Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

LLM-GRec: Large Language Model Enhanced Graph Recommendation via Granular-Ball Computing

Xiaofei Zhu

zxf@cqut.edu.cn

Chongqing University of Technology

Jiale Shen

Chongqing University of Technology

Research Article

Keywords: Recommender Systems, Large Language Models, Granular-Ball Computing,Knowledge Graph 1

Posted Date: November 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11325134/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

LLM-GRec: Large Language Model Enhanced Graph Recommendation via Granular-Ball Computing

Xiaofei Zhu and Jiale Shen

College of Computer Science and Engineering, Chongqing University of Technology, Hongguang Avenue, Chongqing, 400054, Chongqing, China.

*Corresponding author(s). E-mail(s): zxf@cqut.edu.cn;
Contributing authors: shenjiale@stu.cqut.edu.cn;

Abstract

Large Language Models (LLMs) have been introduced into recommender systems to enhance semantic understanding. However, LLM-generated user interests often consist of discrete phrases containing noise, redundancy, and structural instability. Moreover, the semantic gap between generated interests and collaborative recommendation objectives limits effective user–item matching. To address these challenges, we propose LLM-GRec, a robust graph-based recommendation framework enhanced by granular-ball computing. First, historical interactions and collaborative candidate items guide LLMs to generate recommendation-oriented user interests. Second, granular-ball computing organizes fine-grained interests into stable, multi-granularity interest anchors. Third, a unified graph integrates users, items, knowledge entities, and interest anchors for graph representation learning. An anchor alignment constraint encourages semantic consistency between user and item representations in the shared interest space. Furthermore, a quality-aware interest masking and reconstruction mechanism reduces the influence of noisy interests during graph propagation and recommendation learning. Experiments on three real-world datasets demonstrate that LLM-GRec outperforms existing recommendation methods on most evaluation metrics, with particularly notable improvements in sparse scenarios. The proposed framework effectively integrates semantic interest modeling, structured representation learning, and robust graph recommendation. Source code is available at https://github.com/shenjiale534/LLM-GRec.

Keywords: Recommender Systems, Large Language Models, Granular-Ball Computing,Knowledge Graph

1 Introduction

Personalized recommender systems have become essential for alleviating information overload Lai et al (2024). Graph neural networks (GNNs) have emerged as an effective approach for modeling complex interaction structures Wu et al (2021b). Representative methods, such as NGCF Wang et al (2019b) and LightGCN He et al (2020), aggregate neighborhood information over user–item interaction graphs to capture high-order collaborative signals and improve recommendation performance.

However, GNN-based recommendation methods still face data sparsity and limited semantic understanding. Most rely on ID embeddings, which struggle to capture the semantic intentions underlying user behaviors, particularly in sparse scenarios Kang and McAuley (2018). To overcome this limitation, recent studies incorporate LLMs into recommendation systems Wei et al (2024); Ren et al (2024). With strong semantic understanding and reasoning capabilities, LLMs extract rich interest information from historical interactions, providing semantic augmentation for collaborative recommendation.

Despite these advantages, LLM-enhanced recommendation methods face several challenges. First, many approaches generate user interests primarily from historical behavior descriptions without sufficiently incorporating collaborative signals, resulting in incomplete interest representations. Second, LLM-generated interests are typically discrete keywords or phrases containing redundancy, overlap, and semantic noise. Directly integrating these fine-grained interests as graph nodes may destabilize representation learning. Furthermore, LLM hallucinations can introduce interests inconsistent with actual user preferences Ji et al (2023), and such noise may be amplified during graph propagation. Therefore, constructing stable, reliable, and recommendation-oriented interest representations remains a critical challenge.

To address these limitations, we propose LLM-GRec, a robust graph-based recommendation framework integrating collaborative-enhanced interest generation, granular-ball computing, and interest-anchor alignment. First, historical interaction descriptions and collaborative candidate items jointly guide LLMs to generate interests reflecting both semantic preferences and collaborative patterns. Second, granular-ball computing adaptively aggregates fine-grained interest phrases into stable, multi-granularity interest anchors, reducing redundancy and noise. Third, a semantic-enhanced graph integrates user–item interactions, item knowledge semantics, and interest anchors for graph representation learning. An anchor alignment mechanism encourages user and item representations to maintain consistent distributions within a shared interest space. Finally, a quality-aware interest masking and reconstruction mechanism selectively masks low-quality interest anchors and reconstructs their representations, reducing noisy semantic propagation and improving recommendation robustness.

In summary, the main contributions of this work are summarized as follows:

1. We propose a collaborative-enhanced LLM interest generation module that integrates users' behavioral semantics with collaborative filtering signals, transforming implicit interaction patterns into explicit semantic interests and alleviating the semantic gap in traditional ID-based recommendation models.

2. We introduce granular-ball computing into interest modeling for recommender systems, learning robust granular-ball interest anchors through multi-granularity interest clustering, thereby improving model robustness under sparse and noisy environments.
3. We design a graph recommendation learning mechanism based on granular-ball interest anchor alignment, which aligns user and item representations while incorporating a quality-aware interest masking strategy to suppress the negative impact of low-quality semantic signals.
4. Extensive experiments on three real-world datasets demonstrate that LLM-GRec consistently outperforms existing state-of-the-art methods across multiple evaluation metrics and achieves notable improvements in sparse scenarios.

2 Related work

2.1 LLM-based Recommendation Methods

In recent years, LLMs have been increasingly applied to enhance semantic modeling in recommender systems through natural language understanding and knowledge reasoning. Traditional recommendation methods rely on user–item interactions and ID-based representations, which struggle to capture underlying semantic intentions. Thus, LLMs are employed to extract semantic information from historical behaviors, item descriptions, and external knowledge sources.

Existing approaches improve recommendation through semantic representation generation or task adaptation. RLMRec Ren et al (2024) integrates LLM-derived user and item representations with collaborative signals. LLMRec Wei et al (2024) generates semantic information to augment user–item graphs, while TALLRec Bao et al (2023) adapts LLMs through parameter-efficient fine-tuning.

However, LLM-generated interests often consist of discrete, redundant, or noisy phrases lacking stable structural organization and alignment with collaborative representations. Moreover, hallucinated interests may deviate from actual user preferences Ji et al (2023), introducing noise during graph propagation. Although LLM-ESR Liu et al (2024) enhances semantic representations in sparse scenarios, explicit interest quality modeling and structured alignment mechanisms remain insufficient. Therefore, this work develops collaborative-enhanced interest generation, granular-ball interest clustering, and interest anchor alignment to construct stable, robust, and recommendation-oriented interest representations for graph-based recommendation.

2.2 Multi-granularity Interest Modeling Methods

In real-world recommendation scenarios, user interests are diverse and hierarchical, making single-vector representations insufficient to capture complex preferences. Thus, multi-interest modeling has become an important research direction. MIND Li et al (2019) extracts multiple interest vectors through dynamic routing. ComiRec Cen et al (2020) employs self-attention to learn multi-interest representations and select relevant interests. SINE Tan et al (2021) models interest activation patterns through sparse

representations, while MIRN Yu et al (2022b) enhances user behavior representations through multi-interest routing.

Despite their effectiveness, these methods primarily discover interests using fine-grained point representations without explicitly modeling the global interest structure. Such representations are vulnerable to anomalous samples under sparse or noisy conditions. Granular Ball Computing Xia et al (2022) partitions data into adaptive, multi-granularity hyperspherical granules, preserving structural information while improving noise robustness. This offers a promising approach to multi-granularity interest modeling.

However, existing methods mainly learn multiple interest vectors or clusters, with limited consideration of how multi-granularity interest structures can be associated with user and item representations in graph recommendation. Moreover, LLM-generated interest phrases often contain discrete and noisy semantics, requiring more stable structural modeling. Therefore, this work introduces granular-ball interest clustering and incorporates the resulting interest anchors into graph-based recommendation representation learning.

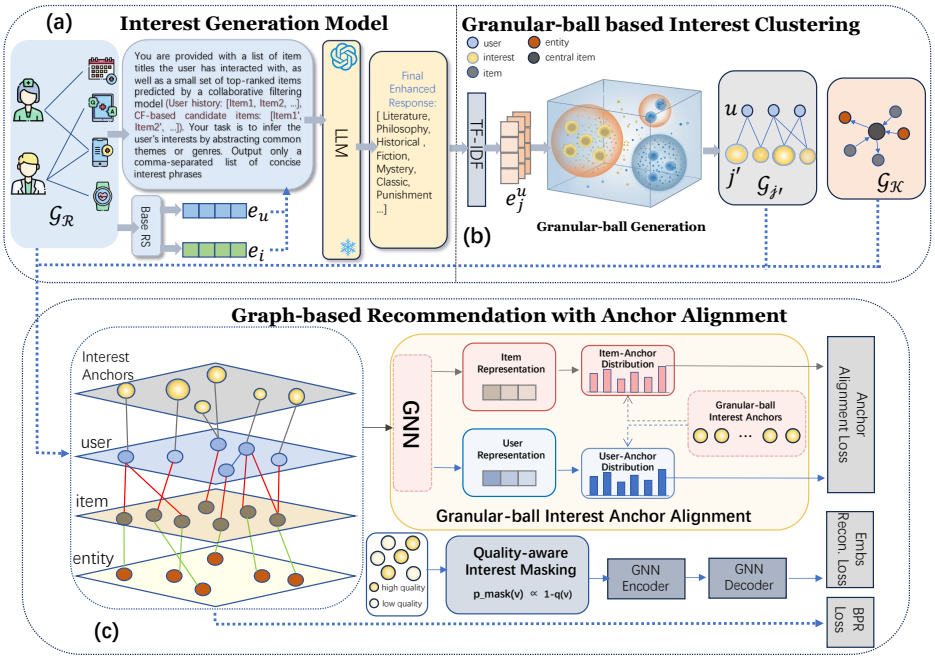


Fig. 1 The overall architecture of LLM-GRec.(a) LLM-based interest generation.(b) Granular-ball-based interest clustering.(c) Graph recommendation with anchor alignment and quality-aware masking

2.3 GNN-based Recommendation Methods

Graph Neural Networks (GNNs) have been widely adopted in recommender systems for modeling high-order user–item relationships. NGCF Wang et al (2019b) captures collaborative signals through multi-layer propagation on user–item graphs. LightGCN He et al (2020) simplifies graph convolution by removing feature transformations and nonlinear activations, retaining neighborhood aggregation to reduce complexity and improve recommendation performance.

Subsequent studies incorporate knowledge graphs, intent disentanglement, and self-supervised learning to enhance graph representations. KGAT Wang et al (2019a) models item semantic relationships using knowledge graphs, while DGCF Wang et al (2020) captures latent intents through disentangled graph convolution. SGL Wu et al (2021a) and SimGCL Yu et al (2022a) improve representation robustness through graph augmentation and contrastive learning. LightGCL Cai et al (2023) enhances graph representation consistency, whereas DiffRec Wang et al (2023) reconstructs user–item interaction representations using diffusion models.

Despite these advances, existing graph recommendation models mainly rely on fine-grained interaction edges and insufficiently capture the semantic structures underlying user interests. Although knowledge graphs and contrastive learning provide complementary information, explicit interest structure constraints remain limited. Consequently, unreliable neighbors and semantic noise may destabilize node representations, particularly in sparse scenarios. To address these limitations, this work introduces granular-ball interest anchors and integrates structured user interests into graph recommendation learning, aligning user and item representations within a unified interest semantic space.

3 Methods

3.1 Problem Definition

Let the user set and item set be denoted as $\mathcal{U} = u_1, u_2, \ldots, u_{|\mathcal{U}|}$ and $\mathcal{I} = i_1, i_2, \ldots, i_{|\mathcal{I}|}$, respectively. The user–item interaction data is represented as $\mathcal{R}_{ui} \subseteq \mathcal{U} \times \mathcal{I}$, where $(u, i) \in \mathcal{R}_{ui}$ indicates that user u has a historical interaction with item i. Accordingly, the user–item interaction graph can be formulated as:

$$\mathcal{G}_{ui} = (\mathcal{U} \cup \mathcal{I}, \mathcal{R}_{ui}). \tag{1}$$

For each user $u \in \mathcal{U}$, the historical interaction sequence is denoted as $\mathcal{H}_u = [i_1, i_2, \ldots, i_{n_u}]$, where n_u represents the length of user u's historical interactions. Let t_i denote the textual information of item i. The textual history of user u can then be represented as $\mathcal{T}_u = [t_{i_1}, t_{i_2}, \ldots, t_{i_{n_u}}]$. Based on users' historical behaviors and item textual information, we LLMs to generate user interests and further obtain an interest anchor set $\mathcal{J}$ through granular-ball computing. The associations between users and interest anchors are denoted as $\mathcal{R}_{uj} \subseteq \mathcal{U} \times \mathcal{J}$.

Meanwhile, to enrich item-side semantic information, we incorporate a knowledge graph $\mathcal{G}_{kg} = (\mathcal{E}, \mathcal{R}_{kg}, \mathcal{T}_{kg})$, where $\mathcal{E}$ represents the entity set, $\mathcal{R}_{kg}$ denotes the relation

set, and $\mathcal{T}_{kg}$ denotes the set of knowledge triples. Based on user–item interactions, item-side knowledge graph information, and user–interest anchor relationships, we construct a unified graph $\mathcal{G}$, upon which user and item representations are learned. The ultimate objective is to predict users' preference scores for unobserved items:

$$\hat{y}_{ui} = f_{\Theta}(u, i|\mathcal{G}), \tag{2}$$

where $\hat{y}_{ui}$ denotes the predicted preference score of user u towards item i, $f_{\Theta}(\cdot)$ represents the proposed recommendation model, and Θ denotes the model parameters. The candidate items are ranked according to the predicted scores, and the Top-N recommendation results are generated for each user.

3.2 Overall Framework

The overall framework of LLM-GRec is illustrated in Fig. 1, comprising three modules. First, the user interest generation module (Fig. 1(a)) combines historical interactions and collaborative candidate items to guide LLMs in generating user interests. Second, the granular-ball-based interest modeling module (Fig. 1(b)) aggregates fine-grained interests into stable interest anchors. Finally, the graph recommendation module (Fig. 1(c)) integrates user–item interactions, knowledge entities, and interest anchors for GNN-based representation learning. Anchor alignment and quality-aware interest masking further enhance semantic consistency and robustness through joint optimization of recommendation, alignment, and reconstruction losses.

3.3 Interest Generation Model

As illustrated in Fig. 1(a), we design the Interest Generation Model to integrate collaborative filtering signals and textual semantic information, enabling Large Language Models to generate explicit user interest descriptions. Unlike existing interest generation approaches that solely rely on users' historical interaction texts, we further incorporate collaborative candidate items as auxiliary contexts, allowing LLMs to simultaneously capture users' explicit historical preferences and latent collaborative preferences.

3.3.1 Collaborative Signal Extraction

To mine latent collaborative relationships from user–item interaction data, we first learn collaborative representations of users and items based on the user–item interaction graph $\mathcal{G}_R$. Specifically, a basic recommendation model BaseRS is employed to model the interaction graph and obtain user and item collaborative representations:

$$\mathbf{e}_u^c, \mathbf{e}_i^c = \text{BaseRS}(\mathcal{G}_R), \tag{3}$$

where $\mathbf{e}_u^c$ and $\mathbf{e}_i^c$ denote the collaborative representations of user u and item i, respectively.

Based on these representations, we calculate the matching scores between users and candidate items to identify items that users may potentially be interested in:

$$s^c(u,i) = \mathrm{sim}(\mathbf{e}_u^c, \mathbf{e}_i^c), \tag{4}$$

where $\mathrm{sim}(\cdot)$ denotes the similarity function. Subsequently, we select the Top-K items with the highest scores from the set of items that the user has not interacted with, forming the collaborative candidate item set:

$$\mathcal{C}_u = \mathrm{TopK}\left(s^c(u,i) \mid i \notin \mathcal{I}_u^+\right), \tag{5}$$

where $\mathcal{I}_u^+$ denotes the set of items historically interacted with by user u. The collaborative candidate set $\mathcal{C}_u$ provides additional latent preference signals from the perspective of interaction structures, offering richer contextual information for subsequent LLM-based interest reasoning. In particular, it helps alleviate the insufficiency of user interest information in sparse interaction scenarios.

3.3.2 LLM-based Semantic Interest Generation

After obtaining the collaborative neighbor set $\mathcal{C}_u$, we further leverage LLM to perform semantic modeling of user interests, thereby compensating for the limitation of traditional collaborative filtering methods in semantic representation. Specifically, for each user u, the input to the LLM consists of two components. The first is the user's historical interaction text sequence $\mathcal{T}_u$, which reflects explicit user preferences. The second is the textual set corresponding to collaborative neighboring items, denoted as $\mathcal{T}_u^{cf} = [t_{c_1}, t_{c_2}, \ldots, t_{c_K}]$, where $c_k \in \mathcal{C}_u$, which serves to supplement potential interest information. Based on this, the two types of textual information are jointly constructed as a prompt and fed into the LLM for semantic reasoning:

$$\mathbf{z}_u = \mathrm{LLM}(\mathcal{T}_u, \mathcal{T}_u^{cf}), \tag{6}$$

where $\mathbf{z}_u$ denotes the semantic representation of user interests. Unlike approaches that rely solely on user historical behaviors, the incorporation of collaborative neighbor information provides structured supplementary signals for LLMs, enabling them to generate more comprehensive and stable semantic representations, especially under sparse interaction scenarios.

3.4 Granular-ball based Interest Clustering

To extract more robust and structured semantic representations from the fine-grained interest phrases generated by LLMs, we propose the Granular-ball based Interest Clustering module, as illustrated in Fig. 1(b). Since LLM-generated interests are typically represented as discrete phrases, they may contain redundant, overlapping, and noisy semantic information. Directly incorporating such fine-grained interests as graph nodes into recommendation models may reduce the stability of representation learning. Therefore, inspired by the idea of granular ball computing Xia et al (2025), we

adaptively aggregate interest phrases with different granularities and organize fine-grained interests into more stable interest anchors, thereby improving the robustness of interest structures.

3.4.1 Interest Phrase Representation and Local Semantic Modeling

Based on the user interest phrase set generated by the LLM in the previous section, we regard each generated keyword or semantic phrase as a basic interest unit for subsequent interest structure modeling. To capture the semantic relationships among different interest phrases, we employ the TF-IDF method to encode each interest phrase and obtain its vector representation:

$$\mathbf{x}_j^u = \mathrm{TFIDF}(z_j^u), \tag{7}$$

where $\mathbf{x}_j^u \in \mathbb{R}^d$ denotes the vector representation of the j-th interest phrase generated for user u.

Based on these representations, we measure the semantic similarity between different interest phrases using cosine similarity:

$$s(\mathbf{x}_j^u, \mathbf{x}_k^v) = \frac{(\mathbf{x}_j^u)^\top \mathbf{x}_k^v}{|\mathbf{x}_j^u||\mathbf{x}_k^v|}. \tag{8}$$

Subsequently, for each interest phrase, we retain only the Top-K most similar neighbors to construct a local semantic neighborhood structure. This neighborhood structure provides the foundation for subsequent granular ball partitioning, enabling semantically related interest phrases to be aggregated into the same high-level interest anchors.

3.4.2 Granular Ball Partitioning and Interest Anchor Graph Construction

After obtaining the vector representations of interest phrases, we further apply granular ball computing to partition the interest space into multiple granularities. Formally, a granular ball can be represented as a set consisting of multiple semantically similar interest phrases:

$$\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{|\mathcal{B}|}\}, \tag{9}$$

where the interest phrases within $\mathcal{B}$ exhibit high semantic consistency in the embedding space. Each granular ball is regarded as a high-level interest anchor j', representing a group of semantically related fine-grained interests.

To obtain high-quality granular ball partitions, we adopt a quality-driven recursive partition strategy. For any granular ball $\mathcal{B}$, its quality is defined as:

$$Q(\mathcal{B}) = \frac{2|\mathcal{E}_\mathcal{B}|}{|\mathcal{V}_\mathcal{B}|} + \mathrm{Transitivity}(\mathcal{B}), \tag{10}$$

where $\mathcal{V}_{\mathcal{B}}$ and $\mathcal{E}_{\mathcal{B}}$ denote the node set and edge set within the granular ball, respectively. This quality function simultaneously measures the internal connectivity density and local structural consistency of a granular ball.

During the recursive partition process, if splitting a granular ball $\mathcal{B}$ into two sub-granular balls $\mathcal{B}_1$ and $\mathcal{B}_2$ leads to an improvement in the overall quality, the splitting operation is performed:

$$\text{Split}(GB) = \begin{cases} \text{True}, & \text{if } Q(B_1)+ \\ & Q(B_2) > Q(B) \\ \text{False}, & \text{otherwise.} \end{cases} \tag{11}$$

This process is recursively conducted until the granular ball quality can no longer be improved or the minimum granularity constraint is satisfied. Through this procedure, fine-grained interest phrases are aggregated into a more stable interest anchor set $\mathcal{J}'$.

Based on the granular ball partition results, we further construct a user–interest anchor graph. Specifically, if an interest phrase of user u is assigned to an interest anchor $j' \in \mathcal{J}'$, an edge is established between user node u and interest anchor j'. The resulting user–interest anchor graph is formulated as:

$$\mathcal{G}_{j'} = (\mathcal{U} \cup \mathcal{J}', \mathcal{E}_{uj'}), \tag{12}$$

where

$$\mathcal{E}_{uj'} = (u, j') \mid \exists z_j^u \in j'. \tag{13}$$

This graph structure transforms discrete LLM-generated interest phrases into structured interest anchors suitable for graph representation learning, providing semantic foundations for subsequent anchor alignment-based recommendation learning. Meanwhile, we incorporate the item-side knowledge graph $\mathcal{G}_K$ to supplement item semantic information, which is jointly utilized with the user–interest anchor graph in the subsequent graph recommendation learning module.

3.5 Graph-based Recommendation with Anchor Alignment

To unify user behavioral signals, item semantic knowledge, and granular-ball interest anchors into a recommendation learning framework, we propose the Graph-based Recommendation with Anchor Alignment module, as illustrated in Fig. 1(c). This module first constructs a semantic-enhanced heterogeneous graph by integrating the user–item interaction graph, item–entity knowledge graph, and user–interest anchor graph, and then employs GNNs to learn user and item representations. Subsequently, the granular-ball interest anchors are utilized as a shared semantic space to align the anchor distributions of user and item representations, enabling them to maintain semantic consistency at the high-level interest space. Finally, to alleviate the noise issue in automatically generated interests, we further design a quality-aware interest masking mechanism and jointly optimize the recommendation loss, anchor alignment loss, and embedding reconstruction loss.

3.5.1 Semantic-enhanced Graph Representation Learning

Based on the user–interest anchor graph $\mathcal{G}_{j'}$ obtained in the previous section, we integrate it with the user–item interaction graph and the item-side knowledge graph to construct a semantic-enhanced heterogeneous graph:

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}), \quad \mathcal{V} = \mathcal{U} \cup \mathcal{I} \cup \mathcal{E}_K \cup \mathcal{J}', \tag{14}$$

where $\mathcal{U}$, $\mathcal{I}$, $\mathcal{E}_K$, and $\mathcal{J}'$ denote the user set, item set, entity set, and granular-ball interest anchor set, respectively. The edge set $\mathcal{E}$ consists of user–item interaction edges, item–entity association edges, and user–interest anchor edges. Through this graph structure, the model can simultaneously capture collaborative signals, knowledge semantic signals, and interest anchor semantic signals within a unified representation space.

For any node $v \in \mathcal{V}$, we initialize its embedding representation as $\mathbf{h}_v^{(0)} \in \mathbb{R}^d$. Then, a GNN is employed to perform multi-layer message propagation over the heterogeneous graph. At the l-th layer, the representation of node v is updated as:

$$\mathbf{h}_v^{(l+1)} = \sigma\left(\mathbf{W}_0^{(l)}\mathbf{h}_v^{(l)} + \sum_{u \in \mathcal{N}(v)} \frac{1}{\sqrt{d_v d_u}} \mathbf{W}_1^{(l)}\mathbf{h}_u^{(l)}\right), \tag{15}$$

where $\mathcal{N}(v)$ denotes the neighbor set of node v, d_v and d_u represent the degrees of nodes v and u, respectively, $\mathbf{W}_0^{(l)}$ and $\mathbf{W}_1^{(l)}$ are learnable parameters at the l-th layer, and $\sigma(\cdot)$ denotes the nonlinear activation function.

After L layers of graph propagation, we aggregate representations from all layers to obtain the final node representation:

$$\mathbf{z}_v = \frac{1}{L+1} \sum_{l=0}^{L} \mathbf{h}_v^{(l)}. \tag{16}$$

3.5.2 Granular-ball Interest Anchor Alignment

Although the semantic-enhanced graph can jointly model multi-source information, there may still exist representation discrepancies between user-side interest anchors and item-side semantic knowledge. To address this issue, we utilize the granular-ball interest anchors $\mathcal{J}'$ as a shared semantic anchor space and compute the anchor distributions of user and item representations, respectively. The embedding matrix of granular-ball interest anchors is defined as:

$$\mathbf{A} = [\mathbf{a}_1; \mathbf{a}_2; \ldots; \mathbf{a}_{|\mathcal{J}'|}] \in \mathbb{R}^{|\mathcal{J}'| \times d}, \tag{17}$$

where $\mathbf{a}_{j'}$ denotes the embedding representation of the j'-th interest anchor.

For user u and item i, we calculate their correlations with all granular-ball interest anchors and obtain the anchor distributions through softmax:

$$\mathbf{p}_u^A = \text{softmax}\left(\frac{\mathbf{z}_u \mathbf{A}^\top}{\tau}\right), \quad \mathbf{p}_i^A = \text{softmax}\left(\frac{\mathbf{z}_i \mathbf{A}^\top}{\tau}\right), \tag{18}$$

where τ denotes the temperature coefficient, and $\mathbf{p}_u^A$ and $\mathbf{p}_i^A$ represent the semantic distributions of users and items over the granular-ball interest anchor space, respectively. These distributions characterize the associations between user preferences, item semantics, and different high-level interest anchors.

To ensure that users and items with interaction relationships maintain consistent semantic distributions in the interest anchor space, we design an anchor alignment loss:

$$\mathcal{L}_{\text{align}} = \frac{1}{|\mathcal{B}|} \sum_{(u,i)\in\mathcal{B}} D_{\text{KL}}\left(\mathbf{p}_u^A \| \mathbf{p}_i^A\right), \tag{19}$$

where $\mathcal{B}$ denotes the user–item interaction set within a training batch, and $D_{\text{KL}}(\cdot\|\cdot)$ represents the KL divergence. By minimizing this loss, the model transfers user-side granular-ball interest semantics into item representation learning, enabling user and item representations to achieve semantic alignment within the shared interest anchor space.

3.5.3 Quality-aware Interest Masking

Since interest anchors are obtained by aggregating LLM-generated interest phrases, they may still contain semantically ambiguous, insufficiently supported, or low-quality interest nodes. Assigning identical importance to all interest anchors may introduce noisy interests into the graph propagation process. Therefore, we introduce a quality-aware interest masking mechanism, which adaptively determines the masking and reconstruction probability of each interest anchor according to its quality.

For any interest anchor $j' \in \mathcal{J}'$, we calculate its quality score based on structural support and semantic consistency. Specifically, the item neighbor set of interest anchor j' is defined as:

$$\mathcal{N}_I(j') = \{i \in \mathcal{I} \mid (j', i) \in \mathcal{E} \text{ or } (i, j') \in \mathcal{E}\},$$

and its connection degree is denoted as $d_{j'} = |\mathcal{N}_I(j')|$. The average representation of neighboring items is calculated as:

$$\bar{\mathbf{h}}_{j'} = \frac{1}{d_{j'} + \epsilon} \sum_{i\in\mathcal{N}_I(j')} \mathbf{h}_i^{(0)}. \tag{20}$$

Here, ϵ is a smoothing term to avoid division by zero. Intuitively, an interest anchor connected with more items has stronger behavioral support. Meanwhile, if the representation of an interest anchor is more consistent with the representations of its neighboring items, the interest anchor is considered to have higher semantic reliability.

Therefore, we define the support quality and consistency quality of each interest anchor. The support quality is obtained by logarithmic normalization of its connection

degree, while the consistency quality is measured by the cosine similarity between the interest anchor representation and the average representation of neighboring items:

$$\begin{aligned} s(j') &= \mathrm{MinMax}\left(\log(1+d_{j'})\right), \\ c(j') &= \mathbb{I}(d_{j'}>0)\cdot\frac{1+\cos\left(\mathbf{h}_{j'}^{(0)},\bar{\mathbf{h}}_{j'}\right)}{2}. \end{aligned} \tag{21}$$

where $s(j') \in [0,1]$ denotes the structural support quality and $c(j') \in [0,1]$ denotes the semantic consistency quality. When an interest anchor has no item neighbors, its consistency quality is set to 0, indicating that the interest lacks neighborhood semantic information for validation.

Finally, the quality score of interest anchor j' is defined as the weighted combination of support quality and consistency quality:

$$q(j') = \alpha s(j') + (1-\alpha)c(j'), \tag{22}$$

where α is a hyperparameter. A larger weight assigned to support quality indicates that we emphasize whether an interest anchor has sufficient item-level structural support, while consistency quality further constrains the semantic agreement between interest anchors and their neighboring items.

Based on this quality score, we introduce a quality-aware interest masking strategy. Lower-quality interest anchors are more likely to contain noisy or weak semantic signals and therefore should be masked and reconstructed with higher probability. The masking probability is defined as:

$$p_{\mathrm{mask}}(j') = \frac{1-q(j')+\epsilon}{\sum_{k'\in\mathcal{J}'}\left(1-q(k')+\epsilon\right)}. \tag{23}$$

During training, the model samples a masked anchor set $\mathcal{M}$ from the interest anchor set according to $p_{\mathrm{mask}}(j')$. By preferentially masking low-quality interest nodes, the model is encouraged to recover unstable interest semantics from graph structural contexts, thereby reducing the influence of noisy interests on recommendation representation learning.

Specifically, the representations of masked interest nodes are replaced with mask vectors. Then, a GNN encoder–decoder architecture is employed to reconstruct the original representations of masked interest anchors from the remaining graph structure:

$$\hat{\mathbf{H}} = \mathrm{GNN}_{\mathrm{dec}}\left(\mathrm{GNN}_{\mathrm{enc}}(\tilde{\mathbf{H}},\tilde{\mathcal{E}}),\tilde{\mathcal{E}}\right), \tag{24}$$

where $\tilde{\mathbf{H}}$ and $\tilde{\mathcal{E}}$ denote the masked node representations and graph structure, respectively, and $\hat{\mathbf{H}}$ denotes the reconstructed representations generated by the decoder.

To constrain the recovery of semantic information for masked interest anchors, we adopt an embedding reconstruction loss:

$$\mathcal{L}_{\text{rec}} = \frac{1}{|\mathcal{M}|} \sum_{j' \in \mathcal{M}} \left\| \hat{\mathbf{h}}_{j'} - \mathbf{h}_{j'}^{(0)} \right\|_2^2. \tag{25}$$

This loss encourages the model to learn stable semantic representations of interest anchors from neighborhood contexts, thereby reducing the interference of low-quality interests during recommendation representation learning.

For recommendation optimization, we adopt the BPR loss to model the preference difference between positive and negative items:

$$\mathcal{L}_{\text{BPR}} = - \sum_{(u,i,i^-) \in \mathcal{O}} \log \sigma \left(\hat{y}_{ui} - \hat{y}_{ui^-} \right), \tag{26}$$

where $\mathcal{O}$ denotes the training sample set, i is an item interacted with by user u, and i^- is a negative item.

Finally, the overall optimization objective is defined as:

$$\mathcal{L} = \mathcal{L}_{\text{BPR}} + \lambda_1 \mathcal{L}_{\text{align}} + \lambda_2 \mathcal{L}_{\text{rec}} + \lambda_3 \|\Theta\|_2^2, \tag{27}$$

where λ_1, λ_2, and λ_3 are hyperparameters, and Θ denotes all learnable parameters in the model. By jointly optimizing the above objectives, the model not only preserves the ranking capability required by recommendation tasks, but also achieves semantic alignment between users and items through granular-ball interest anchors and improves the robustness of interest representations via quality-aware interest masking.

4 Experiments

To comprehensively evaluate the effectiveness, robustness, and interpretability of the proposed LLM-GRec, we conduct extensive experiments on three real-world datasets. Specifically, we investigate the following seven research questions:

- **RQ1:** Compared with existing state-of-the-art recommendation methods, can LLM-GRec achieve superior recommendation performance?
- **RQ2:** Do all key components of LLM-GRec make meaningful contributions to the final recommendation performance?
- **RQ3:** Is LLM-GRec sensitive to variations in key hyperparameters, and can it maintain stable performance?
- **RQ4:** Can LLM-GRec remain robust when facing low-quality interest anchors or semantic noise?
- **RQ5:** Are the learned granular-ball interest anchors and their alignment relationships interpretable, and can they reflect users' actual preferences?
- **RQ6:** While achieving improved recommendation performance, does LLM-GRec introduce significant additional training or inference costs?

Table 1 Performance comparison on three datasets. The best results are highlighted in bold, while the second-best results are underlined.

MODEL	DBbook2014				Book-Crossing				MovieLens-1M			
	R@50 ↑	R@100 ↑	N@50 ↑	N@100 ↑	R@50 ↑	R@100 ↑	N@50 ↑	N@100 ↑	R@50 ↑	R@100 ↑	N@50 ↑	N@100 ↑
LightGCN(SIGIR20)	0.4214	0.5358	0.2230	0.2481	0.1607	0.2246	0.0765	0.0911	0.4086	0.5574	0.3977	0.4432
SGL(SIGIR21)	0.4258	0.5249	0.2350	0.2567	0.1613	0.2154	0.0860	0.0982	0.4094	0.5505	0.4004	0.4433
SimGCL(SIGIR22)	0.4280	0.5369	0.2376	0.2610	0.1653	0.2211	0.0860	0.0988	0.4131	0.5560	0.4048	0.4482
CKE(KDD16)	0.3419	0.4448	0.1832	0.2060	0.1284	0.1691	0.0708	0.0811	0.3948	0.5421	0.3841	0.4295
CFKG(Al18)	0.3586	0.4559	0.1627	0.1816	0.1216	0.1771	0.0503	0.0617	0.3930	0.5439	0.3672	0.4142
KGAT(KDD19)	0.4174	0.5309	0.2110	0.2360	0.1564	0.2133	0.0865	0.1004	0.4059	0.5527	0.3928	0.4378
KGIN(WWW21)	0.4147	0.5337	0.2200	0.2462	0.1468	0.2075	0.0681	0.0819	0.4095	0.5581	0.4004	0.4453
KGCL(SIGIR22)	0.4308	0.5271	0.2453	0.2666	0.1562	0.2049	0.0859	0.0971	0.4083	0.5496	0.3996	0.4430
MCCLK(SIGIR22)	0.4376	0.5491	0.2054	0.2273	0.1476	0.2039	0.0656	0.0774	0.4176	0.5655	0.3876	0.4339
KGRec(KDD23)	0.4415	0.5497	0.2373	0.2611	0.1473	0.2075	0.0699	0.0836	0.4136	0.5633	0.4072	0.4524
CIKGRec(AAAI25)	0.4642	**0.5756**	0.2494	0.2739	0.1791	0.2455	0.0889	0.1041	0.4250	0.5718	0.4162	0.4609
LLM-GRec	**0.4671**	0.5741	**0.2576**	**0.2798**	**0.1866**	**0.2501**	**0.0967**	**0.1107**	**0.4341**	**0.5806**	**0.4230**	**0.4680**

- **RQ7:** Can LLM-GRec maintain superior recommendation performance compared with other models in sparse interaction scenarios?

4.1 Experimental Settings

4.1.1 Datasets

We evaluate LLM-GRec on three real-world datasets: **DBbook2014**, **Book-Crossing**, and **MovieLens-1M**, covering different scales, sparsity levels, and semantic characteristics. Explicit ratings are converted into implicit feedback, and each dataset is divided into training, validation, and test sets. Dataset statistics are presented in Table 2.

Table 2 Statistics of datasets.

Dataset	DBbook2014	Book-Crossing	MovieLens-1M
#Users	5,576	6,616	6,040
#Items	2,680	8,853	3,260
#Interactions	65,961	110,662	998,539
#Density	0.44%	0.19%	5.07%
#Relations	13	4	20
#Entities	8,762	1,404	14,377
#Triplets	134,223	1,137	415,104

4.1.2 Baselines

To comprehensively evaluate the performance of the proposed model, we select several representative recommendation methods as baselines. According to the utilized information sources and modeling paradigms, these methods can be categorized into the following three groups:

(1) Graph collaborative filtering-based recommendation methods: These methods mainly model high-order collaborative signals based on user–item interaction graphs.

- **LightGCN** He et al (2020): Simplifies graph convolution operations by retaining only neighborhood aggregation, enabling efficient collaborative signal modeling.
- **SGL** Wu et al (2021a): Introduces self-supervised learning through graph augmentation and contrastive learning to improve representation robustness.
- **SimGCL** Yu et al (2022a): Enhances representation generalization ability through simple noise-based contrastive learning.

(2) Knowledge graph-enhanced recommendation methods: These methods incorporate external knowledge graph information to capture semantic relationships among items, alleviating the semantic deficiency caused by relying solely on interaction data.

- **CKE** Zhang et al (2016): Jointly learns collaborative filtering signals and knowledge graph embeddings.
- **CFKG** Ai et al (2018): Unifies collaborative filtering and knowledge graph information into a unified graph structure for representation learning.
- **KGAT** Wang et al (2019a): Fuses user–item interactions and knowledge graph information through graph attention mechanisms.
- **KGIN** Wang et al (2021): Models user preferences under different relational semantics through intent disentanglement.
- **KGCL** Yang et al (2022): Combines knowledge graphs with contrastive learning to improve representation consistency and robustness.
- **MCCLK** Zou et al (2022): Jointly models collaborative signals and knowledge graph information through multi-view contrastive learning.
- **KGRec** Yang et al (2023): Further improves recommendation performance through knowledge-aware representation learning.

(3) LLM-enhanced recommendation methods: These methods exploit the semantic understanding and reasoning capabilities of LLMs to enhance recommendation representation learning.

- **CIKGRec** Hu et al (2025): Jointly models LLM-generated semantic information and knowledge-enhanced recommendation frameworks.

4.1.3 Evaluation Metrics

We adopt **Recall@K** and **NDCG@K** ($K \in \{50, 100\}$) to evaluate Top-N recommendation performance. Recall@K measures the proportion of relevant items retrieved, while NDCG@K evaluates ranking quality by assigning higher weights to relevant items ranked higher. Candidate items are ranked by predicted scores, and both metrics are calculated for each user and averaged across all users.

4.1.4 Implementation Details

We use consistent experimental settings for fair comparisons. The embedding dimension is $d = 128$, with a batch size of 512. Adam optimization and validation-based early stopping are adopted. GNN layers are selected from $\{2, 3\}$, learning rates from $\{1 \times 10^{-4}, 5 \times 10^{-4}\}$, and training epochs range from 55 to 120. The BPR loss optimizes recommendation ranking. For anchor alignment, $\lambda_{\text{align}} \in \{0.001, 0.01\}$ and

$\tau \in \{0.2, 0.5, 1.0\}$. The masking module uses a 2-layer encoder and 1-layer decoder, with $\lambda_{\text{rec}} \in \{0.01, 0.05, 0.1\}$ and SCE exponent $\alpha \in \{2, 3\}$. The masking ratio increases from $p_{\min} \in \{0.1, 0.3\}$ to $p_{\max} \in \{0.5, 0.7\}$. LLM-generated interests and granular-ball anchors are constructed offline and remain fixed during recommendation training, avoiding additional LLM inference overhead. All baselines use identical data splits, evaluation protocols, and experimental environments, with hyperparameters tuned on validation data to ensure fairness and reproducibility.

4.2 Overall Performance Comparison (RQ1)

Table 1 compares LLM-GRec with state-of-the-art recommendation methods on three datasets. LLM-GRec achieves the best results on most metrics, demonstrating the effectiveness of semantic interest modeling and robust representation learning. Semantic-enhanced methods generally outperform collaborative filtering approaches by alleviating data sparsity. Although LLM-based methods such as CIKGRec incorporate semantic knowledge, noisy generated interests may limit their effectiveness. LLM-GRec addresses this issue through granular-ball interest modeling and quality-aware denoising, producing more robust representations. Compared with CIKGRec, LLM-GRec achieves improvements on most metrics, including an 8.78% increase in NDCG@50 on Book-Crossing. Consistent gains across the three datasets, particularly in sparse scenarios, demonstrate the benefits of structured semantic augmentation and robust interest learning.

4.3 Ablation Study (RQ2)

We conduct ablation studies by removing or replacing key components of LLM-GRec to evaluate their contributions:

- **w/o LLM**: Removes LLM-generated interests to assess semantic information.
- **w/o CF-guided**: Excludes collaborative candidate items from interest generation to assess collaborative signals.
- **w/o GB**: Replaces granular-ball clustering with K-means to evaluate multi-granularity modeling.
- **w/o Mask**: Removes interest masking and reconstruction to assess noise suppression.
- **w/o GBIAA**: Removes interest anchor alignment to evaluate representation consistency.
- **w/o KG**: Excludes the knowledge graph to assess structured semantic information.

Table 3 presents the ablation results. Removing most components of LLM-GRec degrades recommendation performance, confirming their effectiveness. Removing LLM-generated interests (**w/o LLM**) causes substantial performance drops, particularly on sparse datasets, highlighting the importance of semantic augmentation. Excluding collaborative guidance (**w/o CF-guided**) slightly reduces performance, indicating its supplementary role in interest modeling. Replacing granular-ball clustering with K-means (**w/o GB**) reduces performance, demonstrating the benefits of multi-granularity interest modeling. Removing quality-aware masking (**w/o Mask**)

Table 3 Ablation study results of different LLM-GRec variants on DBbook2014, Book-Crossing, and MovieLens-1M datasets.

Model	DBbook2014				Book-Crossing				MovieLens-1M			
	R@50 ↑	R@100 ↑	N@50 ↑	N@100 ↑	R@50 ↑	R@100 ↑	N@50 ↑	N@100 ↑	R@50 ↑	R@100 ↑	N@50 ↑	N@100 ↑
LLM-GRec	**0.4671**	**0.5741**	**0.2576**	**0.2798**	**0.1866**	**0.2501**	**0.0967**	**0.1107**	**0.4341**	**0.5806**	**0.4230**	**0.4680**
w/o LLM	0.4374	0.5532	0.2357	0.2610	0.1701	0.2340	0.0831	0.0977	0.4091	0.5551	0.4022	0.4462
w/o CF-guided	0.4668	0.5701	0.2556	0.2788	0.1850	0.2485	0.0962	0.1106	0.4315	0.5786	0.4200	0.4657
w/o GB	0.4432	0.5518	0.2392	0.2631	0.1856	0.2476	0.0942	0.1085	0.4323	0.5780	0.4195	0.4650
w/o Mask	0.4662	0.5722	0.2529	0.2763	0.1851	0.2482	0.0937	0.1082	0.4261	0.5734	0.4160	0.4610
w/o GBIAA	0.4565	0.5681	0.2470	0.2715	0.1683	0.2349	0.0831	0.0980	0.4058	0.5527	0.3995	0.4440
w/o KG	0.4544	0.5641	0.2490	0.2730	0.1686	0.2339	0.0820	0.0969	0.4093	0.5712	0.4102	0.4518

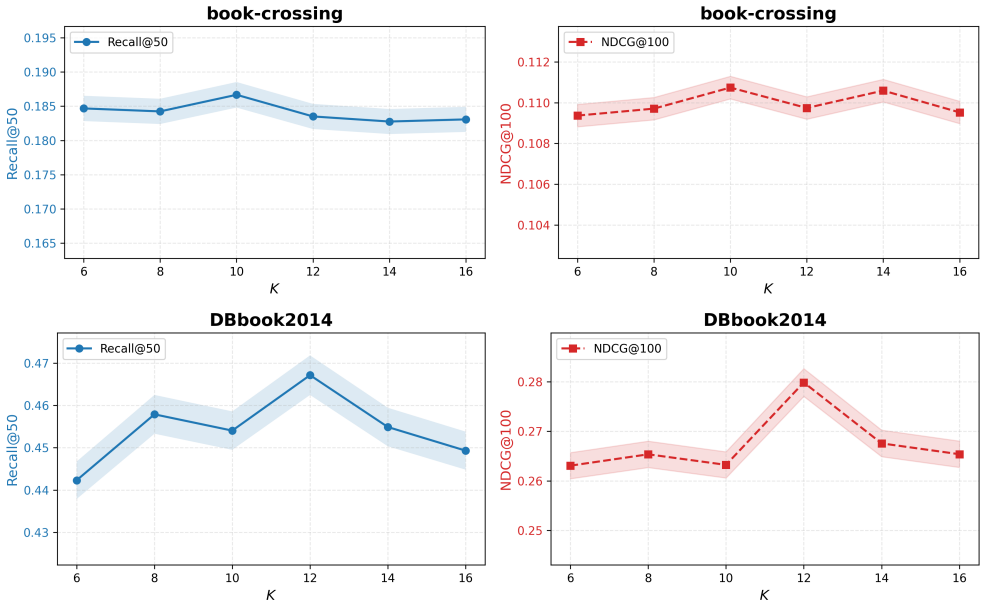


Fig. 2 Performance of granular-ball modeling under different neighborhood sizes on the Book-Crossing and DBbook2014 datasets.

consistently degrades results, confirming its effectiveness in suppressing semantic noise. Without interest anchor alignment (**w/o GBIAA**), performance declines, indicating that alignment improves representation consistency. Similarly, removing knowledge graphs (**w/o KG**) reduces recommendation accuracy, highlighting the contribution of structured item semantics. Overall, these results demonstrate that semantic interest generation, granular-ball modeling, anchor alignment, and noise suppression jointly enhance recommendation robustness and effectiveness.

4.4 Parameter Sensitivity Analysis (RQ3)

We analyze the sensitivity of LLM-GRec to the neighborhood size K in granular-ball modeling, as shown in Fig. 2. Performance initially improves and then declines as K increases, reaching its optimum at $K = 10$ on Book-Crossing and $K = 12$ on

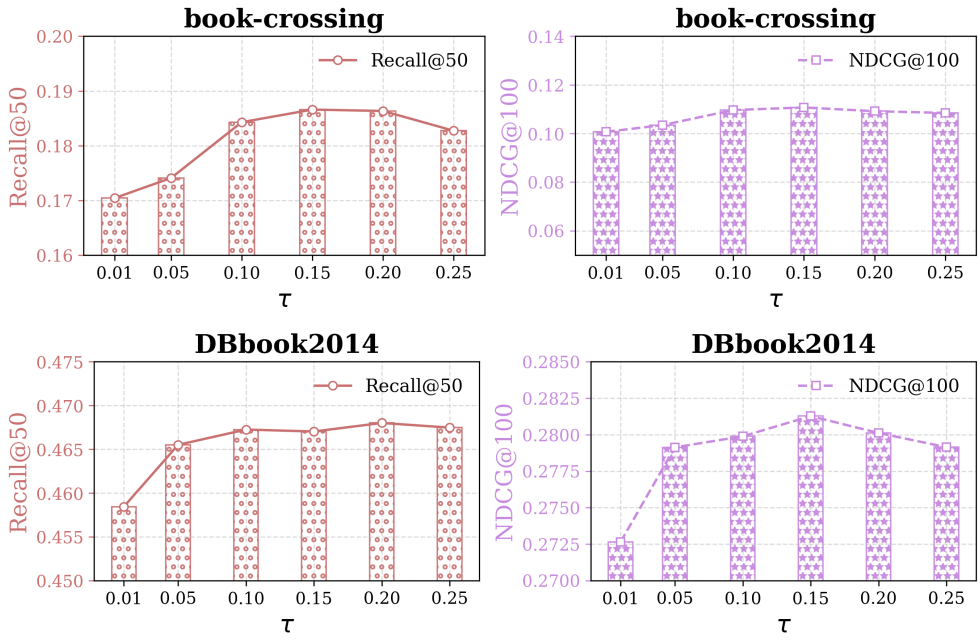


Fig. 3 Performance under different temperature values on the Book-Crossing and DBbook2014 datasets.

DBbook2014. Smaller K values provide insufficient semantic neighborhood information, while excessively large values introduce irrelevant neighbors and semantic noise. Appropriate K values therefore help construct stable granular-ball structures and improve recommendation performance.

We further analyze the impact of the temperature parameter τ on interest anchor alignment, as shown in Fig. 3. Performance initially improves and then declines as τ increases, with better results at relatively small values on both datasets. Small τ produces overly sharp anchor distributions, limiting semantic alignment, while large τ yields nearly uniform distributions, weakening anchor discrimination. An appropriate τ balances distribution smoothness and semantic consistency. Moreover, LLM-GRec maintains relatively stable performance within a certain temperature range, demonstrating the robustness of interest anchor alignment.

4.5 Robustness Analysis (RQ4)

The study shows limited performance degradation for **w/o Mask** under original data, suggesting insufficient noise to fully demonstrate the denoising mechanism. Therefore, we further evaluate robustness by randomly replacing LLM-generated interest phrases with noisy ones at ratios ranging from 0% to 50%, as shown in Fig. 4. As noise increases, all models experience performance degradation. However, LLM-GRec consistently outperforms **w/o Mask** on both Book-Crossing and DBbook2014, exhibiting

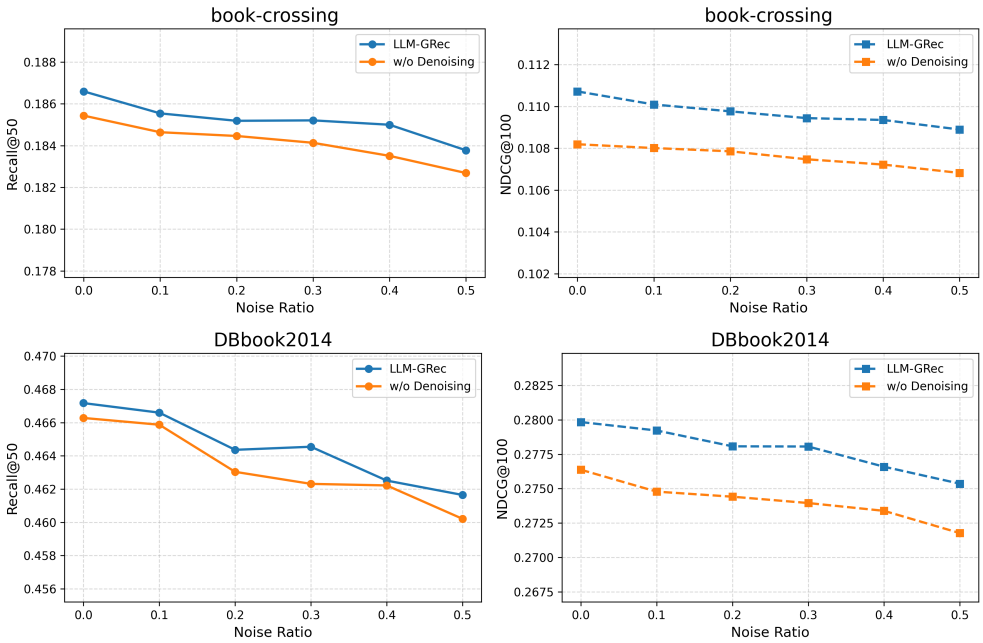


Fig. 4 Performance of LLM-GRec and its variant **w/o Mask** under different noise injection ratios on the Book-Crossing and DBbook2014 datasets.

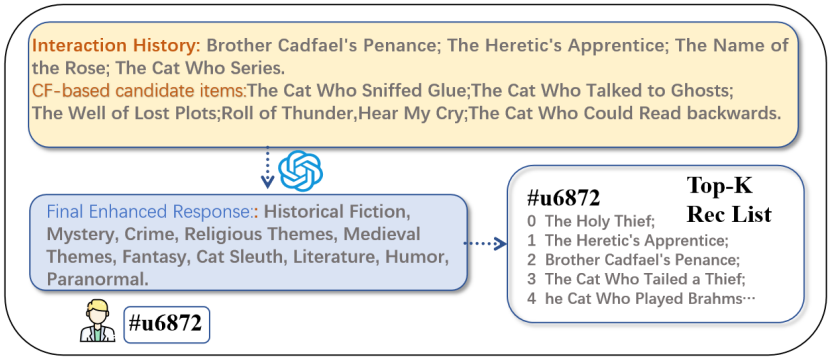


Fig. 5 Case study: Interpretability of user interest representations.

a slower performance decline. The advantage becomes more pronounced at higher noise ratios (40%–50%), demonstrating the effectiveness of quality-aware interest masking in suppressing semantic noise. Moreover, LLM-GRec maintains relatively stable performance under severe perturbations, indicating its ability to learn discriminative user representations despite unreliable interest information. These results confirm that the proposed denoising mechanism enhances recommendation robustness, particularly in noisy semantic environments.

4.6 Case Study (RQ5)

We present a case study to illustrate how LLM-GRec transforms historical interactions into semantic interests and recommendations, as shown in Fig. 5. For user #u6872, historical interactions mainly involve medieval detective novels and the "cat detective" series, including *Brother Cadfael's Penance*, *The Heretic's Apprentice*, and the *The Cat Who* series. These interactions indicate preferences for historical mysteries and distinctive thematic elements. To capture these preferences, LLM-GRec combines historical interactions with collaborative filtering candidate items and feeds them into the Large Language Model to generate semantic interests. The resulting topics, including *Historical Fiction*, *Mystery*, *Crime*, *Medieval Themes*, and *Cat Sleuth*, effectively summarize the user's reading preferences and reveal underlying semantic patterns. The final recommendations closely match these interests, including several works from the Brother Cadfael and The Cat Who series. This consistency demonstrates how semantic interest modeling connects historical behaviors with recommendation outcomes. Overall, this case illustrates the complete recommendation process of LLM-GRec, from historical interactions and collaborative signals to semantic interest generation and recommendation. It further highlights the interpretability of the learned interest representations, which provide meaningful semantic explanations for recommended items.

4.7 Efficiency Analysis (RQ6)

We evaluate the computational efficiency of LLM-GRec by comparing average training time per epoch on DBbook2014 and MovieLens-1M, as shown in Fig. 6. LLM-GRec requires more training time than LightGCN, KGIN, and CIKGRec on both datasets due to additional interest modeling, multi-view graph propagation, and denoising components. On DBbook2014, LightGCN and KGIN exhibit similar training costs, followed by CIKGRec and LLM-GRec. Training time increases for all methods on MovieLens-1M due to its larger scale, while their relative efficiency remains consistent. Despite the additional overhead, LLM-GRec maintains reasonable training efficiency and scalability. Notably, LLM-based interest generation is performed offline, and generated interests remain fixed during recommendation training. Therefore, the reported runtime excludes one-time LLM inference costs and primarily reflects recommendation model training, including interest graph modeling and denoising learning. This design avoids repeated LLM inference during training and limits additional computational overhead.

4.8 Sparsity Analysis (RQ7)

We evaluate LLM-GRec under different interaction sparsity levels by retaining training interactions at ratios $\rho \in \{0.2, 0.4, 0.6, 0.8, 1.0\}$. LightGCN, CIKGRec, and LLM-GRec are compared on DBbook2014, as shown in Fig. 7. As ρ increases, all models improve in Recall@50 and NDCG@100, demonstrating the importance of interaction information. LLM-GRec consistently outperforms both baselines, with more pronounced

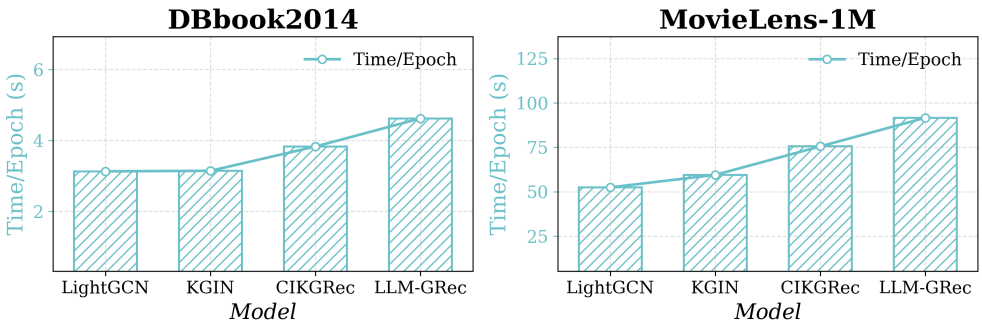


Fig. 6 Efficiency comparison of different models on DBbook2014 and MovieLens-1M.

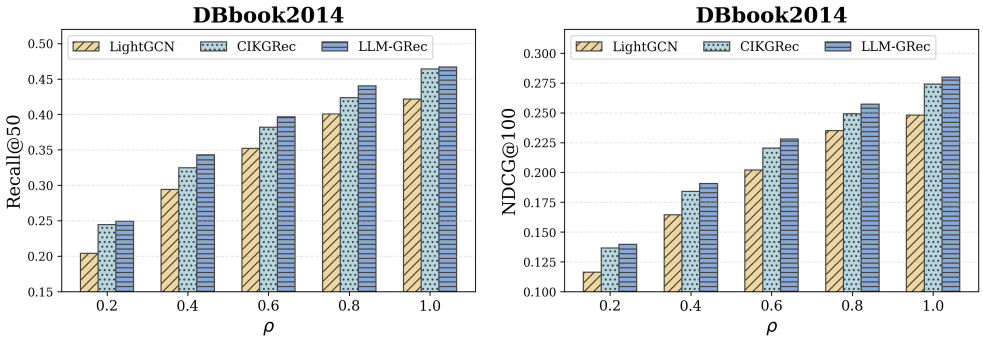


Fig. 7 Performance of different models under different interaction sparsity levels.

advantages under highly sparse conditions ($\rho = 0.2$ and $\rho = 0.4$). Although performance gaps narrow as interaction density increases, LLM-GRec maintains superior results. These findings indicate that semantic interest modeling and multi-granularity representations effectively alleviate data sparsity by providing complementary preference information. Furthermore, quality-aware denoising and interest anchor alignment enhance representation robustness under limited interactions, enabling more reliable recommendation performance in sparse scenarios.

5 Conclusion

In this paper, we propose LLM-GRec, a Large Language Model-enhanced recommendation framework addressing interaction sparsity and insufficient semantic information. By integrating LLM-generated user interests, multi-granularity interest modeling, and graph-based representation learning, LLM-GRec improves user preference representations and recommendation robustness. Specifically, collaborative-enhanced interest generation incorporates collaborative signals to enrich semantic interests.

Granular-ball computing organizes fine-grained interests into multi-granularity structures, providing stable interest representations under sparse and noisy conditions. Furthermore, dynamic mask-based denoising suppresses semantic noise during graph learning, while interest anchor alignment enhances semantic consistency between users and items. Experiments on multiple real-world datasets demonstrate that LLM-GRec outperforms existing methods on most evaluation metrics, with particularly notable improvements in sparse scenarios. These findings highlight the benefits of combining semantic augmentation with structured interest modeling. Future work will investigate finer-grained semantic representations and more efficient online inference mechanisms to improve recommendation effectiveness and practical deployment.

Author Contributions. J.S. Implemented the algorithms and conducted experiments, J.S. and X.Z. wrote the main manuscript text. All authors reviewed the manuscript.

Funding. This research did not receive funding.

Declarations

Ethics approval. The authors state that this research complies with ethical standards. This research does not involve either human participants or animals.

Competing interests. The authors declare no competing interests.

Availability of supporting data. The datasets used during the current study are available. Additionally, the datasets generated, model settings, and training processes are available from the corresponding author upon reasonable request.

References

Ai Q, Azizi V, Chen X, et al (2018) Learning heterogeneous knowledge base embeddings for explainable recommendation. Algorithms 11(9):137. https://doi.org/10.3390/a11090137

Bao K, Zhang J, Zhang Y, et al (2023) Tallrec: An effective and efficient tuning framework to align large language model with recommendation. In: Proceedings of the 17th ACM Conference on Recommender Systems, pp 1007–1014, https://doi.org/10.1145/3604915.3608857

Cai X, Huang C, Xia L, et al (2023) Lightgcl: Simple yet effective graph contrastive learning for recommendation. In: International Conference on Learning Representations (ICLR), URL https://arxiv.org/abs/2302.08191

Cen Y, Zhang J, Zou X, et al (2020) Controllable multi-interest framework for recommendation. In: Proceedings of the 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp 2942–2951, https://doi.org/10.1145/3394486.3403344

He X, Deng K, Wang X, et al (2020) Lightgcn: Simplifying and powering graph convolution network for recommendation. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, pp 639–648, https://doi.org/10.1145/3397271.3401063

Hu Z, Li Z, Jiao Z, et al (2025) Bridging the user-side knowledge gap in knowledge-aware recommendations with large language models. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp 11799–11807, https://doi.org/10.1609/aaai.v39i11.33284

Ji Z, Lee N, Frieske R, et al (2023) Survey of hallucination in natural language generation. ACM Computing Surveys 55(12):1–38. https://doi.org/10.1145/3571730

Kang WC, McAuley J (2018) Self-attentive sequential recommendation. In: 2018 IEEE International Conference on Data Mining (ICDM), pp 197–206, https://doi.org/10.1109/ICDM.2018.00035

Lai R, Chen R, Zhang C (2024) A survey on data-centric recommender systems. arXiv preprint arXiv:240117878 https://doi.org/2401.17878

Li C, Liu Z, Wu M, et al (2019) Multi-interest network with dynamic routing for recommendation at tmall. In: Proceedings of the 28th ACM International Conference on Information and Knowledge Management, pp 2615–2623, https://doi.org/10.1145/3357384.3357814

Liu Q, Wu X, Wang Y, et al (2024) Llm-esr: Large language models enhancement for long-tailed sequential recommendation. arXiv preprint arXiv:240520646 URL https://arxiv.org/abs/2405.20646

Ren X, Wei W, Xia L, et al (2024) Representation learning with large language models for recommendation. In: Proceedings of the ACM Web Conference 2024, pp 3464–3475, https://doi.org/10.1145/3589334.3645458

Tan Q, Zhang J, Yao J, et al (2021) Sparse-interest network for sequential recommendation. In: Proceedings of the 14th ACM International Conference on Web Search and Data Mining, pp 598–606, https://doi.org/10.1145/3437963.3441811

Wang W, Xu Y, Feng F, et al (2023) Diffusion recommender model. In: Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp 832–841, https://doi.org/10.1145/3539618.3591663

Wang X, He X, Cao Y, et al (2019a) Kgat: Knowledge graph attention network for recommendation. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp 950–958, https://doi.org/10.1145/3292500.3330989

Wang X, He X, Wang M, et al (2019b) Neural graph collaborative filtering. In: Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, pp 165–174, https://doi.org/10.1145/3331184.3331267

Wang X, Jin H, Zhang A, et al (2020) Disentangled graph collaborative filtering. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, pp 1001–1010, https://doi.org/10.1145/3397271.3401137

Wang X, Huang T, Wang D, et al (2021) Learning intents behind interactions with knowledge graph for recommendation. In: Proceedings of the Web Conference 2021, pp 878–887, https://doi.org/10.1145/3442381.3450133

Wei W, Ren X, Tang J, et al (2024) Llmrec: Large language models with graph augmentation for recommendation. In: Proceedings of the 17th ACM International Conference on Web Search and Data Mining, https://doi.org/10.1145/3616855.3635853

Wu J, Wang X, Feng F, et al (2021a) Self-supervised graph learning for recommendation. In: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp 726–735, https://doi.org/10.1145/3404835.3462862

Wu Z, Pan S, Chen F, et al (2021b) A comprehensive survey on graph neural networks. IEEE Transactions on Neural Networks and Learning Systems 32(1):4–24. https://doi.org/10.1109/TNNLS.2020.2978386

Xia S, Zhang Q, et al (2022) Granular-ball computing: Efficient, robust, and interpretable data mining. Information Sciences

Xia S, Wang G, Xu G, et al (2025) Gbgc: Efficient and adaptive graph coarsening via granular-ball computing. In: Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence, pp 3489–3497, https://doi.org/10.24963/ijcai.2025/388

Yang Y, Huang C, Xia L, et al (2022) Knowledge graph contrastive learning for recommendation. In: Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp 1434–1443

Yang Y, Huang C, Xia L, et al (2023) Knowledge graph self-supervised rationalization for recommendation. In: Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp 3046–3056, https://doi.org/10.1145/3580305.3599400

Yu J, Yin H, Xia X, et al (2022a) Are graph augmentations necessary? simple graph contrastive learning for recommendation. In: Proceedings of the 45th International

ACM SIGIR Conference on Research and Development in Information Retrieval, pp 1294–1303, https://doi.org/10.1145/3477495.3531937

Yu J, et al (2022b) Mirn: Multi-interest representation network for recommendation. In: Proceedings of the ACM Web Conference (WWW), https://doi.org/10.1145/3357384.3357814

Zhang F, Yuan NJ, Lian D, et al (2016) Collaborative knowledge base embedding for recommender systems. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp 353–362, https://doi.org/10.1145/2939672.2939673

Zou D, Wei W, Mao XL, et al (2022) Multi-level cross-view contrastive learning for knowledge-aware recommender system. In: Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp 1358–1368, https://doi.org/10.1145/3477495.3532025