7 in Cosmetics and 3 in Toys). Instead, in Grocery, they achieve good performance in medium price (e.g., 3), while perform badly in low (e.g., 1, 2) and high price (e.g., 4). We speculate that the discrepancy comes from distinct behavior characteristics in different context (datasets). For instance, in Cosmetics, users usually either choose cheap items for daily use or buy expensive ones for occasional cases. While items with medium price are rarely noticed. As to Grocery, the quality of cheap food is worrying. Also, it is unnecessary to buy expensive ones in most cases. Thus, items with medium price are commonly popular. Obviously, considering item price in e-commerce can enable the business to detect user fine-grained behavior characteristics and offer satisfactory personalized services accordingly.

6.8 Impact of the Session Length (RQ4)

The session length is a key factor that influences model performance in SBR, since it signifies how much information the model can rely on to capture user preference. In this part, therefore, we explore the performance of BiPNet and representative baseline MGS under different session length in all datasets. Based on average session length of the dataset (as shown in Table 2), we choose different ranges to display models' performance patterns in three datasets, i.e., [1, 10] for Cosmetics and [1, 5] for Grocery/Toys. The BiPNet's and MGS's performance curves in terms of Prec@20 and MRR@20 are presented in Figure 7. Note that, due to significant performance gap existing on MRR@20 in Cosmetics, we set two unique scales in y-axis to highlight the performance patterns, i.e., we can refer to the right y-axis to see the performance of length 3–10 for clarity. The following insights can be concluded from Figure 7:

(1) BiPNet outperforms MGS in all cases, which demonstrates the effectiveness of BiPNet in the task of SBR. Compared with MGS, the proposed BiPNet not only considers user interest preference but also emphasizes the significance of user price preference hidden in various item features like price, categories, and brands. We contend that the consistently strong performance of BiPNet on various session length comes from its careful consideration for user price preference. (2) Generally speaking, as the session length increases, the performance of both BiPNet and MGS deteriorates. That is, models perform unsatisfactorily under long sessions. As suggested in previous works, in long sessions, there may exist much more noisy clicks [58] and multiple user intents [57]. As a result, it is challenging for these models to accurately predict user behaviors under such a complex situation. (3) Overall, compared with long sessions, BiPNet achieves larger improvements over MGS in short ones. In sessions with a few items (e.g., session length is 1 or 2), there is a little information available for models to capture user intent. In another word, the models have a serious problem with data sparsity. Fortunately, the incorporation of item side information in BiPNet enriches the data, alleviating data sparsity issue to some extent. It also reminds us the rationale of modeling various user preferences in SBR.

6.9 Hyper-parameter Study (RQ5)

In this part, we investigate the influence of three main hyper-parameters, i.e., the number of price levels ρ, the repeating number of triple-level convolution r and the number of self-attention heads h, on our proposed BiPNet.

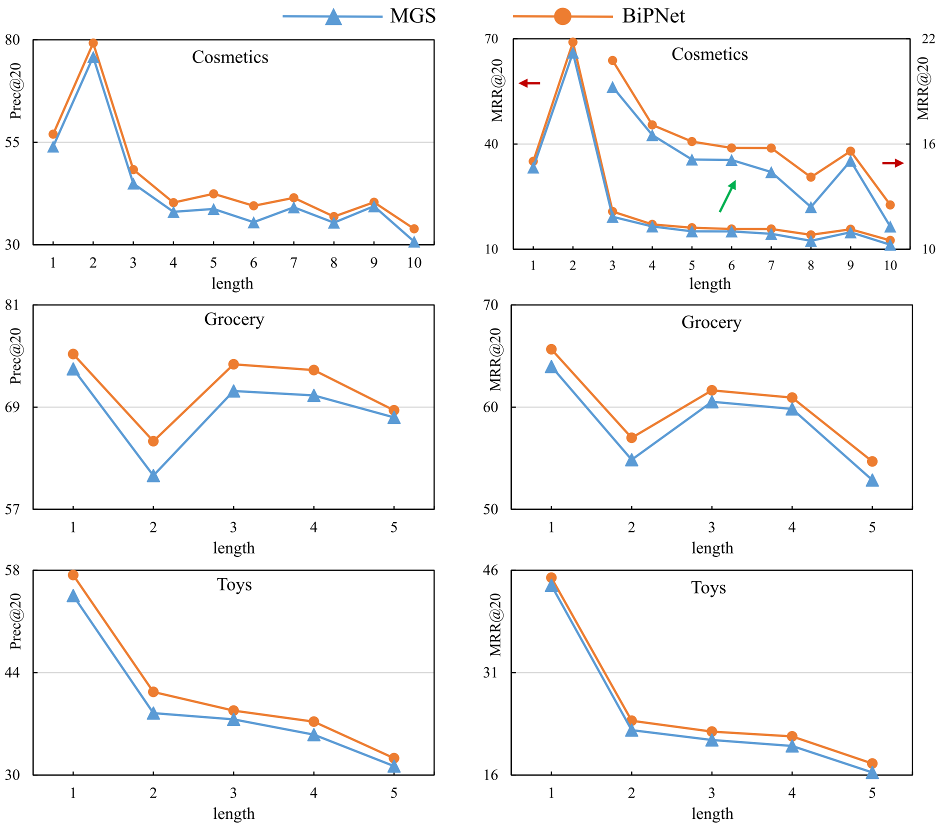
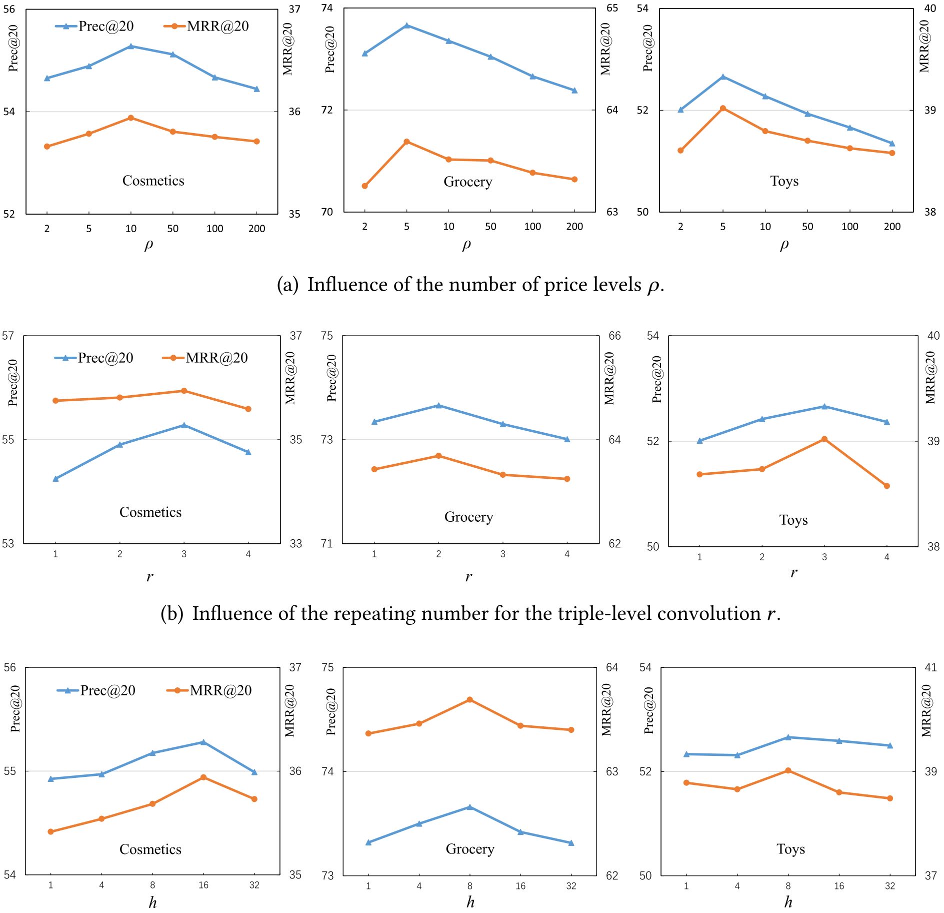


Fig. 7. Impact of the session length.

The number of price levels ρ controls the fineness of price levels. The larger the ρ, the finer the granularity of price discretization, where the price sensitivity of users is assumed higher. As can be seen from Figure 8(a), if we define ρ too small, i.e., 2, which signifies that an item price is coarsely viewed as either expensive or cheap, then BiPNet cannot perceive different price preferences, leading to invalid price modeling and unsatisfactory predictions. In contrast, when the ρ is too large like 200, the similar price is allocated into different price levels, leading to fragmented price levels. In such a case, the difference among price levels is not evident, which is also detrimental to model performance. Moreover, in the view of absolute value, the performance under different ρ is stable. We believe that the stability comes from the method we developed to form price levels. Equally partitioning the price probability distribution, our method effectively allocates absolute price into each level and balances the training data, which helps the model accurately capture various price preferences. In addition, BiPNet achieves best results under different ρ in different datasets, i.e., 10 for Cosmetics and 5 for Grocery/Toys. We guess that users present different price sensitivity under different context (datasets), leading to different optimal values for ρ.

The repeating number of triple-level convolution r determines how much information a node incorporates from its adjacent nodes. Obviously, every time we conduct the operation of triple-level convolution, the target node may collect some information from its neighbors. As shown in Figure 8(b), with r increasing, the BiPNet's performance improves first because that nodes can obtain more useful information from others, and then drops due to over-smooth issues. We choose different r for different datasets, i.e., 3 for Cosmetics/Toys and 2 for Grocery, according



(a) Influence of the number of price levels ρ.

(b) Influence of the repeating number for the triple-level convolution r.

(c) Influence of the number of self-attention heads h.

Fig. 8. BiPNet's performance under different hyper-parameters

to experimental results. We speculate that there are fewer items in Grocery than Cosmetics and Toys (as shown in Table 2), which enables the proposed BiPNet to extract semantics of nodes with fewer triple-level convolution iterations in Grocery.

The number of self-attention heads h determines the granularity of partitioning subspaces when handling price sequences, where the model attends to information from different aspects. As presented in Figure 8(c), too little (e.g., 1) or too much (e.g., 32) heads can have a negative impact on BiPNet's performance. We argue that few heads are unable to capture complex price preference, while many heads make the semantics excessively dispersion, both leading to declining performance. Moreover, referring to Figure 8(c) and Table 2, we observe that dataset with long sessions (Cosmetics) benefits from a large h (16) while datasets with short sessions (Grocery and Toys) prefer a small h (8). It is in line with the fact that self-attention with large number of heads is good at capturing distant dependencies among sequences.

6.10 Complexity Analysis (RQ6)

In this part, we analysis the time complexity of the proposed BiPNet. The most time-consuming part of BiPNet is the representation learning of triple-level convolution in the proposed heterogeneous hypergraph. We can detail its time complexity as $O(n \times r\times (\bar{N_{id}} + \bar{N_p} + \bar{N_c} + \bar{N_b})$). Specifically, n is the total number of items; r is the iterative number of triple-level convolution; and $\bar{N_{id}}$, $\bar{N_p}$, $\bar{N_c}$ as well as $\bar{N_b}$ are the average number of adjacent nodes with type ID, price, category, and brand for items. Referring to Table 2, the number of item price, category, and brand is relatively small, leading to their limited contribution on increasing model complexity. For example, there are totally 5 price levels in Grocery, i.e., the $\bar{N_p}$ is no more than 5. That is, although we incorporate various kinds of information to depict user behaviors, the model is increased on a modest scale. Besides, we can also reduce the number of side features, like merging similar categories or pruning some niche brands, to improve model efficiency. We believe that such merits of BiPNet contribute to its application in real scenarios.

7 CONCLUSION AND FUTURE WORK

The user purchase behaviors are determined by not only interest preference but also price preference in real-world scenario. Unfortunately, the existing methods for session-based recommendation only focus on user interest preference. Therefore, in this work, we propose a novel approach BiPNet to capture user interest and price preference simultaneously via exploring various kinds of information for session-based recommendation. As the pioneer work considering item price in SBR, BiPNet identifies two challenges that get in the way of utilizing price information: (1) price preference is hidden behind rich heterogeneous information and (2) price and interest preference collectively determine user choice. To copy with the first challenge, in BiPNet, we construct a novel heterogeneous hypergraph that encodes rich information about user price preference including item ID, item price, item category, and item brand. Based on the heterogeneous hypergraph, a customized triple-level convolution is innovatively devised to handle heterogeneous information for interest and price preferences learning. As to the second challenge, we propose a bi-preference learning schema under the multi-task learning architecture. With exploring the mutual relations between interest preference and price preference, the bi-preference learning schema is able to model these two preferences simultaneously and accurately capture user purchase intents. We conduct extensive experiments on three real-world datasets and the results demonstrate the superiority of the proposed BiPNet over representative baselines. Further study also supports that item price is of great significance to determining user behaviors in session-based recommendation.

In the future, first, we plan to introduce more available information like item ratings, item images, and user reviews to further explore item features and capture user complex preferences. Second, it is also a promising direction that distinguishing the importance of distinct user preferences under difference context for further improving prediction accuracy. For instance, automatically assigning different weights for distinct kinds of losses based on their utility. It can endow the model with the potential to identify the most influential factors on user behaviors. Last, the proposed heterogeneous hypergraph with triple-level convolution can be easily extended to other tasks, where it is necessary to model various kinds of information with complex relations.

REFERENCES

[1] Erik Brynjolfsson and Michael D. Smith. 2000. Frictionless commerce? A comparison of Internet and conventional retailers. *Manage. Sci.* 46, 4 (2000), 563–585.

[2] Jingfan Chen, Guanghui Zhu, Haojun Hou, Chunfeng Yuan, and Yihua Huang. 2022. AutoGSR: Neural architecture search for graph-based session recommendation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'22).* 1694–1704.

[3] Qian Chen, Zhiqiang Guo, Jianjun Li, and Guohui Li. 2023. Knowledge-enhanced multi-view graph neural networks for session-based recommendation. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval.* 352–361.

[4] Tianwen Chen and Raymond Chi-Wing Wong. 2020. Handling information loss of graph neural networks for session-based recommendation. In *Proceedings of the 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'20)*. 1172–1180.

[5] Mingyue Cheng, Zhiding Liu, Qi Liu, Shenyang Ge, and Enhong Chen. 2022. Towards automatic discovering of deep hybrid network architecture for sequential recommendation. In *Proceedings of the ACM Web Conference (WWW'22)*. 1923–1932.

[6] Minjin Choi, Jinhong Kim, Joonseok Lee, Hyunjung Shim, and Jongwuk Lee. 2022. S-Walk: Accurate and scalable session-based recommendation with random walks. In *Proceedings of the 15th ACM International Conference on Web Search and Data Mining (WSDM'22)*. 150–160.

[7] Alexandru M. Degeratu, Arvind Rangaswamy, and Jianan Wu. 2000. Consumer choice behavior in online and traditional supermarkets: The effects of brand name, price, and other search attributes. *Int. J. Res. Market.* 17, 1 (2000), 55–78.

[8] Shaohua Fan, Junxiong Zhu, Xiaotian Han, Chuan Shi, Linmei Hu, Biyu Ma, and Yongliang Li. 2019. Metapath-guided heterogeneous graph neural network for intent recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD'19)*. 2478–2486.

[9] Yifan Feng, Haoxuan You, Zizhao Zhang, Rongrong Ji, and Yue Gao. 2019. Hypergraph neural networks. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI'19)*. 3558–3565.

[10] Asnat Greenstein-Messica and Lior Rokach. 2018. Personal price aware multi-seller recommender system: Evidence from eBay. *Knowl. Bas. Syst.* 150 (2018), 14–26.

[11] Junpeng Guo, Zeng Gao, Na Liu, and Yi Wu. 2018. Recommend products with consideration of multi-category inter-purchase time and price. *Fut. Gener. Comput. Syst.* 78 (2018), 451–461.

[12] Jiayan Guo, Yaming Yang, Xiangchen Song, Yuan Zhang, Yujing Wang, Jing Bai, and Yan Zhang. 2022. Learning multi-granularity consecutive user intent unit for session-based recommendation. In *Proceedings of the 15th ACM International Conference on Web Search and Data Mining (WDSM'22)*. 343–352.

[13] Jiayan Guo, Peiyan Zhang, Chaozhuo Li, Xing Xie, Yan Zhang, and Sunghun Kim. 2022. Evolutionary preference learning via graph nested GRU ODE for session-based recommendation. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. 624–634.

[14] Ruining He and Julian J. McAuley. 2016. Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. In *Proceedings of the 25th International Conference on World Wide Web*. 507–517.

[15] Balázs Hidasi and Alexandros Karatzoglou. 2018. Recurrent neural networks with top-k gains for session-based recommendations. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management, (CIKM'18)*. 843–852.

[16] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. 2016. Session-based recommendations with recurrent neural networks. In *Proceedings of the 4th International Conference on Learning Representations (ICLR'16)*.

[17] Yupeng Hou, Shanlei Mu, Wayne Xin Zhao, Yaliang Li, Bolin Ding, and Ji-Rong Wen. 2022. Towards universal sequence representation learning for recommender systems. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'22)*. 585–593.

[18] 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 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP'19)*. 4820–4829.

[19] Dietmar Jannach and Malte Ludewig. 2017. When recurrent neural networks meet the neighborhood for session-based recommendation. In *Proceedings of the 11th ACM Conference on Recommender Systems (RecSys'17)*. ACM, 306–310.

[20] Jianwen Jiang, Yuxuan Wei, Yifan Feng, Jingxuan Cao, and Yue Gao. 2019. Dynamic hypergraph neural networks. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI'19)*. 2635–2641.

[21] Toshihiro Kamishima and Shotaro Akaho. 2011. Personalized pricing recommender system: Multi-stage epsilon-greedy approach. In *Proceedings of the 2nd International Workshop on Information Heterogeneity and Fusion in Recommender Systems*. 57–64.

[22] Wang-Cheng Kang and Julian J. McAuley. 2018. Self-attentive sequential recommendation. In *Proceedings of the IEEE International Conference on Data Mining (ICDM'18)*. 197–206.

[23] Yehuda Koren. 2008. Factorization meets the neighborhood: A multifaceted collaborative filtering model. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD'08)*. 426–434.

[24] Siqi Lai, Erli Meng, Fan Zhang, Chenliang Li, Bin Wang, and Aixin Sun. 2022. An attribute-driven mirror graph network for session-based recommendation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'22)*. 1674–1683.

[25] Jing Li, Pengjie Ren, Zhumin Chen, Zhaochun Ren, Tao Lian, and Jun Ma. 2017. Neural attentive session-based recommendation. In *Proceedings of the ACM on Conference on Information and Knowledge Management (CIKM'17)*. ACM, 1419–1428.

[26] Zihao Li, Xianzhi Wang, Chao Yang, Lina Yao, Julian J. McAuley, and Guandong Xu. 2023. Exploiting explicit and implicit item relationships for session-based recommendation. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining (WSDM'23).* 553–561.

[27] Qiao Liu, Yifu Zeng, Refuoe Mokhosi, and Haibin Zhang. 2018. STAMP: Short-term attention/memory priority model for session-based recommendation. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD'18).* ACM, 1831–1839.

[28] Yuanxing Liu, Zhaochun Ren, Wei-Nan Zhang, Wanxiang Che, Ting Liu, and Dawei Yin. 2020. Keywords generation improves e-commerce session-based recommendation. In *Proceedings of the Web Conference (WWW'20).* 1604–1614.

[29] Malte Ludewig, Noemi Mauro, Sara Latifi, and Dietmar Jannach. 2020. Empirical analysis of session-based recommendation algorithms. *User Model. User-Adapt. Interact.* (2020), 1–33.

[30] Yitong Pang, Lingfei Wu, Qi Shen, Yiming Zhang, Zhihua Wei, Fangli Xu, Ethan Chang, Bo Long, and Jian Pei. 2022. Heterogeneous global graph neural networks for personalized session-based recommendation. In *Proceedings of t he 15th ACM International Conference on Web Search and Data Mining (WSDM'22).* 775–783.

[31] Ruihong Qiu, Zi Huang, Tong Chen, and Hongzhi Yin. 2022. Exploiting positional information for session-based recommendation. *ACM Trans. Inf. Syst.* 40, 2 (2022), 35:1–35:24.

[32] Ruihong Qiu, Zi Huang, Jingjing Li, and Hongzhi Yin. 2020. Exploiting cross-session information for session-based recommendation with graph neural networks. *ACM Trans. Inf. Syst.* (2020), 22:1–22:23.

[33] Pengjie Ren, Zhumin Chen, Jing Li, Zhaochun Ren, Jun Ma, and Maarten de Rijke. 2019. RepeatNet: A repeat aware neural recommendation machine for session-based recommendation. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI'19).* 4806–4813.

[34] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence, (UAI'09).* AUAI Press, 452–461.

[35] Badrul Munir Sarwar, George Karypis, Joseph A. Konstan, and John Riedl. 2001. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th International World Wide Web Conference (WWW'01).* ACM, 285–295.

[36] J. Ben Schafer, Joseph Konstan, and John Riedl. 1999. Recommender systems in e-commerce. In *Proceedings of the 1st ACM Conference on Electronic Commerce.* 158–166.

[37] Walid Shalaby, Sejoon Oh, Amir Afsharinejad, Srijan Kumar, and Xiquan Cui. 2022. M2TRec: Metadata-aware multi-task transformer for large-scale and cold-start free session-based recommendations. In *Proceedings of the 16th ACM Conference on Recommender Systems (RecSys'22).* 573–578.

[38] Guy Shani, David Heckerman, and Ronen I. Brafman. 2005. An MDP-based recommender system. *J. Mach. Learn. Res.* 6 (2005), 1265–1295.

[39] Jiajie Su, Chaochao Chen, Weiming Liu, Fei Wu, Xiaolin Zheng, and Haoming Lyu. 2023. Enhancing hierarchy-aware graph networks with deep dual clustering for session-based recommendation. In *Proceedings of the ACM Web Conference.* 165–176.

[40] Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. 2019. BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM'19).* 1441–1450.

[41] Xiangguo Sun, Hongzhi Yin, Bo Liu, Hongxu Chen, Jiuxin Cao, Yingxia Shao, and Nguyen Quoc Viet Hung. 2021. Heterogeneous hypergraph embedding for graph classification. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining (WSDM'21).* 725–733.

[42] Jiaxi Tang and Ke Wang. 2018. Personalized top-N sequential recommendation via convolutional sequence embedding. In *Proceedings of the 11th ACM International Conference on Web Search and Data Mining (WSDM'18).* 565–573.

[43] Trinh Xuan Tuan and Tu Minh Phuong. 2017. 3D convolutional networks for session-based recommendation with content features. In *Proceedings of the 11th ACM Conference on Recommender Systems (RecSys'17).* 138–146.

[44] Panniello Umberto. 2015. Developing a price-sensitive recommender system to improve accuracy and business performance of ecommerce applications. *Int. J. Electr. Commerce Stud.* 6, 1 (2015), 1–18.

[45] Mengting Wan, Di Wang, Matthew Goldman, Matt Taddy, Justin Rao, Jie Liu, Dimitrios Lymberopoulos, and Julian J. McAuley. 2017. Modeling consumer preferences and price sensitivities from large-scale grocery shopping transaction logs. In *Proceedings of the 26th International Conference on World Wide Web.* 1103–1112.

[46] Jianling Wang, Kaize Ding, Ziwei Zhu, and James Caverlee. 2021. Session-based recommendation with hypergraph attention networks. In *Proceedings of the SIAM International Conference on Data Mining (SDM'21).* 82–90.

[47] Meirui Wang, Pengjie Ren, Lei Mei, Zhumin Chen, Jun Ma, and Maarten de Rijke. 2019. A collaborative session-based recommendation approach with parallel memory modules. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'19).* 345–354.

[48] Shoujin Wang, Longbing Cao, Yan Wang, Quan Z. Sheng, Mehmet A. Orgun, and Defu Lian. 2022. A survey on session-based recommender systems. *ACM Comput. Surv.* (2022), 154:1–154:38.

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

[50] Chunyu Wei, Bing Bai, Kun Bai, and Fei Wang. 2022. GSL4Rec: Session-based recommendations with collective graph structure learning and next interaction prediction. In *Proceedings of the ACM Web Conference (WWW'22).* 2120–2130.

[51] Han Wu, Hongzhe Zhang, Liangyue Li, Zulong Chen, Fanwei Zhu, and Xiao Fang. 2022. Cheaper is better: Exploring price competitiveness for online purchase prediction. In *Proceedings of the 38th IEEE International Conference on Data Engineering (ICDE'22).* 3399–3412.

[52] Shu Wu, Yuyuan Tang, Yanqiao Zhu, Liang Wang, Xing Xie, and Tieniu Tan. 2019. Session-based recommendation with graph neural networks. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI'19).* 346–353.

[53] Xin Xia, Hongzhi Yin, Junliang Yu, Yingxia Shao, and Lizhen Cui. 2021. Self-supervised graph co-training for session-based recommendation. In *Proceedings of the 30th ACM International Conference on Information and Knowledge Management (CIKM'21).* 2180–2190.

[54] Xin Xia, Hongzhi Yin, Junliang Yu, Qinyong Wang, Lizhen Cui, and Xiangliang Zhang. 2021. Self-supervised hypergraph convolutional networks for session-based recommendation. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI'21).* 4503–4511.

[55] Chengfeng Xu, Pengpeng Zhao, Yanchi Liu, Victor S. Sheng, Jiajie Xu, Fuzhen Zhuang, Junhua Fang, and Xiaofang Zhou. 2019. Graph contextualized self-attention network for session-based recommendation. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI'19).* 3940–3946.

[56] Yuhao Yang, Chao Huang, Lianghao Xia, Chunzhen Huang, Da Luo, and Kangyi Lin. 2023. Debiased contrastive learning for sequential recommendation. In *Proceedings of the ACM Web Conference.* 1063–1073.

[57] Peiyan Zhang, Jiayan Guo, Chaozhuo Li, Yueqi Xie, Jaeboum Kim, Yan Zhang, Xing Xie, Haohan Wang, and Sunghun Kim. 2023. Efficiently leveraging multi-level user intent for session-based recommendation via atten-mixer network. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining (WSDM'23).* 168–176.

[58] Xiaokun Zhang, Hongfei Lin, Bo Xu, Chenliang Li, Yuan Lin, Haifeng Liu, and Fenlong Ma. 2022. Dynamic intent-aware iterative denoising network for session-based recommendation. *Inf. Process. Manage.* (2022), 102936.

[59] Xiaokun Zhang, Hongfei Lin, Liang Yang, Bo Xu, Yufeng Diao, and Lu Ren. 2021. Dual part-pooling attentive networks for session-based recommendation. *Neurocomputing* (2021), 89–100.

[60] Xiaokun Zhang, Bo Xu, Fenglong Ma, Chenliang Li, Liang Yang, and Hongfei Lin. 2023. Beyond co-occurrence: Multi-modal session-based recommendation. *IEEE Trans. Knowl. Data Eng.* (2023).

[61] Xiaokun Zhang, Bo Xu, Liang Yang, Chenliang Li, Fenglong Ma, Haifeng Liu, and Hongfei Lin. 2022. Price DOES matter! Modeling price and interest preferences in session-based recommendation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'22).* 1684–1693.

[62] Yu Zhang and Qiang Yang. 2021. A survey on multi-task learning. *IEEE Trans. Knowl. Data Eng.* (2021).

[63] Yu Zheng, Chen Gao, Xiangnan He, Yong Li, and Depeng Jin. 2020. Price-aware recommendation with graph convolutional networks. In *Proceedings of the 36th IEEE International Conference on Data Engineering (ICDE'20).* 133–144.

[64] Dengyong Zhou, Jiayuan Huang, and Bernhard Schölkopf. 2006. Learning with hypergraphs: Clustering, classification, and embedding. In *Proceedings of the 20th Annual Conference on Neural Information Processing Systems.* 1601–1608.

Received 5 April 2023; revised 1 August 2023; accepted 24 October 2023