

# Dynamic Multi-Interest Graph Neural Network for Session-Based Recommendation

Mingyang Lv<sup>1,3\*</sup>, Xiangfeng Liu<sup>1,3\*</sup>, Yuanbo Xu<sup>1,2†</sup>

<sup>1</sup>MIC Lab, College of Computer Science and Technology, Jilin University, China

<sup>2</sup>College of Computer Science and Technology, Jilin University, China

<sup>3</sup>College of Software, Jilin University, China

yuanbox@jlu.edu.cn, {lvmy5521, liuxf5521}@mails.jlu.edu.cn

## Abstract

Session-based recommendation (SBR) is widely used in e-commerce and streaming services, with the task of performing real-time recommendations based on short-term anonymous user history data. Most existing SBR frameworks follow the pattern of learning a single representation for a specific session, which makes it difficult to capture potential multiple interests, thus preventing discriminative recommendation. Multi-Interest learning has emerged as an effective approach for addressing this issue on sequential data in recent years. However, the current Multi-Interest frameworks perform poorly on session data because they may generate excessive interests. To address these issues, we proposed a model named **Dynamic Multi-Interest Graph Neural Network (DMI-GNN)**, which introduces the Multi-Interest learning framework into SBR and refines it by proposing a multiple positional patterns (MPP) learning method and a Dynamic Multi-Interest (DMI) regularization. Specifically, the MPP learning layer ensures the model to obtain representations with different positional information for sessions. The DMI regularization, on the other hand, mitigates the influence of excessive interests. Experiments on three bench-mark datasets demonstrate that our methods achieve better performance on different metrics.

**Code** — <https://github.com/MICLab-Rec/DMI-GNN>

## Introduction

With the rapid growth of information on the Internet, recommendation systems play a crucial role in many applications, helping users alleviate information overload and providing personalized delivery, making it easier for users to choose information that they are interested in. Many existing recommendation methods typically rely on long-term user behavior sequences, such as collaborative filtering, matrix factorization-based methods (Sarwar et al. 2001), and Markov chain-based approaches (Rendle, Freudenthaler, and Schmidt-Thieme 2010). These methods are simple and easy to implement, and have been widely applied in practical recommendation scenarios. However, in recent years, due to the limitations of real-world scenarios with unlogged

\*These authors contributed equally.

†Corresponding author



Figure 1: Scenarios of Multi-Interest in session-based recommendation.  $S_1$  and  $S_2$  are different sessions. The two  $v_5$  items are the next items of these sessions, respectively. The heatmaps in the right corner represent the similarity scores between the next items and their corresponding session items, respectively.

users and increased awareness of user privacy protection, the performance of these methods has been quite limited. As a result, session-based recommendation techniques (using anonymous behavioral data) have gained significant attention, generating recommendations for the next item based solely on the ongoing session.

Existing session-based recommendation methods are mainly based on sequence structures or graph structures. Sequence-based methods model sessions as sequences and predict the next item the user is likely to interact with. Methods based on recurrent neural networks, such as GRU and LSTM, became early leaders in this field due to their inherent ability to process sequences and their representational capabilities. However, as mentioned in (Qiu et al. 2019), methods based on recurrent neural networks cannot effectively model the relationships between items, as the actual transition patterns are far more complex than simple temporal ordering. Graph-based methods (Li et al. 2016; Wu et al. 2019; Xu et al. 2019) have been proposed in recent years to model more complex relationships within sessions. On one hand, GNNs expand the expressive power of item representations in the spatial dimension by explicitly aggregating neighborhood information of items; on the other

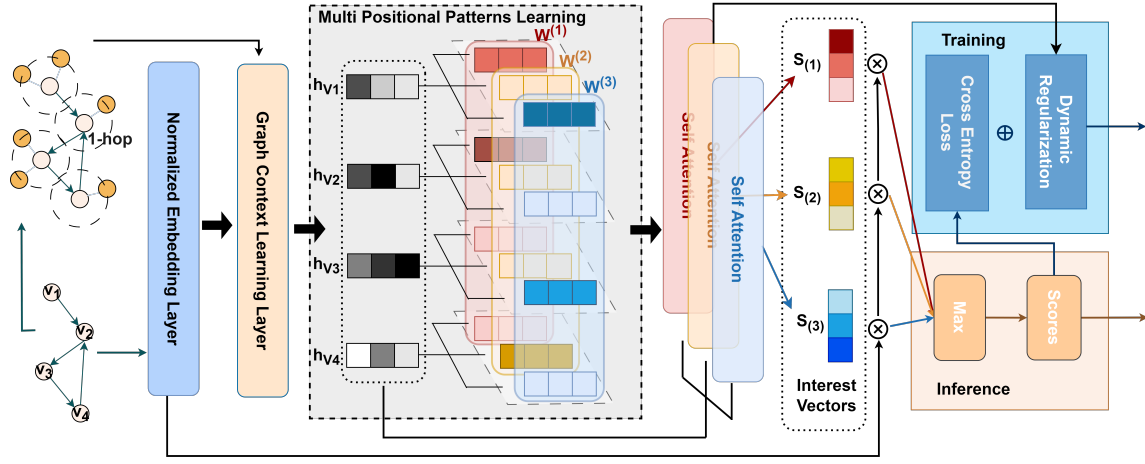


Figure 2: Framework of DMI-GNN

hand, GNNs can model complex relationships between items based on their weighted graph structure, such as cases where multiple interactions exist between two items.

Considering the limitations of session length, existing methods have adopted a strategy of generating only one specific embedding for the entire session in the final prediction stage, using this embedding to represent the current session’s interest in the next item. However, in practical recommendation scenarios, as shown in figure 1, items within one session can also be split into multiple interests. For example, although the items in  $S_1$  in the figure are all related to digital products, they can still be further subdivided into three interests: mobile phones, computers, and headphones. Moreover, figure 1 also illustrates that in actual behavior patterns, the relevance of the next item of a session to all items in that session does not follow a single positional pattern. The pattern in  $S_1$  represents targeted browsing behavior, specifically manifested as browsing some other related items after viewing the target item, but ultimately returning to the initial target item. The pattern in  $S_2$  represents non-targeted browsing behavior, which often does not lead to browsing suitable items at the beginning of the session, but is more likely to browse target items as time progresses. Unfortunately, in previous methods, the single positional pattern modeling and single interest modeling fail to capture richer behavioral patterns. This ultimately results in the interest representation either significantly biasing towards one type of item or not biasing towards any type of item; the former may lead to accurate but more monotonous results, while the latter ensures diversity in recommendations but sacrifices accuracy.

To address this issue, we introduce a multi-interest learning module to the existing session-based recommendation framework. Multi-interest learning methods (Li et al. 2019; Chai et al. 2022; Sabour, Frosst, and Hinton 2017) have been proposed in recent research and have shown high potential in many sequence-based approaches. Instead of focusing on computing a specific interest, these methods explicitly construct multiple different interest representations based on the user’s behavior sequence,

breaking through the performance bottleneck of single interest representation. Specifically, these methods often learn interest representations through the dynamic routing method of CapsuleNet (Sabour, Frosst, and Hinton 2017) or multi-head attention methods (Cen et al. 2020) after learning contextual information from item representations. Based on this, we propose a Dynamic Multi-Interest Graph Neural Network for session-based recommendation. We first use a multi-interest extractor coupled with a multiple positional patterns learning layer to obtain multiple interest representations. Then, to ensure that these interest vectors have rich representational capabilities while avoiding redundant low-quality interests representations, we propose a dynamic multi-interest regularization term to adjust the distances between interest representations, which can help adapt multi-interest framework to SBR. Finally, we adopt a strategy of taking the maximum of multiple interest scores to obtain the score for each item, and then select the top-k items as the candidate set. Our contributions in this paper can be summarized as follows:

- To the best of our knowledge, this is the first work to leverage session length information to guide the learning of interest numbers for effective session-based recommendation.
- We propose a dynamic multi-interest (DMI) regularization term to mitigate the effect of excessive interests and refine the multi-interest framework with a multiple positional patterns (MPP) learning method to obtain richer representations.
- Our proposed method, DMI-GNN, has shown superiority over the state-of-the-art baselines on three public benchmark datasets in terms of HR@20, MRR@20, and Cov@20.

## Related Works

### Graph-based SBR

Recently, due to its powerful representation learning and information aggregation capabilities, Graph Neural Networks

(GNNs) have gained widespread attention in the field of recommender systems, including Session-Based Recommendation (SBR). Among these, SR-GNN (Wu et al. 2019) was the first to apply GNNs to the SBR problem, modeling session data as a graph and using gated GNNs to capture complex transitions within sessions. SGNN-HN (Pan et al. 2020) captures complex transition relationships between items by establishing a star-shaped graph structure for sessions and utilizes a high-speed network to avoid overfitting issues. DHCN (Xia et al. 2021) introduces hypergraphs to model and capture high-order relationships between items and incorporates self-supervised learning to enhance model performance. GCE-GNN (Wang et al. 2021) captures local information within sessions and global information across all sessions through two levels of representation learning. MGIR (Han et al. 2022) employs GNNs to learn various item relationships and encodes these relationships through different aggregation layers, ultimately integrating both positive and negative relationships to generate enhanced session representations.

### Multi-Interest Learning

The increase in computational power has led recent research to trend towards using more complex and computationally expensive methods to extract richer representations from user data. Multi-Interest is one such method in recommender system applications, modeling user interests as multiple distinct vectors to capture complex interaction patterns between users and items. With the support of CapsuleNet-based Dynamic Routing technology (Sabour, Frosst, and Hinton 2017), MIND (Li et al. 2019) was the first to propose a recommender system based on multi-interest learning, using multiple capsules to represent different interests. Similarly, ComiRec (Cen et al. 2020) introduced another multi-interest extraction module based on the self-attention mechanism. There are also works like PIMIRec (Chen et al. 2021) that include temporal information and interactivity within a similar framework. Works like MGNM (Tian et al. 2022) focus more on improving the performance of multi-interest in the ranking stage. Recent studies have applied multi-interest learning to SBR. MI-GNN (Wang et al. 2023b) builds an interest graph from both historical and current behavior sequences, while TMI-GNN (Shen et al. 2022) models interests as nodes to increase item-item graph density. Unlike these methods, ours generates multiple interests from a single session.

### Proposed Method

We propose a Dynamic Multi-Interest Graph Neural Network (DMI-GNN) for session-based recommendation. Figure 2 illustrates the overall framework of DMI-GNN. First, a normalized item embedding is fed into a graph context learning layer, where the inter-session and intra-session information of each node are learned. Next, we introduce a multi-interest recommendation framework based on the multiple positional patterns learning. By designing the Dynamic Multi-Interest loss, we obtain richer interest representations replacing the traditional approach of generating only

a single session representation in SBR. Finally, the predicted probability of candidate items for recommendation is output by taking the maximum score among different interests.

### Problem Formulation

Session-based recommendation aims to predict the next item a user will interact with based on their short-term interactions. Here, we provide the formulation of the problem.

Let  $V = \{v_1, v_2, \dots, v_n\}$  be the set of items, representing all items that appear across all sessions. A session of an anonymous user can be represented as  $S = \{v_{s,1}, v_{s,2}, \dots, v_{s,l}\}$ , where  $l$  denotes the length of the session and  $v_{s,i}$  represents the  $i$ -th item in session  $s$ . The goal of a session-based recommendation system is to predict  $v_{s,l+1}$  based on a given session  $s$ . For a session-based recommendation system, we output the probability of each item  $v$  being the next interaction item for a session  $s$ , using this probability as the score. The items with the top- $K$  scores are returned as the recommendation results.

### Graph Context Learning Layer

For each item in a session, using only a single layer of embedding makes it difficult to capture and learn complex transition patterns. Due to the inherent connectivity in sessions, we use a GNN-based context learning layer to obtain rich representations for each item.

Our graph context learning layer architecture is based on GCE-GNN. It primarily provides richer representations through two aspects of information. 1). Global context of items (inter-session neighbor information). 2). Intra-session neighbor information. This can be written as follows:

$$\mathbf{h}'_v = \mathbf{h}_v^{G,k} + \mathbf{h}_v^S \quad (1)$$

Where  $\mathbf{h}_v^{G,k}$  and  $\mathbf{h}_v^S$  are the global-level and session-level graph context representations, respectively. The specific calculation methods for  $\mathbf{h}_v^{G,k}$  and  $\mathbf{h}_v^S$  are entirely based on the fundamental construction framework of the model, with reference to GCE-GNN (Wang et al. 2021).

### Multi-Interest Extractor

In the context of industrial recommendation systems, the size of the item set is often in the hundreds of thousands or even millions. Therefore, the matching of candidate items plays a crucial role in improving recommendation efficiency. The quality of representations calculated based on user behavior sequences is a decisive factor in the matching phase. In SBR, users are often anonymous, and session data replace long-term user behavioral sequences, making user interest modeling a challenging problem.

Given the limitations on session length, existing session-based recommendation models typically compute a single interest representation for each session. This results in interest embeddings that lack diversity and expressiveness. However, each session may involve items with significantly different and diverse purposes in practice, highlighting the necessity of applying Multi-Interest learning to SBR.

Various methods for extracting multiple interests currently exist, and our approach employs a multi-head self-attention mechanism. Given the graph context enhanced

item embeddings within a session:  $\mathbf{H} = [\mathbf{h}'_{v_1^s}, \mathbf{h}'_{v_2^s}, \dots, \mathbf{h}'_{v_l^s}]$  where  $l$  is the length of current session. It is notable that different items' contributions to prediction varies as we have illustrated in Figure 1. A single positional encoding cannot represent such information and what's more, it even may have negative impact on multi-interest representations, preventing them from being effective.

Our method, utilizing a multi-interest recommendation framework, learns different positional patterns for each interest. We use a learnable positional embedding matrix  $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_l]$ , where  $\mathbf{p}_i \in \mathbb{R}^d$  is a positional vector for specific position  $i$ . Item representations with positional information can be calculated as follows:

$$\mathbf{z}_i = \tanh(\mathbf{W}_3[\mathbf{h}'_{v_i^s} \parallel \mathbf{p}_{l-i+1}] + \mathbf{b}_3) \quad (2)$$

where  $\mathbf{W}_3 \in \mathbb{R}^{d \times 2d}$  and  $\mathbf{b}_3 \in \mathbb{R}^d$  are learnable parameters which act as transformation matrix and transformation vector.

Next, we use self-attention mechanism to obtain the attention weights  $\alpha_i$  of the  $i$ -th item to compute a single interest:

$$\alpha_i = \mathbf{q}_2^T \sigma(\mathbf{W}_4 \mathbf{z}_i + \mathbf{b}_4) \quad (3)$$

where  $\mathbf{W}_4 \in \mathbb{R}^{d \times d}$  and  $\mathbf{q}_2, \mathbf{b}_4 \in \mathbb{R}^d$  are learnable parameters. Because each positional pattern corresponds to a specific interest, so the parameters can't be simply stacked together as a matrix. Thus the weights vectors of multiple interests can be calculated as:

$$\alpha_i^{(u)} = \mathbf{q}_2^{(u)T} \sigma(\mathbf{W}_4^{(u)} \mathbf{z}_i + \mathbf{b}_4^{(u)}), u \in \{1, 2, \dots, U\} \quad (4)$$

where  $U$  is the number of interests, and each interest representation can be denoted as:

$$\mathbf{s}_u = \sum_{i=1}^l \alpha_i^{(u)} \mathbf{h}'_{v_i^s}, u \in \{1, 2, \dots, U\} \quad (5)$$

## Prediction Layer

Based on the interest representations, the candidate items match process relies on the probability for each candidate item, which can be represented by the similarity score of its initial embedding and the interest representation.

Given that we have multiple interest representations, and each of them can retrieve top-N items respectively, we adopt a policy to obtain the overall top-N items. This policy uses the maximum score among scores of item  $i$  with all interest vectors to denote the overall score of item  $i$ . It can be calculated as:

$$\text{score}(i) = \max_{1 \leq u \leq U} (\mathbf{s}_u^T \mathbf{h}_{v_i}) \quad (6)$$

and the corresponding probability  $\hat{y}_i$  can be denoted as :

$$\hat{y}_i = \text{softmax}(\text{score}(i)) \quad (7)$$

The loss function is the cross-entropy of the prediction results  $\hat{y}_i$ :

$$\mathcal{L}_{CE} = - \sum_{i=1}^n y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i) \quad (8)$$

where  $y$  denotes the one-hot encoding of the ground truth.

## Dynamic Regularization for Multi-Interest Learning

The multi-interest extractor is a crucial module in the multi-interest learning process. From the perspective of maximizing recommendation performance and diversity, we aim to obtain multiple interest vectors that are richer representations, which means that these multiple interests should not be too similar to each other. Therefore, based on the original loss function, we need to add a simple regularization term as follows to ensure the diversity of interest vectors:

$$\mathcal{L}_{sim} = \frac{2}{U(U-1)} \sum_{i=1}^{U-1} \sum_{j=i+1}^U \frac{|\mathbf{a}_i \mathbf{a}_j|}{|\mathbf{a}_i| |\mathbf{a}_j|} \quad (9)$$

where  $\mathbf{a}_i = [\alpha_1^{(i)}, \alpha_2^{(i)}, \dots, \alpha_L^{(i)}]$ ,  $L$  indicates the max length of the sessions, so the same as  $\mathbf{a}_j$ .

However, since the number of interests in the interest extractor often determine the structure of networks, that means it can not be revised during training. How to determine the appropriate number of interests has become one of the most challenging issues in applying Multi-Interest recommendation methods to SBR. We know that session data are often short in length, with an average of about 5-10 items, meaning that a considerable portion of session data actually only have 1-2 interests. Previous Multi-Interest recommendation frameworks struggle to handle the data where the actual number of interests is less than the preset number of interests. Excessive interests may also produce redundant, low-quality interests, thereby degrading the actual recommendation performance.

To mitigate this phenomenon, we then propose a simple yet effective solution. We discovered that these interest representations do not have to be overly discriminative; we can let the model automatically adjust the distances between interest representations. In cases where the session length is short, we want all representations to be more similar or some of them to be more similar, and when the session is long in length, the interest representations are expected to be less similar so as to capture richer representations. Very similar interest representations can be approximately considered as the same interest, as shown in Figure 3. This way, the model can automatically adjust the number or differentiation degree of Multi-Interest representations according to different datasets, thereby achieving better recommendation results. Specifically, we adjust the regularization term as follows:

$$\mathcal{L}_{DMI} = \sigma\left(\frac{2(l-\eta)}{U(U-1)} \sum_{i=1}^{U-1} \sum_{j=i+1}^U \frac{|\mathbf{a}_i \mathbf{a}_j|}{|\mathbf{a}_i| |\mathbf{a}_j|}\right) \quad (10)$$

where  $l$  is the length of current session,  $\eta$  is the hyperparameter and  $\sigma$  is the sigmoid function.

Thus the overall training loss can be denoted as :

$$\mathcal{L} = \mathcal{L}_{CE} + \beta \mathcal{L}_{DMI} \quad (11)$$

where  $\beta$  is the hyperparameter that balances two losses.

## Experiments

In this section, we report our experimental setting, including datasets, baselines, evaluation metrics, and an analysis of

experimental results. We aim to answer the following questions:

- **RQ1.** How does our proposed method performance compared with the state-of-the-art (SotA) methods in session recommendation and Multi-Interest recommendation?
- **RQ2.** How do different modules of DMI-GNN affect the recommendation?
- **RQ3.** Is the DMI regularization actually help to mitigate the problems of applying Multi-Interest learning to SBR.
- **RQ4.** How do hyperparameters influence model performance?

Dataset	# training	# test	# items	Avg.Lens
Tmall	351,268	25,898	40,727	6.69
RetailRocket	433,643	15,132	36,968	5.43
LastFM	2,837,330	672,833	38,615	11.78

Table 1: Statistics of the used datasets.

## Datasets and Preprocessing

We conducted an evaluation of the proposed method on three real-world benchmark datasets. The Tmall<sup>1</sup> dataset, sourced from the IJCAI-15 competition, comprises anonymized shopping logs from the Tmall online platform. The LastFM<sup>2</sup> dataset, captures the music listening behaviors of users. The RetailRocket<sup>3</sup> dataset, derived from a Kaggle competition, encompasses user browsing activity recorded over a six-month period. For a fair comparison, we follow the pre-processing method proposed by SR-GNN (Wu et al. 2019). The statistics of the three datasets after preprocessing are detailed in Table ??.

## Evaluation Metrics

We employ two widely used ranking based metrics as previous methods (Liu et al. 2018; Wang et al. 2021): **HR@K** (Hit Rate at K) and **MRR@K** (Mean Reciprocal Rank at K). Additionally, to measure the diversity of recommendations provided by the system, referring to previous work (Yang et al. 2023), we introduce an extra evaluation metric, **Cov@K** (Coverage at K). It quantifies how well the system can cover a wide range of items from the entire collection.

## Baselines and Implementaion Details

We conducted comparative experiments with our model against several representative categories of SBR methods.

First are the traditional methods that have been widely applied: **FPMC** (2010), **GRU4REC** (2016), **NARM** (2017).

The second category includes classic graph-based methods in the field of session recommendation: **SR-GNN** (2019) and **GCE-GNN** (2021).

The third category consists of classic methods in the multi-interest recommendation field: **MIND** (2019) and **ComiRec** (2020).

The final category includes state-of-the-art (SotA) methods that have shown the strongest performance in session recommendation in recent years: **MGIR** (2022) and **A-Mixer** (2023).

For fair comparison, we aligned our experimental settings with those of GCE-GNN. The Adam optimizer (Kingma and Ba 2015) was chosen, operating at a learning rate of 0.001. Our model was configured with an embedding size of 100 and trained within 20 epochs, processing data in batches of 100. For DMI-GNN, we tune the balance coefficient  $\beta$  among  $\{0.001, 0.005, 0.01, 0.05\}$ ,  $U$  among  $\{2, 3, 4, 5\}$ , and searched  $\eta$  from 8 to 18 in 2 increments.

Finally we select the parameter group based on the combined performance of the HR@20 and MRR@20 metrics. We conducted the experiment on a NVIDIA 3080Ti, using PyTorch version 1.11.0 + cu113.

## Overall Comparison (RQ1)

Table ?? presents the full results of nine baseline methods and our approach across three metrics on three real-world datasets, with the best result in each column highlighted in bold and the second-best result underlined. It can be observed that our method outperforms the baseline methods in every metric for each dataset. This also demonstrates the effectiveness of our proposed approach.

Among traditional methods, FPMC, based on Markov chains, performs worst as its modeling of chronological order is not complex enough. GRU4Rec, demonstrates its effectiveness on Tmall and RetailRocket by utilizing Gated Recurrent Units (GRUs) to model user sequences. By incorporating attention mechanism into SBR, NARM, improved over GRE4Rec, shows the best performance among traditional methods on all three datasets.

In contrast to traditional methods, recent approaches based on GNNs have achieved significantly better results. For instance, SR-GNN, which was the first to apply GNNs to session-based recommendation (SBR), performs well on LastFM. However, its performance advantage mainly relies on RNN-based methods. In comparison, GCE-GNN, with its more extensive utilization of graph-based inter-session and intra-session information, obtains better performance on Tmall and RetailRocket.

Unlike the traditional method of learning a single representation for a specific session, Multi-Interest learning generates multiple representations for each session to capture users’ potential interests. MIND is one of the most classic multi-interest recommendation models. Different from its excellent performance on sequential recommendations, MIND acts poorly on all three datasets of SBR because it may generate excessive interests. ComiRec utilizes self-attention mechanisms to generate multiple interest vectors and achieves impressive results on the Cov@20 metric of three datasets.

Among all baseline methods, two SotA methods, MGIR integrates various relationships to enhance the performance. A-Mixer effectively utilizes multi-level user intent informa-

<sup>1</sup><https://tianchi.aliyun.com/dataset/dataDetail?dataId=42>

<sup>2</sup><http://ocelma.net/MusicRecommendationDataset/lastfm-1K.html>

<sup>3</sup><https://www.kaggle.com/retailrocket/ecommerce-dataset>

Datasets	Tmall			RetailRocket			LastFM		
Methods	HR@20	MRR@20	Cov@20	HR@20	MRR@20	Cov@20	HR@20	MRR@20	Cov@20
FPMC	2.59	0.97	58.24	12.05	4.05	61.98	8.60	2.26	80.16
GRU4Rec	23.79	11.20	41.35	50.22	25.65	66.17	22.40	7.95	72.41
NARM	28.08	15.06	48.88	50.64	27.41	70.33	22.54	7.40	71.87
SR-GNN	26.75	13.36	53.16	50.48	26.45	69.33	22.90	9.01	57.66
GCE-GNN	32.66	15.07	68.54	55.28	28.34	71.31	24.64	8.82	66.78
MIND	28.69	14.18	47.85	43.87	23.11	59.27	18.96	7.11	39.48
ComiRec	34.22	15.64	<u>71.16</u>	53.92	28.29	<u>72.94</u>	22.23	8.08	82.64
MGIR	35.98	17.27	50.61	<u>56.52</u>	<u>29.91</u>	72.42	<u>24.72</u>	8.84	<u>84.19</u>
A-Mixer	<u>37.74</u>	<u>18.13</u>	47.42	56.28	28.76	70.34	24.50	<u>9.05</u>	67.34
DMI-GNN	<b>40.26</b>	<b>18.58</b>	<b>94.00</b>	<b>57.53</b>	<b>31.27</b>	<b>76.26</b>	<b>25.66</b>	<b>9.39</b>	<b>93.86</b>

Table 2: Model performance on all three datasets

tion to improve the inference capabilities. In comparison, MGIR achieves a comprehensive lead on RetailRocket, A-Mixer performs better on Tmall, while on LastFM, the results of both methods are similar.

Our method, DMI-GNN, combines the strengths of both approaches, outperforming graph-based and Multi-Interest-based methods. Specifically, DMI-GNN outperforms MGIR and A-Mixer across three datasets. The introduction of the Multi-Interest framework significantly enhances the diversity of recommendations, as evidenced by a substantial increase in Cov@20. Additionally, DMI-GNN maintains a high level of HR@20 and MRR@20, especially on Tmall.

Methods	Tmall		LastFM	
	HR@20	MRR@20	HR@20	MRR@20
w/o DMI reg	38.27	18.32	25.62	9.34
w/o Multi-Pos	36.93	17.84	25.63	9.34
w/o Multi-Interest	33.33	16.74	24.42	9.02
max $\rightarrow$ sum	33.70	17.10	24.32	9.01
DMI-GNN	<b>40.26</b>	<b>18.58</b>	<b>25.66</b>	<b>9.39</b>

Table 3: The ablation study of DMI-GNN on dataset Tmall and LastFM

### Ablation Study (RQ2)

We conducted an ablation study on each design choice in DMI-GNN to demonstrate their effectiveness. Specifically, these factors include the DMI regularization, the multiple positional patterns learning layer, and the entire Multi-Interest extractor. We tested a variant of the model’s inference strategy by changing the candidate selection from picking the highest score across all interest representations to using a voting method, where the scores from multiple interests are summed for each item.

Table ?? presents the ablation study of various modules in DMI-GNN. Firstly, removing the DMI regularization has

a negative impact on the recommendation system’s hit rate and ranking metrics on Tmall. This indicates that DMI regularization, by controlling the similarity of interest representations, achieves higher improvements in hit rate and ranking. This notably demonstrates the effectiveness of this regularization term. Furthermore, we can observe that replacing multiple positional patterns learning with single positional encoding results in a clear decline in metrics such as MRR@20 on Tmall, reasonably suggesting that MPP has an important influence on the order of candidate items. Finally, the results obtained using the Multi-Interest method consistently outperform those using single-interest modeling across all metrics and datasets, with substantial improvements in each metric.

The result of changing the selection strategy is similar to the effect after removing the Multi-Interest module, which might because the voting mechanism instead reduces the discriminative power and diversity of the representation, thereby degrading the model’s performance.

As mentioned earlier, multi-interest can acquire diverse user behavior information and handle more complex samples by learning richer representations.

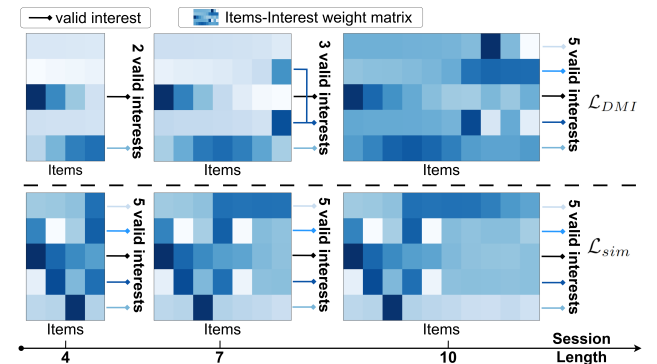


Figure 3: Comparison of the item-interest weight matrices generated by models using our proposed DMI regularization and similarity regularization respectively on sessions of different lengths.



## Case Study of DMI (RQ3)

Our model generates an item-interest weight matrix during the multi-interest extraction phase, which represents the association between items and each interest. We visualized this weight matrix in the form of a heatmap on several real cases to detect whether the DMI regularization term actually mitigates the problem of excessive interests. Figure 3 demonstrates the matrices generated by models with different regularization terms on the same real-world cases. The three included cases are sessions composed of the first 4 items, the first 7 items, and the first 9 items from the same session data, respectively. The upper part of the figure is the matrices generated by our proposed DMI regularization term. It can be observed that when the session length is 4, among the five generated interest representations that make up the matrix, three of the vectors have values almost entirely close to 0, resulting in only two valid interest representations being generated. In the matrix generated from the session of length 7, four interest representations that are valid in values can be observed, of which representations 2 and 4 are very similar. They can be considered as the same interest representation in real recommendations, thus the model ultimately generates only three different interest representations.

In the lower part of the image, where the model is merely guided by the similarity regularization term, five valid interest representations are generated for both of first two short sessions. This inevitably includes redundant interests, which may weaken their representativeness and subsequently reduce the effectiveness of the recommendations. This demonstrates that our proposed DMI regularization term effectively mitigates the redundancy issue in Multi-Interest learning for session data by controlling the differentiation degree of interest representations. For sessions of length 10, both models generate five valid interest representations to adapt to the potentially more diverse interests in longer sequences, indicating that the DMI regularization also maintains good performance in longer sequence scenarios.

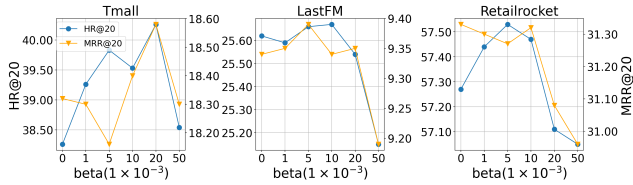


Figure 4: The impact of balance coefficient  $\beta$

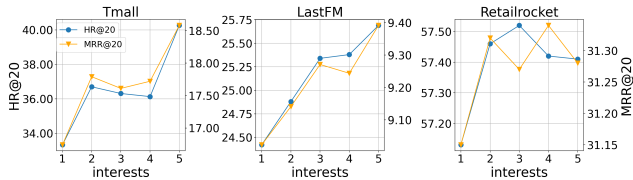


Figure 5: The impact of interests number  $U$

## Impact of hyperparameters (RQ4)

We tested the impact of different hyperparameters on model performance across three datasets as follows.

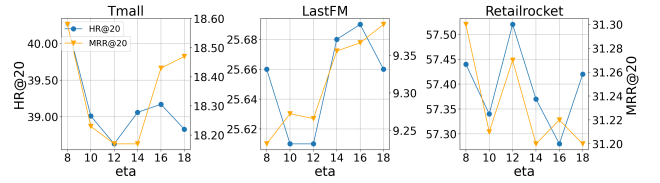


Figure 6: The impact of the hyperparameter  $\eta$

**Impact of  $\beta$ .** First, we investigated the effect of the regularization balance coefficient beta on model performance. It can be observed from Figure 4 that for both metrics across all three datasets, moderate balance coefficient values tend to achieve better results. This is because a low balance coefficient fails to leverage the regularization term to guide optimization, while an excessively high balance coefficient causes the model to "overcorrect", deviating from the initial optimization direction primarily based on cross-entropy loss, leading to decreased model performance.

**Impact of  $U$ .** Next, we also explored the impact of the number of interests on model performance. Figure 5 shows that for the three datasets we used, three or more interests tend to yield better results. The model's performance has shown a growing trend as the interests number increases. This may be due to the DMI-loss regularization term we introduced, which aims to mitigate potential performance degradation caused by excessive interests.

**Impact of  $\eta$ .** Finally, we also considered the influence of the hyperparameter  $\eta$ . The significance of  $\eta$  is to make the model focus more on distinguishing interests generated from session data with lengths near  $\eta$ . From Figure 6, it can be observed that this parameter has different effects on different datasets, we can see that the LastFM dataset, which has a longer average length, performs better at higher  $\eta$  values, while the other two datasets with shorter average lengths perform better with lower  $\eta$  values. This may be due to the different distributions of session lengths in the datasets. For example, the LastFM dataset contains more session data with lengths greater than 10, so the model performs better when it focuses more on sessions of these lengths.

## Conclusion

This paper investigates the introduction of multi-interest learning in session-based recommendation systems. This is a challenging task, as it is often difficult to extract an appropriate number of interest representations from session data, which tend to be relatively short on average. This paper proposes a recommendation system architecture based on GNNs and Multi-Interest learning: DMI-GNN. Specifically, it employs a multi-head self-attention mechanism based on multiple positional patterns learning to capture richer user interests. It then uses a DMI regularization term that adapts the distances between interest representations to control their degree of differentiation, allowing Multi-Interests learning to be better suited for SBR. Extensive experiments demonstrate that DMI-GNN outperforms nine baseline models across three benchmark datasets.

## Acknowledgments

This work is supported by the Natural Science Foundation of China No. 62472196, Jilin Science and Technology Research Project 20230101067JC.

## References

- Cen, Y.; Zhang, J.; Zou, X.; Zhou, C.; Yang, H.; and Tang, J. 2020. Controllable Multi-Interest Framework for Recommendation. In Gupta, R.; Liu, Y.; Tang, J.; and Prakash, B. A., eds., *KDD '20: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, CA, USA, August 23-27, 2020*, 2942–2951. ACM.
- Chai, Z.; Chen, Z.; Li, C.; Xiao, R.; Li, H.; Wu, J.; Chen, J.; and Tang, H. 2022. User-Aware Multi-Interest Learning for Candidate Matching in Recommenders. In Amigó, E.; Castells, P.; Gonzalo, J.; Carterette, B.; Culpepper, J. S.; and Kazai, G., eds., *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022*, 1326–1335. ACM.
- Chen, G.; Zhang, X.; Zhao, Y.; Xue, C.; and Xiang, J. 2021. Exploring Periodicity and Interactivity in Multi-Interest Framework for Sequential Recommendation. In Zhou, Z., ed., *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI 2021, Virtual Event / Montreal, Canada, 19-27 August 2021*, 1426–1433. ijcai.org.
- Han, Q.; Zhang, C.; Chen, R.; Lai, R.; Song, H.; and Li, L. 2022. Multi-Faceted Global Item Relation Learning for Session-Based Recommendation. In Amigó, E.; Castells, P.; Gonzalo, J.; Carterette, B.; Culpepper, J. S.; and Kazai, G., eds., *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022*, 1705–1715. ACM.
- Hidasi, B.; Karatzoglou, A.; Baltrunas, L.; and Tikk, D. 2016. Session-based Recommendations with Recurrent Neural Networks. In Bengio, Y.; and LeCun, Y., eds., *4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings*.
- Jiang, Y.; Xu, Y.; Yang, Y.; Yang, F.; Wang, P.; Li, C.; Zhuang, F.; and Xiong, H. 2024a. TriMLP: A Foundational MLP-Like Architecture for Sequential Recommendation. *ACM Trans. Inf. Syst.*, 42(6).
- Jiang, Y.; Yang, Y.; Xu, Y.; and Wang, E. 2024b. Spatial-Temporal Interval Aware Individual Future Trajectory Prediction. *IEEE Transactions on Knowledge and Data Engineering*, 36(10): 5374–5387.
- Kingma, D. P.; and Ba, J. 2015. Adam: A Method for Stochastic Optimization. In Bengio, Y.; and LeCun, Y., eds., *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.
- Li, C.; Liu, Z.; Wu, M.; Xu, Y.; Zhao, H.; Huang, P.; Kang, G.; Chen, Q.; Li, W.; and Lee, D. L. 2019. Multi-Interest Network with Dynamic Routing for Recommendation at Tmall. In Zhu, W.; Tao, D.; Cheng, X.; Cui, P.; Rundensteiner, E. A.; Carmel, D.; He, Q.; and Yu, J. X., eds., *Proceedings of the 28th ACM International Conference on Information and Knowledge Management, CIKM 2019, Beijing, China, November 3-7, 2019*, 2615–2623. ACM.
- Li, J.; Ren, P.; Chen, Z.; Ren, Z.; Lian, T.; and Ma, J. 2017. Neural Attentive Session-based Recommendation. In Lim, E.; Winslett, M.; Sanderson, M.; Fu, A. W.; Sun, J.; Culpepper, J. S.; Lo, E.; Ho, J. C.; Donato, D.; Agrawal, R.; Zheng, Y.; Castillo, C.; Sun, A.; Tseng, V. S.; and Li, C., eds., *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, CIKM 2017, Singapore, November 06 - 10, 2017*, 1419–1428. ACM.
- Li, Y.; Tarlow, D.; Brockschmidt, M.; and Zemel, R. S. 2016. Gated Graph Sequence Neural Networks. In Bengio, Y.; and LeCun, Y., eds., *4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings*.
- Liu, Q.; Zeng, Y.; Mokhosi, R.; and Zhang, H. 2018. STAMP: Short-Term Attention/Memory Priority Model for Session-based Recommendation. In Guo, Y.; and Farooq, F., eds., *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018, London, UK, August 19-23, 2018*, 1831–1839. ACM.
- Pan, Z.; Cai, F.; Chen, W.; Chen, H.; and de Rijke, M. 2020. Star Graph Neural Networks for Session-based Recommendation. In d'Aquin, M.; Dietze, S.; Hauff, C.; Curry, E.; and Cudré-Mauroux, P., eds., *CIKM '20: The 29th ACM International Conference on Information and Knowledge Management, Virtual Event, Ireland, October 19-23, 2020*, 1195–1204. ACM.
- Qiu, R.; Li, J.; Huang, Z.; and Yin, H. 2019. Rethinking the Item Order in Session-based Recommendation with Graph Neural Networks. In Zhu, W.; Tao, D.; Cheng, X.; Cui, P.; Rundensteiner, E. A.; Carmel, D.; He, Q.; and Yu, J. X., eds., *Proceedings of the 28th ACM International Conference on Information and Knowledge Management, CIKM 2019, Beijing, China, November 3-7, 2019*, 579–588. ACM.
- Rendle, S.; Freudenthaler, C.; and Schmidt-Thieme, L. 2010. Factorizing personalized Markov chains for next-basket recommendation. In Rappa, M.; Jones, P.; Freire, J.; and Chakrabarti, S., eds., *Proceedings of the 19th International Conference on World Wide Web, WWW 2010, Raleigh, North Carolina, USA, April 26-30, 2010*, 811–820. ACM.
- Sabour, S.; Frosst, N.; and Hinton, G. E. 2017. Dynamic Routing Between Capsules. In Guyon, I.; von Luxburg, U.; Bengio, S.; Wallach, H. M.; Fergus, R.; Vishwanathan, S. V. N.; and Garnett, R., eds., *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, 3856–3866.
- Sarwar, B. M.; Karypis, G.; Konstan, J. A.; and Riedl, J. 2001. Item-based collaborative filtering recommendation algorithms. In Shen, V. Y.; Saito, N.; Lyu, M. R.; and Zurko, M. E., eds., *Proceedings of the Tenth International World Wide Web Conference, WWW 10, Hong Kong, China, May 1-5, 2001*, 285–295. ACM.



- Shen, Q.; Zhu, S.; Pang, Y.; Zhang, Y.; and Wei, Z. 2022. Temporal aware Multi-Interest Graph Neural Network for Session-based Recommendation. In Balasubramanian, V. N.; and Tsang, I. W., eds., *Asian Conference on Machine Learning, ACML 2022, 12-14 December 2022, Hyderabad, India*, volume 189 of *Proceedings of Machine Learning Research*. PMLR.
- Tian, Y.; Chang, J.; Niu, Y.; Song, Y.; and Li, C. 2022. When Multi-Level Meets Multi-Interest: A Multi-Grained Neural Model for Sequential Recommendation. In Amigó, E.; Castells, P.; Gonzalo, J.; Carterette, B.; Culpepper, J. S.; and Kazai, G., eds., *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022*, 1632–1641. ACM.
- Wang, E.; Xu, Y.; Yang, Y.; Jiang, Y.; Yang, F.; and Wu, J. 2023a. Zone-Enhanced Spatio-Temporal Representation Learning for Urban POI Recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 35(9): 9628–9641.
- Wang, T.; Chen, C.; Huang, J.; and Huang, S. 2023b. Modeling Cross-session Information with Multi-interest Graph Neural Networks for the Next-item Recommendation. *ACM Trans. Knowl. Discov. Data*, 17(1): 1:1–1:28.
- Wang, Z.; Wei, W.; Cong, G.; Li, X.; Mao, X.; and Qiu, M. 2021. Global Context Enhanced Graph Neural Networks for Session-based Recommendation. *CoRR*, abs/2106.05081.
- Wu, S.; Tang, Y.; Zhu, Y.; Wang, L.; Xie, X.; and Tan, T. 2019. Session-Based Recommendation with Graph Neural Networks. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019*, 346–353. AAAI Press.
- Xia, X.; Yin, H.; Yu, J.; Wang, Q.; Cui, L.; and Zhang, X. 2021. Self-Supervised Hypergraph Convolutional Networks for Session-based Recommendation. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, 4503–4511. AAAI Press.
- Xu, C.; Zhao, P.; Liu, Y.; Sheng, V. S.; Xu, J.; Zhuang, F.; Fang, J.; and Zhou, X. 2019. Graph Contextualized Self-Attention Network for Session-based Recommendation. In Kraus, S., ed., *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*, 3940–3946. ijcai.org.
- Xu, Y.; Wang, E.; Yang, Y.; and Chang, Y. 2022a. A Unified Collaborative Representation Learning for Neural-Network Based Recommender Systems. *IEEE Transactions on Knowledge and Data Engineering*, 34(11): 5126–5139.
- Xu, Y.; Wang, E.; Yang, Y.; and Xiong, H. 2024a. GS-RS: A Generative Approach for Alleviating Cold Start and Filter Bubbles in Recommender Systems. *IEEE Transactions on Knowledge and Data Engineering*, 36(2): 668–681.
- Xu, Y.; Yang, Y.; Wang, E.; Zhuang, F.; and Xiong, H. 2022b. Detect Professional Malicious User With Metric Learning in Recommender Systems. *IEEE Transactions on Knowledge and Data Engineering*, 34(9): 4133–4146.
- Xu, Y.; Zhuang, F.; Wang, E.; Li, C.; and Wu, J. 2024b. Learning without Missing-At-Random Prior Propensity-A Generative Approach for Recommender Systems. *IEEE Transactions on Knowledge and Data Engineering*, 1–13.
- Yang, H.; Choi, Y.; Kim, G.; and Lee, J. 2023. LOAM: Improving Long-tail Session-based Recommendation via Niche Walk Augmentation and Tail Session Mixup. In Chen, H.; Duh, W. E.; Huang, H.; Kato, M. P.; Mothe, J.; and Poblete, B., eds., *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2023, Taipei, Taiwan, July 23-27, 2023*, 527–536. ACM.
- Zhang, P.; Guo, J.; Li, C.; Xie, Y.; Kim, J.; Zhang, Y.; Xie, X.; Wang, H.; and Kim, S. 2023. Efficiently Leveraging Multi-level User Intent for Session-based Recommendation via Atten-Mixer Network. In Chua, T.; Lauw, H. W.; Si, L.; Terzi, E.; and Tsaparas, P., eds., *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining, WSDM 2023, Singapore, 27 February 2023 - 3 March 2023*, 168–176. ACM.