
Research on the Application of Graph Neural Networks Based on Multiple Attention Mechanisms in Personalized Recommendation

Lingling Kong*, Hongyan Deng and Jiayi Huang

School of Computer Science and Technology, Hainan University, Haikou, China
E-mail: 20223003432@hainanu.edu.cn; 20223003536@hainanu.edu.cn;
20213006619@hainanu.edu.cn

**Corresponding Author*

Received 15 April 2025; Accepted 24 May 2025

Abstract

Graph neural networks (GNNs) have been widely applied due to their ability to model interactions among different objects. However, from the perspective of mathematical graph theory, the existing GNN frameworks still face challenges when dealing with specific graph structure problems. Nevertheless, the existing graph neural networks are unable to accurately identify and capture user characteristics based on common interests, have difficulty in flexibly handling diverse user interests and the differences in interests among users, and cannot effectively extract and utilize the feature information of intermediate nodes. To address these issues, this paper proposes a heterogeneous graph recommendation model based on a multi-level attention mechanism (MAHGRM-HGNN). The MAHGRM model consists of three major modules: the node-level aggregation and feature fusion module, the semantic-level aggregation module, and the importance analysis module. By introducing dual attention mechanisms at the node level and semantic level, MAHGRM can effectively identify and fuse multi-hop neighbor information

Journal of Web Engineering, Vol. 24_5, 773–804.

doi: 10.13052/jwe1540-9589.2454

© 2025 River Publishers

related to user interests, while modeling the semantic features represented by different paths. Additionally, MAHGRM adopts an innovative feature fusion method, integrating intermediate heterogeneous nodes and their related different paths according to the topological structure of the graph, thereby avoiding the loss of intermediate node information and enriching the feature representation of the target node. In the importance analysis module, MAHGRM introduces a strategy for evaluating the importance of product nodes by calculating the importance scores of different product nodes, selecting the most popular products as the candidate set, and randomly selecting some products from it for recommendation to users. MAHGRM combines the node-level aggregation and feature fusion, semantic-level aggregation, and importance analysis modules closely. The key advantage lies in its ability to effectively integrate multi-level information in heterogeneous graphs and the collaborative optimization effect brought by the cross-module feature sharing mechanism, making the final recommendation results more targeted and timely. This cross-module collaborative effect ensures the precise capture of user interests and the efficiency of product recommendations, preventing the repetition of recommending the same product to users. The experimental results were extensively tested on multiple real-world datasets. The results showed that the performance of MAHGRM was significantly superior to that of the comparison models such as GCN, GAT, HAN, HPN, OSGNN and ie-HGCN. On the MovieLens-1M dataset, the AUC, ACC and F1-score of MAHGRM reached 0.931, 0.867 and 0.863 respectively, achieving the best performance and fully demonstrating the superiority of MAHGRM in terms of performance.

Keywords: Graph neural networks, personalized recommendation systems, attention mechanism, heterogeneous graphs.

1 Introduction

At present, as a crucial application in the field of artificial intelligence, the recommendation system plays an important role in various online platforms. Its core is to accurately match user preferences with massive data, so as to resolve the problem of information overload and help users easily obtain the information they need. Traditional recommendation systems mostly rely on generalized linear models. Such models have many advantages such as simple structure, efficient operation, easy parallel processing and strong interpretability. However, their disadvantages are also significant. Over-reliance

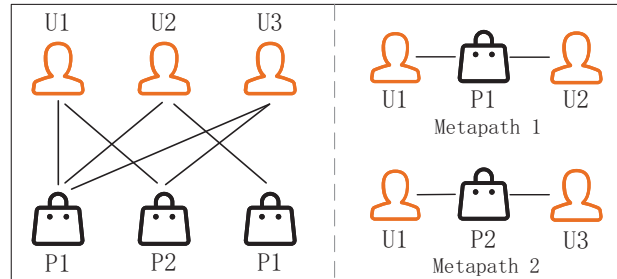


Figure 1 Schematic diagram of a heterogeneous graph and metapath.

on limited user behavior preferences and historical click sequences greatly limits the model's ability to deeply model feature interactions.

In view of this dilemma, session-based recommendation (SBR) came into being. Its characteristic is that it works only with limited and anonymous user behavior data, and predicts the user's interest-oriented items with the help of implicit user feedback. Compared with traditional recommendation methods, the performance of SBR methods is obvious when the user-item interaction data is scarce. Markov-based methods [1] regard the recommendation process as a sequence optimization problem and try to solve it to deduce the user's subsequent behavior. However, its defect is that it only focuses on the sequence conversion modeling of two consecutive items, and a large amount of historical interaction information is ignored, which seriously affects the prediction accuracy; probabilistic matrix factorization (PMF) is committed to decomposing the user-item evaluation matrix into two low-rank matrices to show the potential characteristics, but it simply defines preferences based on user positive clicks, and the results are unsatisfactory.

With the booming development of deep learning, recurrent neural networks (RNNs) have emerged in the field of SBR. Hidasi's team pioneered the introduction of RNNs with gated recurrent units (GRUs); GRU4REC [2], as the first conversational recommendation model that uses RNN networks with GRUs to model interaction sequences, has opened up a new path; Tan et al. cleverly used data augmentation and pre-training to fully consider the temporal changes in user behavior, greatly improving the robustness of training [3]; Li et al. innovatively proposed NARM, embedding the attention mechanism in the RNN architecture to simultaneously capture users' continuous behavior and interests [4]; Liu et al. took into account both global and current preferences, integrating the attention mechanism with the multi-layer perceptron (MLP) network to create STAMP [5]; Song et al.

integrated RNN with variational autoencoder (VAE) to accurately extract user preferences [6]; Wu et al. [7] innovatively proposed SR-GNN, a GNN with a gating mechanism. The model cleverly learns item embeddings by using the self-attention mechanism, calculates the correlation between each item and the last item to obtain representative session embeddings. Even so, most of these models focus on simulating the current session item transitions; Qiu et al. use FGNN with multi-weighted graph attention layer (WGAT) to achieve item representation and feature extraction; Chen et al.'s MAE-GNN combines dual-gated graph neural network with multi-head attention mechanism to lock user preferences in multiple dimensions [8]; Dong et al.'s newly proposed GPAN uses high- and low-order session perceptrons and self-attention layers to derive long- and short-term user preferences, but there is still room for improvement compared to GCE-GNN and MAE-GNN.

Although methods based on graph neural networks (GNNs) have achieved excellent performance in recommender systems, there are still some limitations. For instance, models such as HAN and SR-HGN typically employ a single graph structure or a fixed aggregation method, which makes it difficult to fully model the multi-scale evolution characteristics of user interests. They often assume that users' interests are homogeneous and stable, ignoring the fact that users may have multiple potential interests simultaneously, and these interests may change dynamically over time or in different contexts. Moreover, the attention mechanism aggregating information from all nodes will retain irrelevant noise information. The existing methods mainly face the following key challenges:

1. Difficulty in capturing features based on common interests: Direct interactions between users often indicate that users share similar interests and hobbies. For example, in Figure 1, user U1 is connected to user U2 via common hobby P1 through meta-path 1, which indicates that U1 and U2 have an intersection in hobby P1. However, in complex relationship interactions, it is difficult to identify and mine common interests in user groups.
2. Diverse information in the graph handles complex problems: Each user's interests and hobbies are usually diverse, and they may share multiple different hobbies with different user groups. As shown in Figure 1, U1 is not only connected to U2 through P1, but also connected to U3 through P2, which reflects the diversity of user interests. How to capture this multivariate relationship is a key challenge for HGNN.
3. The problem of fusion of heterogeneous node information. When constructing an interactive network, the intermediate node commodities

between users also play an important role, but they are often overlooked in the modeling process. For example, in Figure 1, U1 is connected to U2 through meta-path 1, but the intermediate heterogeneous node P1 is ignored in this process. As the direct neighbor of U1, the intermediate node P1 contains important information about user interests. Therefore, effectively capturing the information of intermediate nodes is an important problem to be solved.

4. The problem of popular products being ignored: Users often have the habit of chasing popular products. Therefore, how to effectively incorporate popular products that have no interaction with the user's current interests into the user's field of vision is a problem worth considering.

Based on the above analysis, this paper proposes a heterogeneous graph recommendation model named MAHGRM-HGNN. MAHGRM-HGNN consists of a node-level aggregation and feature fusion module, a semantic-level aggregation module, an importance analysis module, and a fusion recommendation strategy. Firstly, the node-level aggregation and feature fusion module fuses the features of the target node and its similar preference neighbors through the node-level attention mechanism to generate precise interest feature embeddings. By combining the meta-path topological structure, the feature fusion of intermediate heterogeneous nodes is carried out, effectively retaining the information of intermediate nodes and enhancing the feature richness. This mechanism is different from the existing methods that directly ignore or simply process the intermediate nodes, enabling the model to more fully explore the potential relationships in the heterogeneous graph and ensuring that the information of intermediate nodes is not lost, enriching the feature representation. Then, in the semantic-level aggregation module, for different meta-paths shared by the target node and different neighbor nodes with various hobbies, semantic-level attention is used for fusion, and multi-semantic fusion feature embeddings are generated for different semantic features. In the importance analysis module, MAHGRM uses graph importance to calculate the importance score of nodes in the graph, selects the Top-K popular products to form a candidate set, and enhances the diversity of user relationships through random edge connections, enriching the connection relationships of users. This method is different from the simple extreme popular strategy, ensuring that the recommended content contains both popular elements and diverse potential interests, enhancing the richness and novelty of the recommendation. Additionally, the random edge connection technology avoids bias towards a single hot product while improving the robustness of the model. Finally, through the fusion recommendation strategy,

personalized recommendations and popular product nodes are integrated. Firstly, personalized recommendation lists are generated based on node-level and semantic-level aggregation, and from the popular products selected by the importance analysis, several products are randomly selected to form a subset. The personalized recommendation list (from feature aggregation) and the popular product subset (from importance analysis) are combined, and the final recommendation list is generated according to a dynamic ratio, balancing personalization and popular trends. This strategy ensures that the recommended content not only conforms to the user's interests but also contains popular trends. This method effectively improves the personalization and diversity of the recommendation.

In summary, the model has made the following key contributions:

1. This paper proposes a new multi-level attention heterogeneous graph recommendation model (MAHGRM), which captures the features between different nodes and different meta-paths through node-level attention and semantic-level attention, and fuses them to generate feature embedding.
2. This paper uses the feature fusion method to fuse the intermediate heterogeneous nodes in the heterogeneous graph with the different meta-paths containing them according to the topological structure of the graph, avoiding the problem of intermediate node loss.
3. This paper calculates the importance scores of different product nodes and selects popular products as candidate sets for random recommendations to users, which helps to increase the novelty of recommended content and prevent duplicate products from being recommended to users.
4. This paper calculates the importance scores of different product nodes and selects popular products as candidate sets for random recommendations to users, which helps to increase the novelty of recommended content and prevent duplicate products from being recommended to users.

2 Related Work

2.1 Recommendation Methods Based on Deep Learning

Applying graph neural networks to recommendation systems can more effectively explore users' interests and needs. Therefore, this direction has become an important focus in current recommendation system research and has attracted the attention of scholars in various fields.

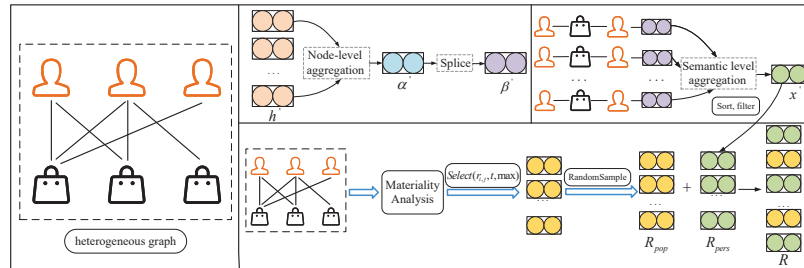


Figure 2 MAHGRM model flowchart.

Du et al. proposed an image recommendation algorithm that combines time sorting and feature vector construction, processed by an LSTM neural network, which is significantly superior to traditional methods [9]. Fang et al. integrated deep neural networks with collaborative filtering, enhanced feature extraction through short-term memory networks, and effectively reduced the root mean square error and average absolute error, reducing them by 2.015% and 2.222% respectively [10].

Guo and Wang proposed a social recommendation framework that aims to solve the problem of insufficient consideration of item feature correlation in IoT user recommendation systems. The framework considers the correlation between IoT user features by abstractly encoding user and item feature spaces, thereby providing more personalized information services [11]. The hierarchical attention recommendation system designed by Zhao et al. uses user behavior information in social networks and integrates heterogeneous human network data to improve the effectiveness of graph neural network recommendation algorithms in dealing with heterogeneity. The research results show that this system can significantly improve the flexibility of the recommendation system [12].

Zhu et al. developed a neural attention-based tourism package recommendation system that combines users' long-term and short-term behaviors, aiming to enhance the effectiveness of matching users' preferences. This system integrates the tourism package encoder and the user encoder to fuse preferences, thereby obtaining more tailored recommendations [13]. Huang et al. proposed an interactive recommendation system optimized by deep reinforcement learning, which outperforms traditional methods in improving the hit rate of long-term recommendations [14]. Zhu developed a machine learning-based system that uses gradient descent and backpropagation for training, improving the accuracy of human resource recommendations significantly [15]. Yuan proposed a network educational

resource recommendation model based on path sorting, optimizing link prediction and sorting, effectively enhancing the recommendation and sharing of online resources [16].

The above studies show that the graph neural network algorithm combining attention mechanism and short-term preference can effectively enhance the performance of the recommendation system. Considering that relevant research in this field is still insufficient, this study will focus on integrating attention mechanism and short-term preference to improve the effect of recommendation algorithm through feature extraction.

2.2 GNN-based Methods

Recently, graph neural networks (GNNs) [17, 18] have shown good performance in sequential recommender system (SBR) models, opening up new research directions in this field. Most GNN-related methods represent conversation sequences as conversation graphs and aggregate information from adjacent nodes through GNNs.

Taking SR-GNN as an example, by combining RNN, GGNN and attention mechanism, the user-item sequence is constructed as a graph to capture the potential transitions of items in the conversation. Subsequently, Xu et al. proposed GC-SAN, which utilizes GNNs and a self-attention mechanism to capture the long-term dependencies between items [19]. Later, TAGNN optimized the attention module and considered the historical behavior correlation of the target item. Wang et al. proposed GCE-GNN, which learns the transitions of items in the current and historical conversations by expanding the information range and performs excellently. Recently, Dong et al. proposed GPAN, which models the item transitions in high and low-level conversations using high-order conversation perceptrons and combines session location information to strengthen the sequence-preference relationship [20].

3 Proposed Model

This chapter will introduce the four main components of the MAHGRM model in detail, namely: a node-level aggregation and feature fusion module, a semantic-level aggregation module, an importance analysis module and a fusion recommendation strategy. The overall process of the MAHGRM model is shown in Figure 2.

In heterogeneous graphs, different types of nodes have significant heterogeneity, which means that they differ in structure, function and feature

representation. Therefore, using only the original features of the nodes may not be enough to capture the relationship and similarity between them. To solve this problem, this paper defines a corresponding feature transformation matrix for each type of node in the heterogeneous graph, mapping the features of each type of node into a unified feature space. Specifically, different types of node features correspond to different feature transformation matrices to ensure that the features of different categories of nodes are fully reflected in the new feature space. For example, in the recommendation system of a social network, some nodes may represent users, while others represent products. By projecting the features of user and product nodes into the same feature space, the user's potential preference for products can be more accurately identified. The specific feature transformation process is shown in formula (1):

$$h'_p = h_p M_p \quad (1)$$

Let $h'_p \in \mathbb{R}^{n \times m'}$ be the feature matrix of type p nodes after feature transformation, and $h_p \in \mathbb{R}^{n \times m}$ be the original feature matrix of type p nodes, where n is the number of type p nodes. The transformation matrix $M_p \in \mathbb{R}^{m \times m'}$ is used for the dimensional transformation specific to type p nodes, where m is the original embedding dimension of type p nodes and m' is the feature dimension after projection. After the above operations, all types of nodes are projected into the same dimensional space.

3.1 Node-level Aggregation and Feature Fusion Module

In a heterogeneous graph, users with common interests and hobbies form a pair of meta-path-based node pairs. For these node pairs, the MAHGRM model adopts a node-level attention mechanism, which aims to effectively fuse the features of neighbor users n_j with the features of the target node n_i . This processing method can capture the interest characteristics of individual users and enhance the accuracy and personalization of recommendations by using the information of similar users around them. Specifically, the node-level attention mechanism measures their influence on the target node by assigning different weights to each neighbor node. Some neighbor users may be given higher weights because they are highly similar to the target node in interests, thereby more significantly affecting the feature representation of the target node during feature fusion. In this way, MAHGRM can more effectively identify and utilize common interests in the user group and generate more representative feature embeddings for the target node. However, feature fusion that relies solely on neighbor users may also lead to the

loss of the target node's own features. Therefore, MAHGRM introduces a residual connection strategy. This can preserve the original features of the target node n_i during the feature fusion process. Specifically, when fusion of features occurs, the model not only combines the features of neighboring users with the features of the target node, but also directly adds the original features of the target node to the fused feature representation. The specific implementation process is as follows:

$$\delta_{i,j} = \frac{\exp(\sigma(W \cdot [h'_i \| h'_j]))}{\sum_{k \in K_i} \exp(\sigma(W \cdot [h'_i \| h'_k]))} \quad (2)$$

$$\alpha'_i = \sigma \left(\sum_{j \in K_i} \delta_{i,j} \cdot h'_j \right) + h'_i \quad (3)$$

where σ is the activation function \tanh , $\|$ represents the concatenation operation, W is the node-level attention vector used to compute the attention value $\delta_{(i,j)}$, h'_j is the feature of node $n_j \in N_i$, and α'_i is the new feature embedding generated by node n_j that incorporates neighbor features.

In heterogeneous graphs, the features of intermediate nodes contain important information about themselves and also reflect the relationship with other nodes. If the features of these intermediate nodes are ignored during the modeling process, it may lead to inaccurate judgment of user preferences, thereby reducing the accuracy of the recommendation system. In order to solve the problem of intermediate node feature loss, MAHGRM analyzes the topological structure of the heterogeneous graph, identifies the connection relationship between the intermediate heterogeneous nodes and the meta-paths containing them, and effectively fuses the intermediate heterogeneous nodes in the heterogeneous graph with the target node features under different meta-paths related to them according to the topological structure of the graph. Considering that the number of intermediate nodes between the target node n_i and the neighbor node n_j under the path ϕ_i may be more than one, it is assumed here that there are k intermediate nodes between the target node and its neighbors. The specific implementation process is shown in formula (4):

$$\beta'_i = \alpha'_i \| h_1 \| h_2 \| \dots \| h_k \quad (4)$$

where α'_i is the feature embedding of node n_i under path ϕ_i , and β'_i combines the features of intermediate nodes. h_1, h_2, \dots, h_k are the features of target nodes n_i in path ϕ_i between itself and neighboring nodes n_j .

This fusion not only enriches the feature representation of the intermediate nodes, but also improves the modeling ability of user preferences by integrating information from different interaction modes. In addition, this topology-based fusion strategy helps avoid information redundancy and loss, so that the features of each node can be retained and optimized.

3.2 Semantic-level Aggregation

For the various meta-paths connecting the target user node n_i and other different neighboring users, the MAHGRM model introduces a semantic-level attention mechanism to aggregate different features, thereby improving the understanding of user interests and the accuracy of recommendations. In a heterogeneous graph, the target user node and its neighboring users are often connected by multiple meta-paths, each of which represents the relationship and interaction pattern of different types of interests and hobbies between users. However, traditional methods cannot effectively identify and utilize these complex structural information. To address this challenge, MAHGRM aggregates different features through a semantic-level attention mechanism. This strategy first assigns a weight to each meta-path, reflecting the importance of the path in describing the target user's interests. The calculation of this weight is based on the interaction intensity and similarity between the target user node and its neighboring users, ensuring that the model can focus on the paths that are most likely to affect user preferences. Specifically, MAHGRM evaluates the features of neighboring users directly connected to the target user n_i , and weighted sums these features through a semantic-level attention mechanism. In this way, the model can not only integrate information from different neighboring users, but also generate a comprehensive user feature embedding. The specific implementation process is as follows:

$$\eta_{\phi_i} = \frac{1}{|\mathcal{V}|} \sum_{i \in \mathcal{V}} t^T \cdot \tanh(Q \cdot \beta'_i + b) \quad (5)$$

$$\psi_{\phi_i} = \frac{\exp(\eta_{\phi_i})}{\sum_{\phi_i \in \mathcal{X}^R} \exp(\eta_{\phi_i})} \quad (6)$$

$$x_i = \sum_{i=1}^p \psi_{\phi_i} \cdot \beta'_i \quad (7)$$

where $Q \in \mathbb{R}^{d \times d}$ is a learnable weight matrix, $b \in \mathbb{R}^{d \times 1}$ is the bias vector, and $t^T \in \mathbb{R}^{1 \times d}$ is the multi-head attention vector. β'_i is the feature vector of

node n_i after certain transformations, and η_{ϕ_i} is the importance weight of the corresponding path. The attention value ψ_{ϕ_i} is obtained using the softmax function, and x_i is computed as a weighted sum of features incorporating the attention value.

MAHGRM effectively aggregates various meta-path features between the target user node and its neighbor users by adopting a semantic-level attention mechanism, enhancing the comprehensive capture of user interests and providing more accurate recommendations for the recommendation system.

3.3 Importance Analysis

In order to make the model focus on the most useful information and effectively ignore the impact of noise on the features, MAHGRM-HGNN implements a node denoising strategy in the attention aggregation process, aiming to improve the robustness and accuracy of the model and ensure that the recommendation system can extract the most relevant features from complex heterogeneous graph data. First, the attention value of the node is calculated by formula (8):

$$\gamma_{i,j} = \frac{\exp\left(\sigma(W \cdot [h'_i \| h'_j])\right)}{\sum_{k \in \mathcal{K}_i} \exp\left(\sigma(W \cdot [h'_i \| h'_k])\right)} \quad (8)$$

where σ is the activation function \tanh , $\|$ represents the concatenation operation, W is the node-level attention vector, and h'_j is the feature of node $n_j \in N_i$. $\gamma_{i,j}$ is the attention value computed for the node.

In the MAHGRM-HGNN model, the attention values for nodes are evaluated. When the attention value is below a parameter t , the model automatically sets it to zero. Specifically, when a node has a low impact on the target user, its weight in the feature aggregation process is diminished, effectively removing this noise information. This denoising mechanism helps simplify feature representation and increases the model's sensitivity to important information. By filtering out nodes that are irrelevant to the target user or have significant interference effects, MAHGRM-HGNN can focus on neighboring nodes that significantly contribute to the recommendation results. This ensures that the final generated user feature embedding is more accurate and better aligns with the user's true interests. Additionally, the denoising method based on attention values is dynamic, allowing the model to adjust the importance of nodes according to changing contexts and user interactions. The specific implementation process is shown in

Equation (9).

$$S_{i,j} = \text{Select}(\gamma_{i,j}, t, \text{max}) \quad (9)$$

where *Select* is a selection function that selects based on three parameters: $\gamma_{i,j}$, t is the number of nodes selected, and *max* is the selection criterion. First, sort all the nodes in $\gamma_{i,j}$ by their values based on the attention values of the nodes. Then, based on the preset threshold t and the maximum number of selected nodes *max*, the first t nodes are selected from the sorting list to ensure that only the most important nodes are focused on, thereby improving the efficiency and effect of the model. The value of t can be set as the median of the total number of nodes or the 75th percentile, thereby ensuring that only significant nodes are selected for subsequent processing. When t is large, more nodes will be filtered out, which may introduce noise but is helpful for capturing details. The smaller t value is more conservative, emphasizing the core node and reducing noise.

In order to effectively exclude irrelevant node information, improve the efficiency and accuracy of overall learning, and ensure that the node features selected by the model are fully paid attention to, this paper sets the attention value of unselected nodes to 0:

$$\gamma_{i,j} - s_{i,j} \rightarrow 0 \quad (10)$$

where $\gamma_{i,j}$ represents the attention values for all product nodes, and $s_{i,j}$ is the attention value after selecting the top t nodes.

Through the above operations, t popular nodes can be screened out and randomly recommended to users. This process ensures that users receive the latest and most relevant content. The randomness of the recommendation increases diversity, allowing users to discover new information every time they browse. When implementing, you can also consider adding personalized recommendation tags to each popular node based on the user's historical preferences and interests to further enhance the accuracy and attractiveness of the recommendation.

3.4 Fusion Recommendation Strategy

In the MAHGRM-HGNN model, through formulas (5)–(7) of the node-level aggregation and feature fusion module and the semantic-level aggregation module, the feature embedding x_i that can accurately reflect the user's interest preferences was generated. These feature embeddings capture the common interests between the user and their neighboring nodes, and also integrate

the rich information from different meta-paths, generating a comprehensive interest representation for each user. Meanwhile, in the importance analysis module, the MAHGRM-HGNN model evaluated the importance of commodity nodes based on the attention values of the nodes, and selected the top t most popular commodities as the candidate set $s_{i,j}$. This process ensures that the nodes that interact frequently with the majority of users are given priority in the final recommendation.

In order to effectively integrate the features generated by the node-level aggregation and feature fusion module and the semantic-level aggregation module with the t popular product node candidate sets $s_{i,j}$ generated by the importance analysis module and recommend them to users, this paper proposes a comprehensive strategy. First, using the feature embedding x_i obtained from the node-level aggregation and feature fusion module, a preliminary recommendation list is generated. This list is constructed based on the personalized preferences of users and can precisely match different users' interests. Then, we add the t popular product node candidate sets $s_{i,j}$ selected by the importance analysis module to this preliminary recommendation list to form a more diverse recommendation set. Specifically, the integration strategy can be implemented through the following steps:

- Generating user preference recommendations: Based on the aggregated embedding features α_i calculated from the module and the hierarchical fusion model, use the TopK function to determine the top N_{pers} items that best match the target user's interest, generating an ordered list of user preference recommendations R_{pers} :

$$R_{\text{pers}} = \text{TopK}(x_i, N_{\text{pers}}) \quad (11)$$

Here, x_i represents the feature embedding of the node, and the TopK function indicates the selection of the top N_{pers} items that best match the interests of the target user.

- Adding hot item points: From the importance sampling model module, randomly select N_{pop} hot items from the available items, and add them to the generated user preference recommendation list R_{pers} . This ensures the diversity of recommendation results and enhances the variety of recommended content, allowing users to encounter more trending items,

$$R_{\text{pop}} = \text{RandomSample}(S_{i,j}, N_{\text{pop}}) \quad (12)$$

in which the RandomSample function selects a set of N_{pop} hot items from the available set $S_{i,j}$.

- Recommendation list adjustment: In order to ensure the quality and diversity of the final recommendation list, the proportions of personalized recommendations to hot items need to be adjusted. Let α represent the proportion of personalized recommendations and β represent the proportion of hot items, where $\alpha + \beta = 1$. The total recommendation quantity N can be represented as:

$$N_{\text{pers}} = \alpha \cdot N \quad (12)$$

$$N_{\text{pop}} = \beta \cdot N \quad (13)$$

where R represents the final recommendation list, which can be constructed by merging the personalized recommendation list R_{pers} and the hot item points R_{pop} :

$$R = \text{Merge}(R_{\text{pers}}, R_{\text{pop}}) \quad (14)$$

Here, the function Merge combines the personalized recommendation list R_{pers} and the hot item points R_{pop} to form a new set R , while also randomly increasing diversity.

Through this fusion strategy, the MAHGRM-HGNN model captures the complex interests of users and also combines the hot trends in the market, providing users with an accurate and rich recommendation experience. By introducing globally popular products, a reliable recommendation basis can be provided for new users or users with sparse behaviors. As shown in the hyperparameter experiment in Section 4.7, when $\alpha = 0.7$, the model achieves the optimal AUC value, indicating that retaining 30% of the recommended proportion of popular products can balance cold start and personalized demands. This not only ensures that the recommendation system can deeply understand the personal preferences of users, but also flexibly respond to changes in the external environment, thereby enhancing the overall satisfaction of the user experience.

3.5 Model Training

In order to ensure the effectiveness of the model during the training process, it is necessary to reasonably balance the ratio of positive samples and negative samples. Therefore, we extract the same number of negative samples for each user as the number of their positive samples. Specifically, the ratio of positive samples to negative samples for each user in the training set is balanced, so that the model receives balanced signals from positive and negative samples

during learning, avoiding the degradation of model performance due to too many negative samples. This paper defines the loss function of the model, as shown in formula (15) and formula (16).

$$\Gamma(y_{ij}, y_{ij}) = - \sum_{i=1}^p [y_{ij} \log(y_{ij}) + (1 - y_{ij}) \log(1 - y_{ij})] \quad (15)$$

$$L = \sum_i \left(\sum_{j \in \mathcal{I}(i)} \Gamma(y_{ij}, y_{ij}) - \sum_{j \in \mathcal{J}(i)} \Gamma(y_{ij}, y_{ij}) \right) + \lambda \|\Delta\|_2^2 \quad (16)$$

In these equations, p is the set of nodes with labels, y_{ij} is the actual input of node j , and y_{ij} represents the output from node j . The term Γ represents the cross-entropy loss function, Δ is the set of parameters that need to be trained in the model, and P^+ refers to the set of positive items for the user, while P^- represents the set of negative items.

This loss function is designed to quantify the performance of the model in actual recommendations and optimize the model's parameters by minimizing the loss value. This setting ensures that positive and negative samples are given equal attention during training. By extracting an equal amount of negative samples and designing a reasonable loss function, MAHGRM-HGNN can effectively adjust its weights during training, enhancing the model's learning efficiency and recommendation effect.

4 Experiments

4.1 Dataset

This experiment selected three public benchmark datasets from different fields to verify the effectiveness of the proposed model. The following is a brief introduction to the datasets:

- **MovieLens-1M dataset:** Released by the GroupLens Research team in 2003, dedicated to movie recommendation research. The dataset contains basic information of users (such as ID, age, gender, occupation) and movie information (such as ID, title, type, and release year). User ratings are integers from 1 to 5, reflecting the degree of preference for the movie. In addition, the dataset also contains user labels for movies (such as comedy, action, suspense, etc.). This dataset can be used for research such as recommendation systems, user behavior analysis, and movie classification.

- **AL-MAHGRM dataset:** This dataset consists of users (U) and music (M). The number of user nodes is 75,707 and the number of music nodes is 9944. Music is the target type node and is divided into six categories. There are four edge relationships between users and music: download, play, and favorite. There are three edge relationships between users: same hobby, same age, and same gender. In addition, there are two kinds of edge relationships between music, namely the same singer and the same style. For AL-MAHGRM, the meta-relationship paths set in this paper are: P-M-P: download, P-M-P: play, P-M-P: favorite. This information can be used to analyze the correlation between users' music preferences and artists.
- **Book-Crossing dataset:** Published by Cai-Nicolas Ziegler et al. in 2004, the dataset contains basic information of readers (such as ID, age, gender) and book information (such as ISBN, title, author, publisher). Users' ratings of books are integers from 1 to 10, indicating the degree of preference. This dataset can be used to analyze the correlation between reading preferences and books.

According to the statistical information of the experimental dataset in Table 1, the movie ratings (five-point system) in the MovieLens-1M dataset are converted into binary form: when the audience's rating of the movie is equal to or greater than 4, the user's interaction with the movie is marked as a positive example; if the rating is less than 4, it is marked as a negative example. For the LastFM dataset, if the music appears in the listener's listening record, the interaction between such a user and the music is considered a positive example, otherwise it is a negative example. In the Book-Crossing dataset, as long as the user has interacted with the item (that is, the item appears in the user's interaction list), it is marked as a positive example, and if it has not appeared, it is a negative example.

Specifically, for each user, the label of the positive interaction sample is set to 1, and the number of negative samples equal to the positive sample is randomly selected and their labels are set to 0. Subsequently, the total dataset

Table 1 Dataset statistics

Dataset	MovieLens-1M	AL-MAHGRM	Book-Crossing
Users	6036	75,707	17,860
Projects	2445	9944	14,967
Interactions	753,772	632,841	139,746
Relationships	6	6	12

consisting of positive and negative interaction samples is divided into training set, validation set and test set in a ratio of 6:2:2. Among them, the training set is used for model training, the validation set is used for hyperparameter tuning, and the test set is used to evaluate the final performance of the model.

4.2 Baselines and Experimental Setup

To verify the effectiveness of the proposed model, we compare it with the following existing models:

- GCN [21]: This is a classic graph neural network model that aggregates the features of neighboring nodes through convolution operations to generate the representation of each node.
- GAT [22]: This is a graph neural network model based on the attention mechanism that can dynamically adjust the importance of neighboring nodes to enhance the expressiveness of the model.
- HAN [23]: This is a heterogeneous graph neural network model that uses a hierarchical attention mechanism to capture nodes and their semantic information in heterogeneous graphs.
- ie-HGCN [24]: This is an improved heterogeneous graph convolutional network that effectively integrates information of different types of nodes and edges by introducing enhanced graph convolution operations.
- HPN [25]: This is a propagation network for heterogeneous graphs that uses the semantic information of multiple nodes to improve the effectiveness of node representation through a graph propagation mechanism.
- OSGNN [26]: This is an optimized semantic graph neural network that improves the contextual representation of the graph by selecting the optimal semantic features and propagation methods.
- GCORec [27]: This is a recommendation model that integrates short-term and long-term preferences, and uses a hierarchical attention network and group enhancement mechanism to evaluate project attributes.

In the experiment, we set the learning rate of all models to 0.004, the hidden layer size to 64, and the training process lasted for 1000 rounds. To reduce the consumption of computing resources, we adopted the early stopping strategy during the training process and set the patience value to 5. In terms of specific configurations, GAT has set up eight attention heads. HAN adjusted the number and type of meta-paths according to the dataset, while GCORec combined short-term and long-term preferences and set corresponding network parameters respectively. Through these detailed

parameter settings, fair comparisons among the models were ensured, thereby accurately evaluating the advantages of MAHGRM over existing methods.

The experimental results use three indicators: Area under curve (AUC), accuracy (ACC) and F1 measure (F1) to evaluate the performance of different variant models.

- **AUC:** AUC measures the performance of the model in the classification task, which represents the area under the ROC curve and reflects the ability of the model to distinguish between positive and negative samples. The closer the value is to 1, the better the model performance.
- **ACC:** Accuracy refers to the proportion of correctly classified samples to the total number of samples, which is used to measure the overall prediction ability of the model. A high accuracy means that the model makes correct predictions in most cases.
- **F1:** The F1 score combines precision and recall, and is suitable for the evaluation of imbalanced datasets. It takes into account the accuracy and detection rate of the model and is a more comprehensive performance indicator.

4.3 Comparison Experiment with Baselines

In order to evaluate the overall effectiveness of the MAHGRM model, this paper compares it with several representative benchmark models. Table 2 shows the performance results of the model on the three datasets of MovieLens-1M, AL-MAHGRM, and Book-Crossing, where the best performance is marked in bold. In this way, we can clearly see the advantages and disadvantages of the MAHGRM model under different evaluation indicators.

Table 2 Comparison of experimental results of different baseline models

DataSet	Index	GCN	GAT	HAN	HPN	OSGNN	ie-HGCN	GCORec	MAHGRM
MovieLens-1M	AUC	0.825	0.819	0.907	0.915	0.924	0.928	0.925	0.931
	ACC	0.817	0.823	0.829	0.831	0.842	0.853	0.855	0.867
	F1	0.809	0.828	0.835	0.841	0.852	0.859	0.862	0.863
AL-MAHGRM	AUC	0.782	0.789	0.793	0.801	0.824	0.843	0.848	0.852
	ACC	0.725	0.731	0.739	0.745	0.752	0.764	0.767	0.773
	F1	0.731	0.739	0.742	0.748	0.753	0.761	0.764	0.768
Book-Crossing	AUC	0.702	0.706	0.718	0.724	0.732	0.745	0.748	0.752
	ACC	0.658	0.662	0.681	0.695	0.703	0.712	0.720	0.723
	F1	0.658	0.662	0.668	0.672	0.675	0.679	0.677	0.682

It can be seen from the data in the table that the performance of the MAHGRM model on different datasets (including MovieLens-1M, AL-MAHGRM and Book-Crossing) is generally better than that of other comparison models, such as GCN, GAT, HAN, HPN, OSGNN and ie-HGCN. Specifically, on the MovieLens-1M dataset, the AUC, ACC and F1-score of MAHGRM reached 0.931, 0.867 and 0.863 respectively, all of which were the highest values. Similarly, in the AL-MAHGRM and Book-Crossing datasets, MAHGRM also demonstrated relatively high performance metrics, although these improvements varied among different metrics and datasets. This indicates that MAHGRM has significant advantages in handling complex relationships and diverse interests in recommendation systems. HAN relies on predefined meta-paths for semantic fusion, resulting in the attention mechanism taking effect only within the path. However, MAHGRM dynamically fuses various different meta-paths through semantic-level cross-path attention. Traditional homogeneous graph networks ignore the characteristics of heterogeneous nodes and lose the differences in weights in behaviors. Although GAT introduces node attention, it fails to distinguish the semantics of meta-paths, resulting in the failure of feature fusion. Furthermore, although GCORec and ie-HGCN exhibit better performance in some cases, they are generally inferior to MAHGRM. This result demonstrates the potential of MAHGRM in enhancing the accuracy, relevance and user satisfaction of recommendation systems.

4.4 Clustering Experiment

The final node representation features are reduced to two-dimensional space using t-SNE, and the K-means clustering method is used for node clustering. NMI and ARI are used as evaluation indicators. The experimental results of the above comparison models are listed in Table 3.

Table 3 Clustering experimental results of different models

Dataset	Index	GCN	GAT	HAN	HPN	OSGNN	ie-HGCN	MAHGRM
MovieLens-1M	NMI	0.4127	0.4011	0.7821	0.7634	0.0542	0.0427	0.8014
	ARI	0.3017	0.2627	0.7681	0.7955	0.0041	0.0034	0.8243
AL-MAHGRM	NMI	0.4058	0.3974	0.7688	0.7811	0.0317	0.0314	0.7792
	ARI	0.2548	0.2305	0.7456	0.8005	0.0005	0.0005	0.8054
Book-Crossing	NMI	0.5731	0.4624	0.8214	0.8153	0.0612	0.0573	0.8537
	ARI	0.4108	0.3123	0.8351	0.8204	0.0032	0.0027	0.8614

Table 3 shows the performance comparison of different models using two clustering evaluation metrics (NMI and ARI) on three datasets (MovieLens-1M, AL-MAHGRM and Book-Crossing). It can be seen from the table that for the MovieLens-1M dataset, the MAHGRM model has the highest scores on NMI and ARI, which are 0.8014 and 0.8243 respectively. The semantic-level attention of MAHGRM dynamically fuses multi-dimensional interests, improving the cross-class user separability of the dataset. Baseline models such as HAN and HPN are limited by the semantic fusion of a single path, resulting in the lack of cross-meta-path feature interaction and poor clustering effect of the models. On the AL-MAHGRM dataset, MAHGRM also achieved the best performance, with its NMI and ARI scores of 0.7792 and 0.8054 respectively, slightly higher than those of other models. Finally, on the Book-Crossing dataset, the MAHGRM model also achieved the highest scores on the NMI and ARI metrics, which were 0.8537 and 0.8614 respectively, demonstrating the superior performance of this model on these datasets. MAHGRM filters out low-influence nodes through a dynamic threshold t to effectively eliminate feature noise. The GCN/GAT model causes user nodes to be diffused in the embedding space due to the lack of denoising, while MAHGRM retains key neighbors through the Select function in Equation (9), thereby enhancing the NMI. In contrast, the performance of OSGNN and ie-HGCN on these three datasets is relatively weak. Overall, the MAHGRM model demonstrated leading performance on the considered datasets and evaluation metrics.

4.5 Visualization

According to Figure 3, we can observe the node distribution of different graph neural network models in the feature space after dimensionality reduction on the AL-MAHGRM dataset. The node distribution of GAT is relatively scattered, and there is a certain overlap between nodes of different colors (representing different categories). The clustering effect is general, and the boundaries between some categories are not clear. The node distribution of ie-HGCN, HAN, HPN and OSGNN models is relatively concentrated, and there is an obvious clustering trend between nodes of different colors. The clustering effect is good, and the boundaries between categories are relatively clear. Among them, the node distribution of MAHGRM model is the most compact, and there is a very good clustering effect between nodes of different colors, showing its strong generalization ability and clustering effect on complex graph structure data.

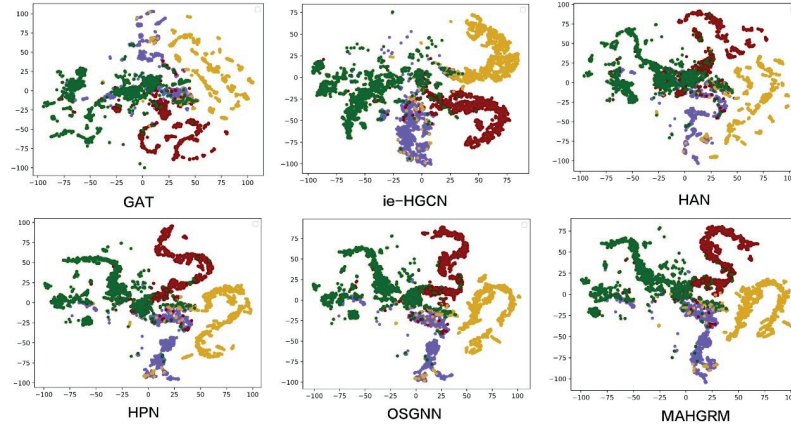


Figure 3 Visualization results of different models on the AL-MAHGRM dataset.

4.6 Ablation Experiment

In order to evaluate the performance of the model more deeply and verify the effectiveness of the method proposed in this paper, this section will conduct ablation experiments on the relevant modules in the MAHGRM-HGNN model. Specifically, this paper will ablate the relationship embedding module to produce the following variant models:

- MAHGRM-HGNN (complete model): a complete model including node-level aggregation and feature fusion module, semantic-level aggregation module and importance analysis module.
- MAHGRM w/o NLA: Remove the node-level aggregation and feature fusion modules, and only use the semantic-level aggregation module and importance analysis module.
- MAHGRM w/o SLA: Remove the semantic-level aggregation module, and only use the node-level aggregation and feature fusion module and importance analysis module.
- MAHGRM w/o IA: Remove the importance analysis module, and only use the node-level aggregation and feature fusion module and semantic-level aggregation module.
- MAHGRM w/o All: Remove all modules and only use the basic graph neural network (GCN) for recommendation.

The experimental results are shown in Table 4:

Table 4 compares the performance of different variants of the MAHGRM model on three datasets (MovieLens-1M, AL-MAHGRM, and

Table 4 Experimental results of different variant models

Dataset	Index	MAHGRM-HGNN	w/o NLA	w/o SLA	w/o IA
MovieLens-1M	AUC	0.917	0.904	0.924	0.912
	ACC	0.867	0.853	0.857	0.839
	F1	0.871	0.862	0.852	0.845
AL-MAHGRM	AUC	0.834	0.820	0.817	0.834
	ACC	0.792	0.764	0.782	0.749
	F1	0.761	0.739	0.745	0.758
Book-Crossing	AUC	0.824	0.813	0.806	0.792
	ACC	0.752	0.693	0.685	0.677
	F1	0.735	0.698	0.703	0.682

Book-Crossing), measured by three evaluation metrics: AUC, ACC, and F1. Overall, the original MAHGRM-HGNN model performs best in most cases, especially on the MovieLens-1M dataset, where it leads other variants in terms of AUC, ACC, and F1. The effects of removing different modules (such as NLA, SLA, and IA) on model performance vary. In some cases, such as the AUC metric of the MovieLens-1M dataset, removing SLA (MAHGRM w/o SLA) is even slightly better than the full MAHGRM-HGNN model, but is inferior to the full version in terms of ACC and F1. On the AL-MAHGRM and Book-Crossing datasets, although the differences between the variants are relatively small, the overall trend shows that the lack of any component will lead to some degree of performance degradation, especially on the AL-MAHGRM dataset, the ACC and F1 scores are significantly reduced, which shows that these components are important for maintaining the overall performance of the model. The removal of the importance Analysis module (IA) led to a 2.6% decrease in the F1 value of the MovieLens-1M dataset, indicating that the introduction of popular products can effectively break through the “information cocoon” of the recommendation system and enhance the surprise of the recommendation results. Comprehensive analysis shows that the MAHGRM-HGNN model achieves optimal or near-optimal performance through the joint action of its three designed modules.

4.7 Hyperparameter Experiments

In order to study the influence of hyperparameters on the performance of the MAHGRM-HGNN model, this paper conducted relevant hyperparameter

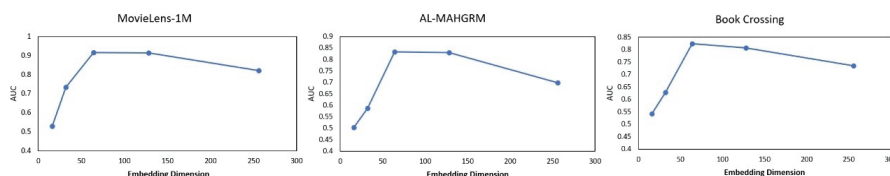


Figure 4 Changes in model performance at different embedding dimensions.

experiments. Specifically, this paper examined the effects of embedding dimension, the number of attention heads, dropout rate and threshold t on the model's performance.

In terms of embedding dimensions, this paper evaluated the model performance under 16, 32, 64, 128 and 256 embedding dimensions. Regarding the number of attention heads, this paper analyzed the impact of 1, 2, 4, 8 and 16 different head numbers on the model performance. For the dropout rate, this paper tested the model performance under dropout rates of 0.0001, 0.0005, 0.001, 0.005 and 0.01. For threshold t , this paper examined the model performance when selecting the top 3%, 4%, 5%, 7% and 10% of nodes as popular nodes. The specific results are shown in Figure 4.

Figure 4 shows the impact of different embedding dimensions on the model AUC indicators on three datasets (MovieLens-1M, AL-MAHGRM, and Book-Crossing). As can be seen from the figure, as the embedding dimension increases, the AUC value increases significantly in the early stage, reaches a peak value, and then gradually decreases. Specifically, on the MovieLens-1M dataset, when the embedding dimension is 50, the AUC value rises rapidly to close to 0.9, then it reaches the highest point at 100 dimensions, and then the AUC value decreases slightly as the dimension continues to increase. Similarly, similar trends are observed on the AL-MAHGRM and Book-Crossing datasets, that is, the AUC value rises rapidly in the initial stage, reaches a peak value at around 100 dimensions, and then slowly decreases as the dimension increases further. This shows that for these three datasets, choosing the right embedding dimension is crucial to optimizing model performance, and too low or too high dimensions may lead to performance degradation.

Figure 5 shows the impact of different numbers of attention heads on the model AUC indicators on three datasets (MovieLens-1M, AL-MAHGRM, and Book-Crossing). As can be seen from the figure, as the number of attention heads increases, the AUC value increases significantly in the early stage, reaches a peak, and then gradually decreases. Specifically, on the MovieLens-1M dataset, when the number of attention heads is 2, the AUC value rises

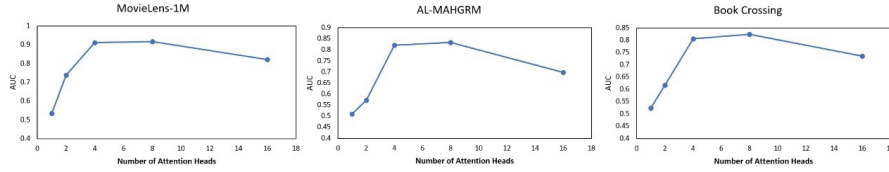


Figure 5 Changes in model performance under different attention heads.

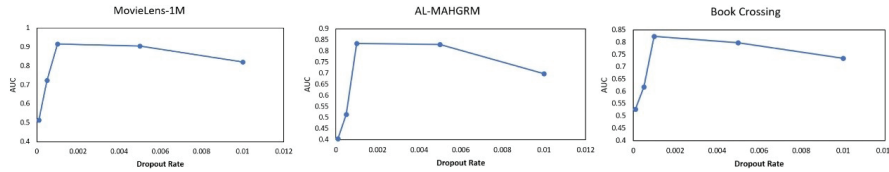


Figure 6 Changes in model performance at different dropout rates.

rapidly to close to 0.9, then it reaches the highest point at 4 attention heads, and then the AUC value decreases slightly as the number of attention heads continues to increase. Similarly, similar trends are observed on the AL-MAHGRM and Book-Crossing datasets, that is, the AUC value rises rapidly in the initial stage, reaches a peak at around 4 attention heads, and then slowly decreases as the number of attention heads increases further. This shows that for these three datasets, choosing the right number of attention heads is crucial to optimizing model performance, and too low or too high a number of attention heads may lead to performance degradation.

Figure 6 shows the impact of different dropout rates on the model AUC indicators on three datasets (MovieLens-1M, AL-MAHGRM, and Book-Crossing). As can be seen from the figure, as the dropout rate increases, the AUC value increases significantly in the early stage, and then gradually decreases after reaching a certain peak. Specifically, on the MovieLens-1M dataset, when the dropout rate is 0.002, the AUC value rises rapidly to close to 0.9, then it reaches the highest point at 0.004, and then the AUC value decreases slightly as the dropout rate continues to increase. Similarly, similar trends are observed on the AL-MAHGRM and Book-Crossing datasets, that is, the AUC value rises rapidly in the initial stage, peaks at around 0.004, and then slowly decreases as the dropout rate further increases. This shows that for these three datasets, choosing an appropriate dropout rate is crucial to optimizing model performance, and too low or too high a dropout rate may lead to performance degradation.

Figure 7 shows the impact of selecting different thresholds t on the AUC metrics of the model on three datasets (MovieLens-1M, AL-MAHGRM,

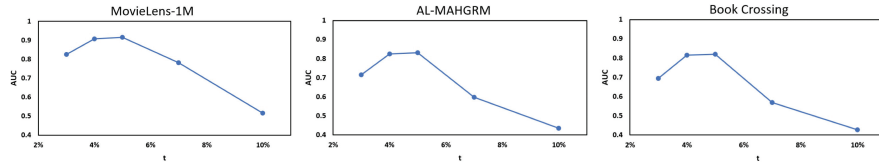


Figure 7 Changes in model performance at different threshold values of t .

Table 5 The clustering experimental results of different models

Indicators	GCN	GAT	HAN	OSGNN	MAHGRM
Times/epoch (s)	0.437	0.524	0.586	1.327	0.916
Memory (GB)	0.609	0.718	0.827	1.205	0.982

and Book-Crossing). The experimental results indicate that the threshold t has a significant impact on the recommendation accuracy (AUC) of the MAHGRM-HGNN model. The AUC values of the three datasets all decrease as t increases. When t reaches 10%, the AUC of MovieLens-1M drops to 0.516, AL-MAHGRM decreases from 0.832 to 0.436, and that for Book-Crossing decreases from 0.82 to 0.427. This indicates that although expanding the proportion of selected popular nodes can improve the diversity of recommendations, it will introduce more noise, resulting in the dilution of core interest features. MovieLens-1M reaches its peak at $t = 5\%$ (AUC = 0.915), while Book-Crossing shows a turning point at $t = 5\%$, which may be due to the stronger data sparsity of Book-Crossing. Moderately relaxing the t value can alleviate the problem of insufficient coverage of long-tail products. The most significant decrease in AUC for Book-Crossing indicates that data sparsity amplifies the impact of threshold selection. The smooth decline trend of AL-MAHGRM suggests that it has stronger robustness to the adjustment of t values.

4.8 Model Complexity Analysis

To assess the complexity of the model, this study conducted a comparison of the training times of different models on a hardware platform equipped with an i5-12400F CPU and an RTX2060 GPU, using the MovieLens-1M dataset. Table 5 summarizes the average training time and memory usage of each model during the training process.

From the data in the table, it can be seen that GCN, GAT and HAN perform most efficiently in terms of training time and memory usage. OSGNN, as a more complex model, significantly outperforms the other models in terms

of training time and memory usage, reaching 1.327 seconds and 1.205 GB respectively. It is worth noting that although the MAHGRM model has higher time and memory consumption due to the use of multiple attention mechanisms for feature fusion and importance analysis, compared to OSGNN, MAHGRM shows a significant advantage in resource utilization. Specifically, the training time of MAHGRM per epoch is 0.916 seconds and the memory consumption is 0.982 GB. This indicates that although MAHGRM has an increased demand for computing resources, it achieves efficient feature representation learning through optimized design and achieves a good balance between performance and resource consumption. In conclusion, although MAHGRM is slightly higher in time and memory consumption than some basic models, its trade-off between complexity and performance makes it an efficient and powerful recommendation system solution, especially suitable for applications that require high accuracy and diversity. This result further proves the feasibility and superiority of MAHGRM in practical applications.

5 Conclusion

The proposed multi-level attention heterogeneous graph recommendation model (MAHGRM) provides an innovative solution to the challenges of existing recommendation systems based on graph neural networks (GNNs), namely, accurately capturing the common hobby features of users, processing users' preferences, and effectively utilizing the information of intermediate node products. By introducing node-level and semantic-level attention mechanisms, MAHGRM can generate feature embeddings that more accurately reflect users' interest preferences and effectively handle users' complex and diverse interests. In addition, by integrating the features of different meta-paths in heterogeneous graphs, the model avoids the loss of intermediate node information and enriches the feature representation of target nodes. In terms of feature fusion, MAHGRM proposes a novel method to comprehensively consider the topological structure of the graph, ensuring that the model can fully utilize all types of node information in the graph. At the same time, in order to address the novelty and diversity of recommended content, MAHGRM introduces a product node importance evaluation strategy, which improves user experience by selecting popular but non-repetitive products for recommendation. Extensive experimental results show that MAHGRM significantly outperforms existing GNN methods on real-world datasets, demonstrating its superior performance in improving recommendation accuracy and personalization. In conclusion, MAHGRM

represents an important advancement in the field of recommendation systems and proposes an effective solution to address the limitations of GNN methods in recommendation systems. Although MAHGRM has demonstrated superior performance on multiple benchmark datasets, there are still some limitations. Although the importance analysis module enhances diversity through the injection of popular products, the coverage of ultra-long-tail products is still insufficient. Moreover, the model may face the challenge of computational cost when dealing with extremely large-scale heterogeneous graphs, as well as the possible dependence on different meta-paths in certain specific application scenarios. Furthermore, in the future, the model can be improved by introducing a more efficient attention mechanism, enhancing the scalability of the model, and combining the dynamic changes of user behavior to improve adaptability.

References

- [1] S. Rendle, C. Freudenthaler, and L. Schmidt-Thieme, “Factorizing personalized Markov chains for next-basket recommendation,” in *Proceedings of the 19th international conference on World wide web*, 2010, pp. 811–820.
- [2] Y. Kiam Tan, X. Xu, and Y. Liu, “Improved recurrent neural networks for session-based recommendations,” *arXiv e-prints*, pp. arXiv–1606, 2016.
- [3] Y. Chen and Y. Tang, “Attentive capsule graph neural networks for session-based recommendation,” in *International Conference on Knowledge Science, Engineering and Management*. Springer, 2022, pp. 602–613.
- [4] Q. Liu, Y. Zeng, R. Mokhosi, and H. Zhang, “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*, 2018, pp. 1831–1839.
- [5] J. Song, H. Shen, Z. Ou, J. Zhang, T. Xiao, and S. Liang, “Islf: Interest shift and latent factors combination model for session-based recommendation,” in *IJCAI*, 2019, pp. 5765–5771.
- [6] S. Wu, Y. Tang, Y. Zhu, L. Wang, X. Xie, and T. Tan, “Session-based recommendation with graph neural networks,” in *Proceedings of the AAAI conference on artificial intelligence*, vol. 33, no. 01, 2019, pp. 346–353.

- [7] Y. Chen, Q. Xiong, and Y. Guo, "Session-based recommendation: Learning multi-dimension interests via a multi-head attention graph neural network," vol. 131. Elsevier, 2022, p. 109744.
- [8] F. Liu, J. Wang, and J. Yang, "A graph neural network recommendation method integrating multi head attention mechanism and improved gated recurrent unit algorithm," *Ieee Access*, vol. 11, pp. 116 879–116 891, 2023.
- [9] J. Fang, B. Li, and M. Gao, "Collaborative filtering recommendation algorithm based on deep neural network fusion," vol. 34, no. 2. Inderscience Publishers (IEL), 2020, pp. 71–80.
- [10] Z. Guo and H. Wang, "A deep graph neural network-based mechanism for social recommendations," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2776–2783, 2020.
- [11] R. Zhao, X. Xiong, X. Zu, S. Ju, Z. Li, and B. Li, "A hierarchical attention recommender system based on cross-domain social networks," *Complexity*, vol. 2020, no. 1, p. 9071624, 2020.
- [12] G. Zhu, Y. Wang, J. Cao, Z. Bu, S. Yang, W. Liang, and J. Liu, "Neural attentive travel package recommendation via exploiting long-term and short-term behaviors," *Knowledge-Based Systems*, vol. 211, p. 106511, 2021.
- [13] L. Huang, M. Fu, F. Li, H. Qu, Y. Liu, and W. Chen, "A deep reinforcement learning based long-term recommender system," *Knowledge-based systems*, vol. 213, p. 106706, 2021.
- [14] H. Zhu, "Research on human resource recommendation algorithm based on machine learning," *Scientific Programming*, vol. 2021, no. 1, p. 8387277, 2021.
- [15] Q. Yuan, "Retracted: Network education recommendation and teaching resource sharing based on improved neural network," *Journal of Intelligent & Fuzzy Systems*, vol. 39, no. 4, pp. 5511–5520, 2020.
- [16] L. Zheng, Z. Tianlong, H. Huijian, and Z. Caiming, "Personalized tag recommendation based on convolution feature and weighted random walk," *International Journal of Computational Intelligence Systems*, vol. 13, no. 1, pp. 24–35, 2020.
- [17] Y. Zhao, H. Liu, and H. Duan, "Hgnn-gams: Heterogeneous graph neural networks for graph attribute mining and semantic fusion," *IEEE Access*, 2024.
- [18] Y. Zhao, S. Xu, and H. Duan, "Hgnn- brfe: Heterogeneous graph neural network model based on region feature extraction," *Electronics*, vol. 13, no. 22, p. 4447, 2024.

- [19] C. Xu, P. Zhao, Y. Liu, V. S. Sheng, J. Xu, F. Zhuang, J. Fang, and X. Zhou, “Graph contextualized self-attention network for session-based recommendation.” in *IJCAI*, vol. 19, no. 2019, 2019, pp. 3940–3946.
- [20] L. Dong, G. Zhu, Y. Wang, Y. Li, J. Duan, and M. Sun, “A graph positional attention network for session-based recommendation,” *IEEE Access*, vol. 11, pp. 7564–7573, 2023.
- [21] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” *arXiv preprint arXiv:1609.02907*, 2016.
- [22] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, “Graph attention networks,” *arXiv preprint arXiv:1710.10903*, 2017.
- [23] X. Wang, H. Ji, C. Shi, B. Wang, Y. Ye, P. Cui, and P. S. Yu, “Heterogeneous graph attention network,” in *The world wide web conference*, 2019, pp. 2022–2032.
- [24] Y. Yang, Z. Guan, J. Li, W. Zhao, J. Cui, and Q. Wang, “Interpretable and efficient heterogeneous graph convolutional network,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 2, pp. 1637–1650, 2021.
- [25] H. Ji, X. Wang, C. Shi, B. Wang, and P. S. Yu, “Heterogeneous graph propagation network,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 1, pp. 521–532, 2021.
- [26] Y. Yan, C. Li, Y. Yu, X. Li, and Z. Zhao, “Osgnn: Original graph and subgraph aggregated graph neural network,” *Expert Systems with Applications*, vol. 225, p. 120115, 2023.
- [27] H. Vaghari, M. Hosseinzadeh Aghdam, and H. Emami, “Group attention for collaborative filtering with sequential feedback and context aware attributes,” *Scientific Reports*, vol. 15, no. 1, p. 10050, 2025.

Biographies



Lingling Kong is a student majoring in Computer Science and Technology at Hainan University. Her current research interests include neural networks.



Hongyan Deng majors in Computer Science and Technology at Hainan University. Her mainly research interests include neural networks.



Jiayi Huang studies Computer Science and Technology at Hainan University. Her research interests include neural networks.

