

A Two-Stage Social Recommendation Framework with Dual-Graph Dynamic Denoising and Bidirectional Cross-View Contrastive Learning

Dong Wang, Haisha Liu*, Yulingheng Wang, Shanlin Liu, Shuai Li

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

*Correspondence Author, 1594195492@stu.cqut.edu.cn

Abstract: Social recommendation models leverage graph neural networks (GNNs) to jointly capture user-item interactions and user-user relationships, thereby improving recommendation performance. However, noise in both graphs may accumulate during message passing and disrupt representation learning. Existing methods often denoise only one graph or apply static and uniform criteria to heterogeneous relations, overlooking the distinct noise characteristics of interaction and social graphs. Moreover, structural and semantic discrepancies may remain between the denoised views, hindering cross-view alignment and information fusion. To overcome these challenges, we propose DGDC-SR, a framework that integrates dual-graph dynamic denoising with bidirectional cross-view contrastive learning. During pre-training, we design separate denoising strategies for the two graphs. For the interaction graph, an implicit gradient-based method adaptively estimates the contributions of observed interactions. For the social graph, a multi-signal dynamic soft denoising strategy estimates edge reliability and suppresses noise propagation. During fine-tuning, bidirectional cross-view contrastive learning aligns user representations across the denoised views, while an adaptive gating fusion mechanism effectively integrates user representations from both views. Extensive evaluations on three public datasets demonstrate the effectiveness of DGDC-SR, which consistently outperforms representative baseline methods.

Keywords: Graph Neural Networks, Contrastive Learning, Social Recommendation, Denoising.

1. Introduction

Social recommendation techniques seek to relieve the sparsity of user-item interaction data by exploiting the relationships that link users to one another. This integration facilitates preference modeling and has drawn substantial interest from the recommendation research community [1, 2]. Social ties not only signal shared interests but also carry social influence, and they allow models to harvest supplementary preference information from a user's social neighbors, ultimately enhancing recommendation performance. Early work in this direction relied chiefly on matrix factorization (MF) [3-5], where social information served as a complement to user modeling. In more recent years, graph neural networks (GNNs) [6-8] have come to dominate social recommendation. These architectures propagate information across the user-user social graph as well as the user-item interaction graph, employing high-order neighborhood aggregation to represent user preferences and social influence jointly. Consequently, the resulting user and item representations become richer, and collaborative signals together with social effects are captured more faithfully [9].

Despite these advantages, the effectiveness of graph-based social recommendation methods hinges to a large extent on the quality of both the interaction and social graphs. In real-world scenarios, these two graphs may contain different types of noisy and redundant relations. In the interaction graph, observed user behaviors may arise from misclicks, accidental browsing, exploratory actions, or temporary interests, and thus may not accurately reflect users' genuine preferences [10]. In addition, abnormal or manipulated interactions may result from shared accounts, automated activities, or fraudulent transactions [11]. By contrast, noise in

the social graph is mainly associated with unreliable user relationships, such as weak ties, inactive links, and preference-inconsistent connections [12, 13].

The noise in the two graphs therefore differs substantially in both source and semantics. Interaction-edge noise is behavior-oriented, and the reliability of an interaction edge mainly depends on whether the observed behavior genuinely reflects user preference and provides useful supervision for recommendation learning. Social-edge noise is relation-oriented, and the reliability of a social edge should be evaluated from multiple aspects, such as user preference consistency, local structural consistency, and propagation stability during message passing. Representative examples of these two types of graph noise are illustrated in Figure 1. Since GNNs repeatedly aggregate neighborhood information, unreliable edges in either graph may propagate noisy information to high-order neighborhoods and amplify its negative effect across propagation layers. This process can distort user and item representations and degrade recommendation performance. Therefore, effectively identifying and suppressing graph-specific noise in the interaction and social graphs is a key challenge in social recommendation.

The first of these operates by pruning or reweighting edges to refine the graph structure [13-16]. The second seeks to bolster the robustness of learned representations through self-supervised techniques, notably contrastive learning together with diffusion-based models [17-21]. More recently, several works [22-24] have investigated the use of large language models to assist with graph denoising and representation learning. Despite the promising outcomes these methods have yielded, a number of limitations still persist.

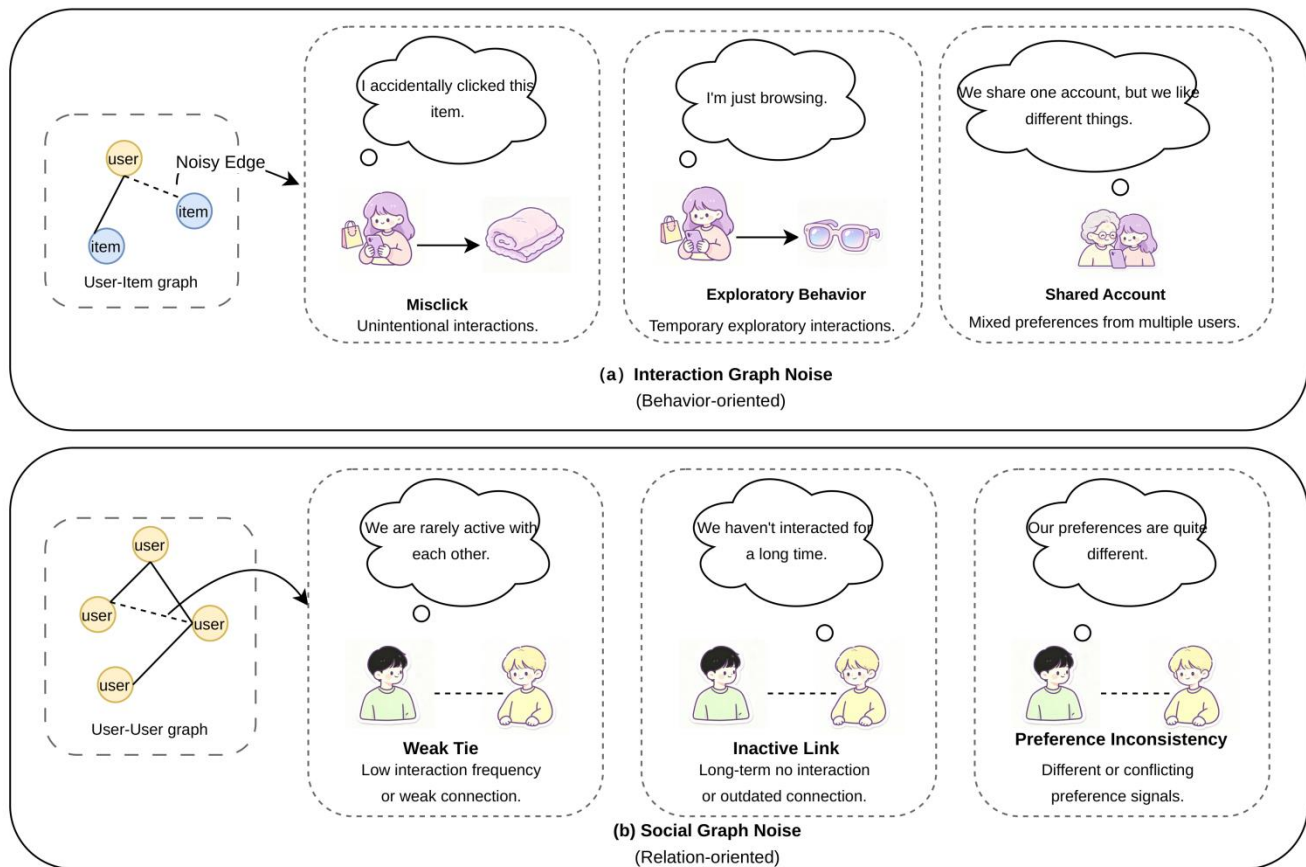


Figure 1: Different noise characteristics in graph-based social recommendation.

First, many existing methods estimate edge reliability using static strategies or a single signal. Such methods commonly assume that edge reliability can be evaluated using a fixed criterion throughout the training process. However, interaction and social edges differ substantially in their semantics and noise characteristics. A criterion designed to assess whether an interaction contributes to preference learning may not be suitable for evaluating whether a social relation is stable and preference-consistent. Therefore, applying a uniform reliability criterion to these heterogeneous relations may produce inaccurate estimates. Moreover, edge reliability may evolve as user and item representations are updated during training. For example, GDSRec [25] relies on a precomputed social-strength coefficient and does not update edge weights during representation learning. This design limits its ability to capture changes in social-edge reliability and identify unreliable relations.

Second, many existing studies focus primarily on denoising the social graph while leaving noise in the interaction graph largely unaddressed. For example, GBSR [26] suppresses irrelevant social information using the information bottleneck principle. GDMSR [27] refines social relations through a curriculum learning strategy. SSMD [28] refines the social graph by mining potential relations and applying cross-view alignment. Although these methods improve the quality of social information, unreliable user-item interactions may still be included in neighborhood aggregation and distort collaborative representations. Such interaction noise may further interfere with social information modeling and cross-view fusion. Therefore, denoising only the social graph is insufficient to control noise propagation in social recommendation. Conversely, directly applying the same

denoising rule to both graphs is also inappropriate because interaction edges and social edges require different reliability assessment criteria. These limitations motivate a joint yet graph-specific denoising strategy that simultaneously processes both graphs while adopting distinct denoising mechanisms for their different noise characteristics.

Third, contrastive learning has become a widely used technique for enhancing the robustness of learned representations. However, most existing methods [17, 18] do not establish an effective bidirectional contrastive mechanism between the interaction and social views. This limits cross-view alignment and weakens multi-view fusion. For example, GDSR [17] uses dual diffusion models to denoise both graphs and aligns representations before and after denoising. However, its contrastive learning remains confined to each individual view. Similarly, SSD-ICGA [18] generates augmented social graphs and applies contrastive learning within the social view. It does not explicitly align the interaction and social views. As a result, insufficient information exchange between the two views may weaken representation consistency and reduce overall model performance.

To address these limitations, we propose DGDC-SR, a two-stage social recommendation framework that integrates dual-graph dynamic denoising with bidirectional cross-view contrastive learning. In the pre-training stage, graph-specific dynamic denoising strategies are applied to the interaction and social graphs. For the interaction graph, edge importance is estimated from its contribution to representation learning. A multi-round score accumulation and reweighting strategy is then used to construct a denoised interaction graph. For the

social graph, edge reliability is estimated using gradient-based stability, cross-view preference consistency, and local structural consistency. The reliability scores are then dynamically updated through an exponential moving average (EMA) mechanism [29] to construct a weighted denoised social graph.

In the fine-tuning stage, user representations are learned from the two denoised graphs. A bidirectional cross-view contrastive learning mechanism is introduced to align user representations between the interaction and social views and reduce their semantic gap. A learnable gating mechanism is further used to adaptively balance interaction and social information. Finally, the model is optimized with a ranking objective. This design improves recommendation performance. Below, we present the key contributions of the present work.

- 1) We propose a two-stage framework that organizes graph denoising and cross-view representation learning into two successive stages, thereby reducing the influence of noise on subsequent representation learning.
- 2) We develop a dual-graph dynamic denoising method that applies graph-specific strategies to the interaction and social graphs, thereby improving graph reliability for subsequent representation learning.
- 3) We introduce a bidirectional cross-view contrastive learning mechanism to align user representations between the interaction and social views. A learnable gating mechanism is further used to adaptively balance the two information sources, thereby improving cross-view consistency and information fusion.

2. Related Work

2.1 Graph-based Social Recommendation

Social recommendation leverages social relationships among users to support user preference modeling, thereby alleviating the data sparsity issue in user-item interactions. Its theoretical foundations primarily rely on the assumptions of social influence [30] and homophily [31], which postulate that socially connected users tend to share similar preferences. Early social recommendation methods [5, 32–34] typically adopted matrix factorization frameworks. These methods incorporated social information into user-item preference learning through strategies such as collaborative decomposition, social regularization, and social-aware matrix factorization, aiming to improve recommendation performance. However, relying on shallow linear models, these methods fail to fully capture the high-order structural dependencies in social and interaction networks. Consequently, their performance is limited in complex scenarios.

Driven by the rapid advancements in Graph Neural Networks (GNNs), graph-based social recommendation has emerged as the mainstream modeling paradigm. These methods model user-user social relationships and user-item interactions as unified graph structures. Then, information propagation and neighborhood aggregation are performed through graph

representation learning, enabling the joint modeling of user preferences and social influence [35–37]. For example, DiffNet [38] enhances user representations by explicitly modeling the diffusion process of social influence. DiffNet++ [21] further introduces a hierarchical attention mechanism to integrate social influence diffusion with user interest propagation. In addition, GraphRec [9] incorporates rating information into the neighborhood aggregation process to improve the expressiveness of node representations. SocialLGN [39], built upon the LightGCN [40] framework, achieves unified modeling of user interactions and social relationships. Furthermore, methods such as SR-HGNN [41] and DGRec [42] extend the representation capability of graph-based social recommendation in complex scenarios from the perspectives of collaborative semantic modeling and dynamic interest modeling.

Overall, graph-based social recommendation methods have made significant progress in capturing high-order collaborative signals and modeling social influence. However, most existing approaches [9, 38] assume that the graph structure is accurate and reliable. They rarely consider the ubiquitous noise in user-item interaction graphs and user-user social graphs, such as redundant connections, weak ties, and unreliable interactions. Such noise is easily amplified during multi-layer propagation of GNNs, thereby degrading representation learning effectiveness and compromising model robustness.

2.2 Contrastive Learning for Recommendation

Contrastive learning is an important self-supervised learning paradigm. It has been widely applied in recommender systems to alleviate data sparsity and insufficient supervision by constructing positive and negative sample pairs and optimizing representation consistency [20]. Early studies mainly focused on graph representation learning. For example, DGI [43] captures graph structural information by contrasting node representations with global graph representations, whereas InfoGraph [44] enhances representation robustness by aligning substructures with graph-level representations. Building on these works, methods such as GraphCL [45] and MVGRL [46] construct contrastive views through various graph augmentation strategies, thereby improving the model's ability to handle structural perturbations.

In recent years, contrastive learning has been further introduced into recommender systems to improve the discriminability and robustness of user and item representations. For example, SimGCL [47] and XSimGCL [48] enhance representation stability by incorporating noise perturbations or cross-layer contrastive mechanisms. NCL [49] models the semantic relationships between users (or items) and their contexts through prototype-level contrastive objectives. CLRec [50] leverages contrastive learning to mitigate exposure bias, thereby improving recommendation performance.

Despite these advancements, most existing contrastive learning methods are limited to a single view or one-directional contrastive mechanisms, which hinders their ability to learn consistent representations across multiple information sources. This limitation is more pronounced in

social recommendation scenarios, where different views exhibit significant differences in both semantic space and structural characteristics. To bridge this gap, this work introduces a bidirectional cross-view contrastive learning mechanism to align user representations across different views, thereby enhancing the consistency and complementarity of multi-view representations.

3. Methodology

This section presents the proposed DGDC-SR framework, which comprises two stages: dual-graph dynamic denoising pre-training and cross-view contrastive fine-tuning. Figure 2 illustrates the overall framework of our method. In the

pre-training stage, the interaction graph is denoised using implicit gradient-based contribution estimation and degree-based penalization, whereas the social graph is denoised according to gradient-based stability, interaction-view preference similarity, and structural consistency. In the fine-tuning stage, the denoised interaction graph is used to learn interaction-view user and item representations, whereas the denoised social graph is used to learn social-view user representations. Bidirectional cross-view contrastive learning then aligns the user representations from the two views. Finally, an adaptive gating mechanism fuses the two user representations, and recommendation scores are computed from the fused user representations and the interaction-view item representations.

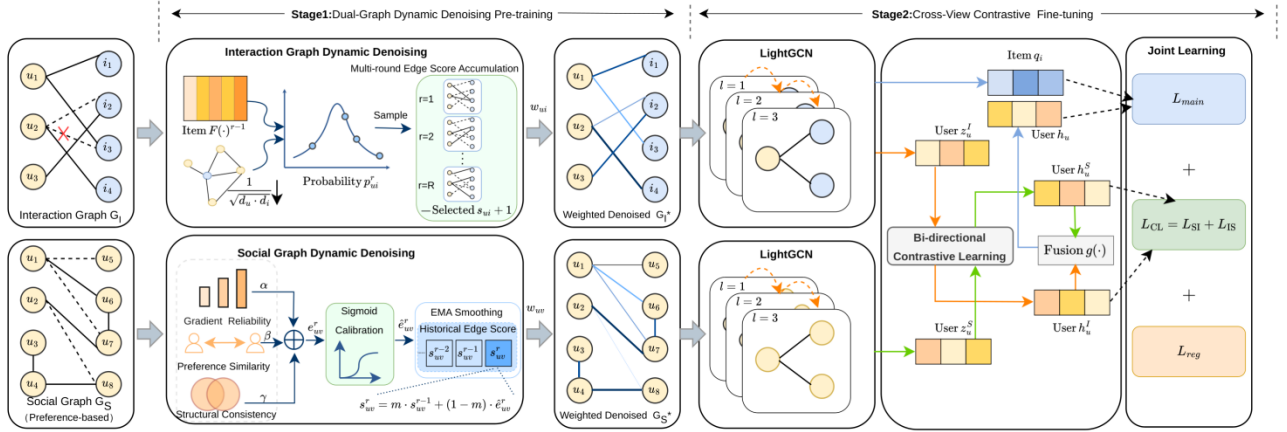


Figure 2: Overview of the proposed DGDC-SR framework.

3.1 Problem Formulation

Consider a recommendation setting that involves N users and M items, collected in the sets $U = \{u_1, u_2, \dots, u_N\}$ and $I = \{i_1, i_2, \dots, i_M\}$, respectively. Social ties between users are represented by a directed social graph $G_S = (U, E_S)$, where E_S is the set of directed social edges such as friendship or follow relationships. Meanwhile, the user-item interactions, including clicks, purchases, and ratings, are captured by a bipartite graph $G_I = (U, I, E_I)$, in which E_I denotes the interaction edges recording these observed behaviors.

Let $\mathbf{A}_S \in \mathbb{R}^{N \times N}$ denote the adjacency matrix of the original social graph, and let $\mathbf{R} \in \mathbb{R}^{N \times M}$ denote the user-item interaction matrix. Owing to the bipartite structure of the interaction graph, its adjacency matrix is defined as

$$\mathbf{A}_I = \begin{bmatrix} \mathbf{0}_{N \times N} & \mathbf{R} \\ \mathbf{R}^T & \mathbf{0}_{M \times M} \end{bmatrix} \in \mathbb{R}^{(N+M) \times (N+M)} \quad (1)$$

Let $\mathbf{e}_u^I \in \mathbb{R}^d$ and $\mathbf{e}_i^I \in \mathbb{R}^d$ denote the learnable base embeddings of user u and item i in the interaction view, respectively, and let $\mathbf{e}_u^S \in \mathbb{R}^d$ denote the learnable base embedding of user u in the social view. Based on these embeddings, multi-layer graph propagation is performed over the interaction and social graphs. We use $\mathbf{q}_u \in \mathbb{R}^d$ and $\mathbf{q}_i \in \mathbb{R}^d$ to denote the propagated interaction-view representations of user u and item i , respectively, and use $\mathbf{p}_u \in \mathbb{R}^d$ to denote the propagated social-view representation of user u . In other words, $\mathbf{e}$ denotes the learnable base embeddings, whereas $\mathbf{q}$ and $\mathbf{p}$ denote the propagated representations in the interaction and social views, respectively. The total number of graph propagation layers is denoted by l .

The objective of this work is to jointly learn robust and high-quality user and item representations from the interaction and social views, thereby improving top- K recommendation performance.

3.2 Dual-Graph Dynamic Denoising Pre-training

3.2.1 Interaction Graph Dynamic Denoising

In social recommendation, the user-item interaction graph $G_I = (U, I, E_I)$ may contain unreliable edges arising from inadvertent clicks, exploratory interactions, and other unintended behaviors. These edges may introduce irrelevant neighborhood information during graph propagation, thereby impairing user preference modeling and degrading top- K recommendation performance. To mitigate this issue, we develop a dynamic denoising strategy that combines implicit gradient-based contribution estimation with a degree-based penalty.

At the end of optimization round $r - 1$, we compute the gradient norm of the Bayesian Personalized Ranking (BPR) loss with respect to each interaction-view item embedding. This norm measures the local sensitivity of the ranking loss to the corresponding item representation and is therefore used as an implicit indicator of the item's contribution to the current optimization process. The contribution indicator of item i in denoising round r is defined as

$$F^{(r)}(i) = \|\nabla_{\mathbf{e}_i^I} \mathcal{L}_{\text{BPR}}^{(r-1)}\|_2 + \varepsilon, \quad (2)$$

where $\mathbf{e}_i^I$ denotes the interaction-view embedding of item i , and $\varepsilon > 0$ is a small constant introduced to avoid zero-valued

sampling scores. For $r = 1$, $\mathcal{L}_{\text{BPR}}^{(0)}$ is obtained from the initial optimization round before the first edge-sampling operation.

Because $F^{(r)}(i)$ is an item-level contribution signal, it is combined with the degrees of the two endpoints to derive an edge-level sampling probability. Specifically, for interaction edge $(u, i) \in E_I$, the sampling probability is defined as

$$p_{ui}^{(r)} = \frac{\frac{F^{(r)}(i)}{\sqrt{d_u d_i}}}{\sum_{(x,j) \in E_I} \frac{F^{(r)}(j)}{\sqrt{d_x d_j}}}, \quad (3)$$

where d_u and d_i represent the degrees of user u and item i , respectively, as measured in the original interaction graph. The degree-aware component mitigates the influence of highly active users and popular items when sampling interaction edges, while the normalization ensures that $\{p_{ui}^{(r)}\}_{(u,i) \in E_I}$ forms a valid probability distribution.

In each denoising round, the model samples $[|E_I|(1-\rho)]$ edges without replacement from E_I according to the sampling distribution, where $\rho \in [0,1)$ denotes the interaction-edge pruning ratio. Let $E_I^{(r)}$ denote the sampled edge set in round r . The accumulated sampling count and the corresponding normalized frequency of each interaction edge are computed as

$$\begin{aligned} s_{ui}^{(r)} &= s_{ui}^{(r-1)} + \mathbb{I}[(u, i) \in E_I^{(r)}], & s_{ui}^{(0)} &= 0, \\ \bar{s}_{ui} &= \frac{s_{ui}^{(R_d)}}{R_d}, & \bar{s}_{ui} &\in [0,1], \end{aligned} \quad (4)$$

where $\mathbb{I}[\cdot]$ denotes the indicator function and R_d denotes the total number of denoising rounds. Multi-round accumulation reduces the influence of sampling fluctuations in any individual round and provides a more stable estimate of interaction-edge importance.

After R_d denoising rounds, the $[|E_I|(1-\rho)]$ edges with the largest normalized sampling frequencies are retained. The pruning ratio ρ governs the count of edges sampled per denoising iteration as well as those preserved within the final denoised interaction graph. Let E_I^* denote the retained edge set. Each retained interaction edge is assigned a continuous weight:

$$w_{ui} = \varepsilon_w + (1 - \varepsilon_w) \bar{s}_{ui}^\gamma, \quad (u, i) \in E_I^*, \quad (5)$$

where $\varepsilon_w \in (0,1)$ denotes the minimum interaction-edge weight, and $\gamma > 0$ controls the nonlinear mapping from normalized sampling frequencies to edge weights.

Let $\mathbf{R}^* \in \mathbb{R}^{N \times M}$ denote the weighted interaction matrix, where $R_{ui}^* = w_{ui}$ for $(u, i) \in E_I^*$ and $R_{ui}^* = 0$ otherwise. The resulting denoised interaction adjacency matrix is constructed as

$$\mathbf{A}_I^* = \begin{bmatrix} \mathbf{0}_{N \times N} & \mathbf{R}^* \\ (\mathbf{R}^*)^\top & \mathbf{0}_{M \times M} \end{bmatrix}. \quad (6)$$

Finally, symmetric normalization is applied to obtain the interaction propagation matrix:

$$\hat{\mathbf{A}}_I^* = (\mathbf{D}_I^*)^{-\frac{1}{2}} \mathbf{A}_I^* (\mathbf{D}_I^*)^{-\frac{1}{2}}, \quad (7)$$

where $\mathbf{D}_I^*$ is the diagonal degree matrix of $\mathbf{A}_I^*$. The resulting

propagation matrix $\hat{\mathbf{A}}_I^*$ is used for interaction-view representation propagation in the subsequent fine-tuning stage.

3.2.2 Social Graph Dynamic Denoising

The user-user social graph may contain unreliable relations, including weak ties, inactive links, and preference-inconsistent connections. During multi-layer graph propagation, these relations may introduce irrelevant neighborhood information and degrade social-view representation learning. To address this issue, we develop a dynamic soft-denoising strategy that estimates social-edge reliability using three complementary signals: gradient-based stability, cross-view preference consistency, and local structural consistency. Rather than directly removing low-confidence social edges, the proposed strategy continuously reweights them, thereby suppressing unreliable information while preserving the connectivity of the social graph.

Let $(u, v) \in E_s$ denote a directed social edge from user u to user v . We first define a gradient-based signal to evaluate the stability of the users connected by each social edge. In denoising round r , the social smoothing loss is defined over the base social graph as

$$\mathcal{L}_{\text{soc}}^{(r)} = \frac{1}{N} \left\| \mathbf{P}^{(r)} - \hat{\mathbf{A}}_S^{(0)} \mathbf{P}^{(r)} \right\|_F^2, \quad (8)$$

where $\hat{\mathbf{A}}_S^{(0)}$ denotes the normalized propagation matrix of the base social graph, and $\mathbf{P}^{(r)} = [\mathbf{p}_1^{(r)\top}; \dots; \mathbf{p}_N^{(r)\top}] \in \mathbb{R}^{N \times d}$ denotes the social-view representation matrix in round r . The smoothing loss measures the discrepancy between each user's representation and its neighborhood-aggregated representation on the base social graph.

The gradient intensity of user u in denoising round r is defined as

$$\mathbf{G}^{(r)}(u) = \left\| \nabla_{\mathbf{e}_u^S} \mathcal{L}_{\text{soc}}^{(r)} \right\|_2 + \varepsilon, \quad (9)$$

where $\mathbf{e}_u^S$ denotes the underlying learnable social-view embedding of user u . A larger value of $\mathbf{G}^{(r)}(u)$ indicates that the corresponding user representation is more sensitive to the social smoothing objective.

To make the gradient responses comparable across users, the gradient intensities are standardized and transformed into a probability distribution using a temperature-scaled softmax. The gradient-based reliability of social edge (u, v) is then computed as

$$\begin{aligned} \tilde{G}^{(r)}(u) &= \frac{G^{(r)}(u) - \mu_G^{(r)}}{\zeta_G^{(r)} + \varepsilon}, \\ \pi_u^{(r)} &= \frac{\exp(\tilde{G}^{(r)}(u)/\tau_g)}{\sum_{x \in U} \exp(\tilde{G}^{(r)}(x)/\tau_g)}, \\ \psi_{uv}^{(r)} &= \frac{\pi_u^{(r)} + \pi_v^{(r)}}{2}, \\ e_{uv}^{g,(r)} &= 1 - \frac{\psi_{uv}^{(r)} - \psi_{\min}^{(r)}}{\psi_{\max}^{(r)} - \psi_{\min}^{(r)} + \varepsilon}. \end{aligned} \quad (10)$$

Here, $\mu_G^{(r)}$ and $\zeta_G^{(r)}$ denote the mean and standard deviation of the user gradient intensities in round r , respectively, and $\tau_g > 0$ is the gradient temperature coefficient. Moreover,

$\psi_{min}^{(r)} = \min_{(x,y) \in E_S} \psi_{xy}^{(r)}$ and $\psi_{max}^{(r)} = \max_{(x,y) \in E_S} \psi_{xy}^{(r)}$. Accordingly, an edge connecting users with relatively large gradient responses is assigned a lower gradient-based reliability score, indicating lower stability under social propagation.

The second signal measures preference consistency between connected users in the interaction view. Let $q_u^{(r)}$ and $q_v^{(r)}$ denote the interaction-view representations of users u and v in round r , respectively. Their preference consistency is defined as

$$e_{uv}^{p,(r)} = \frac{1 + \cos(q_u^{(r)}, q_v^{(r)})}{2}, \quad (11)$$

This transformation rescales cosine similarity from $[-1,1]$ to $[0,1]$, with a larger value indicating greater preference consistency between the two users.

The third signal captures local structural consistency. For social edge (u, v) , the Jaccard similarity [51] between the outgoing neighbor sets of users u and v is defined as

$$e_{uv}^s = \frac{|N^+(u) \cap N^+(v)|}{|N^+(u) \cup N^+(v)|}, \quad (12)$$

Where $N^+(u)$ represents the first-order outgoing neighbor set of user u , and $N^+(v)$ corresponds to that of user v . A larger value of e_{uv}^s indicates greater overlap between their local social neighborhoods.

The three complementary signals are combined to define the overall reliability of social edge (u, v) in denoising round r :

$$e_{uv}^{(r)} = \alpha e_{uv}^{g,(r)} + \beta e_{uv}^{p,(r)} + \eta e_{uv}^s, \quad (13)$$

where α , β , and η are positive trade-off coefficients satisfying $\alpha + \beta + \eta = 1$. They control the contributions of gradient-based stability, cross-view preference consistency, and local structural consistency, respectively.

To improve the separation between high- and low-reliability edges, a centered sigmoid function is applied to the fused reliability score. The calibrated score is then accumulated using an exponential moving average (EMA) mechanism [29]:

$$s_{uv}^{(r)} = m s_{uv}^{(r-1)} + (1 - m) \sigma(\kappa(e_{uv}^{(r)} - \delta)), \quad s_{uv}^{(0)} = 0, \quad (14)$$

where $m \in [0,1]$ denotes the EMA momentum coefficient, $\sigma(\cdot)$ denotes the sigmoid function, $\kappa > 0$ controls its steepness, and $\delta \in (0,1)$ specifies its center. Compared with confidence estimation based on a single round, the EMA update reduces the influence of transient fluctuations and provides more stable social-edge confidence estimates.

After R_d denoising rounds, the accumulated confidence $s_{uv}^{(R_d)}$ is used to determine the final weight of social edge (u, v) :

$$w_{uv} = \varepsilon_s + (1 - \varepsilon_s) s_{uv}^{(R_d)}, \quad (u, v) \in E_S \quad (15)$$

where $\varepsilon_s \in (0,1)$ denotes the minimum social-edge weight. This lower bound prevents low-confidence edges from being assigned zero weight and helps preserve the connectivity of the social graph.

Let $A_S^* \in \mathbb{R}^{N \times N}$ denote the weighted social adjacency matrix, where $A_{S,uv}^* = w_{uv}$ for $(u, v) \in E_S$ and $A_{S,uv}^* = 0$ otherwise. Because the social graph is directed, self-loops and row normalization are applied to obtain the final denoised social propagation matrix:

$$\hat{A}_S^* = (D_S^*)^{-1} (A_S^* + I_N), \quad (16)$$

where D_S^* is the diagonal out-degree matrix of $A_S^* + I_N$, and I_N is the $N \times N$ identity matrix. The resulting propagation matrix $\hat{A}_S^*$ is used for social-view representation propagation in the subsequent fine-tuning stage.

3.3 Cross-View Contrastive Fine-tuning

During the fine-tuning stage, representation propagation is performed separately on the denoised interaction graph and social graph to obtain user representations from the interaction view and the social view. Subsequently, we introduce a bidirectional cross-view contrastive learning mechanism, which treats the representations of the same user across different views as positive pairs. Cross-view consistency constraints are further employed to align the representation spaces of the two views, thereby reducing the semantic discrepancy between the interaction graph and the social graph. On this basis, a gating fusion mechanism is designed to adaptively integrate interaction behavior information and social relationship information. It generates the final user representations, which are then used for ranking optimization, thereby improving Top- K recommendation performance and generalization ability.

3.3.1 Dual-view Representation Propagation on Denoised Graphs

After the dynamic denoising process in Stage 1, the model obtains the denoised interaction graph and denoised social graph. To fully exploit the effective structural information in the two graph views, LightGCN [40] is adopted to perform representation propagation on the denoised interaction graph and social graph during the fine-tuning stage. Here, $\hat{A}_I^*$ and $\hat{A}_S^*$ denote the normalized propagation matrices of the denoised interaction graph and denoised social graph, respectively.

For the interaction view, let $E_I^{(k)}$ denote the joint user-item representation matrix after the k -th propagation layer, where $E_I^{(0)}$ is the initial joint user-item embedding matrix. The propagation process is defined as

$$E_I^{(k)} = \hat{A}_I^* E_I^{(k-1)}. \quad (17)$$

After l propagation layers, the representations from all layers are averaged to obtain the final interaction-view representation matrix:

$$E_I = \frac{1}{l+1} \sum_{k=0}^l E_I^{(k)}. \quad (18)$$

The matrix E_I is then decomposed into the interaction-view user representations q_u and item representations q_i .

For the social view, let $E_S^{(k)}$ denote the user representation matrix after the k -th propagation layer, where $E_S^{(0)}$ is the initial social-view user embedding matrix. The propagation process is defined as

$$E_S^{(k)} = \hat{A}_S^* E_S^{(k-1)}. \quad (19)$$

Similarly, embeddings generated across all propagation layers are aggregated by averaging to derive the ultimate social-view user embedding p_u .

3.3.2 Bidirectional Cross-View Contrastive Learning

Due to the different sources of the interaction graph and the social graph, significant differences exist between the two views in terms of structural characteristics and semantic information. As a result, the representations learned for the same user from different views may exhibit semantic discrepancies. To alleviate the inconsistency between cross-view representations, a bidirectional cross-view contrastive learning mechanism is introduced during the fine-tuning stage. The representations of the same user from the interaction view and the social view are treated as positive pairs, while historical representations stored in the cross-batch negative queues are used as negative samples. Cross-view contrastive constraints are then applied to enhance the consistency between the two views.

First, the user representations learned from the denoised interaction graph and denoised social graph are projected into a shared contrastive learning space:

$$z_u^S = f_S(p_u), \quad z_u^I = f_I(q_u), \quad (20)$$

where $f_S(\cdot)$ and $f_I(\cdot)$ denote the projection functions for the social view and interaction view, respectively.

Next, the projected representations are normalized using L_2 normalization:

$$\hat{z}_u^S = \frac{z_u^S}{\|z_u^S\|_2}, \quad \hat{z}_u^I = \frac{z_u^I}{\|z_u^I\|_2}. \quad (21)$$

Let Q_I and Q_S denote the cross-batch negative queues containing normalized historical projected representations from the interaction view and social view, respectively. In the social-to-interaction direction, $\hat{z}_u^S$ is treated as the anchor, and $\hat{z}_u^I$ is regarded as the positive sample. Historical interaction-view representations stored in Q_I are used as negative samples. The corresponding contrastive loss is defined as

$$\mathcal{L}_{S \rightarrow I} = -\frac{1}{|U_b|} \sum_{u \in U_b} \log \frac{\exp(\text{sim}(\hat{z}_u^S, \hat{z}_u^I) / \tau_{cl})}{\exp(\text{sim}(\hat{z}_u^S, \hat{z}_u^I) / \tau_{cl}) + \sum_{z \in Q_I} \exp(\text{sim}(\hat{z}_u^S, z) / \tau_{cl})}. \quad (22)$$

Here, U_b denotes the user set in the current batch, $\tau_{cl} > 0$ is the contrastive learning temperature, and $\text{sim}(\cdot, \cdot)$ denotes the cosine similarity function. This objective pulls the representations of the same user from the two views closer while pushing the social-view anchor away from the historical interaction-view representations stored in the negative queue, thereby enhancing the semantic consistency between the two views.

Similarly, in the interaction-to-social direction, $\hat{z}_u^I$ is treated as the anchor, and $\hat{z}_u^S$ is regarded as the positive sample. Historical social-view representations stored in Q_S are used as negative samples. The reverse contrastive loss is defined as

$$\mathcal{L}_{I \rightarrow S} = -\frac{1}{|U_b|} \sum_{u \in U_b} \log \frac{\exp(\text{sim}(\hat{z}_u^I, \hat{z}_u^S) / \tau_{cl})}{\exp(\text{sim}(\hat{z}_u^I, \hat{z}_u^S) / \tau_{cl}) + \sum_{z \in Q_S} \exp(\text{sim}(\hat{z}_u^I, z) / \tau_{cl})}. \quad (23)$$

This loss further constrains the interaction-view representation of each user to remain consistent with its corresponding social-view representation while distinguishing it from the historical social-view representations stored in the negative queue, thereby improving the consistency and stability of the dual-view representations.

By jointly considering the cross-view alignment constraints in both directions, the final bidirectional cross-view contrastive learning objective is obtained by equally averaging the two contrastive losses:

$$\mathcal{L}_{cl} = \frac{1}{2} (\mathcal{L}_{S \rightarrow I} + \mathcal{L}_{I \rightarrow S}). \quad (24)$$

Through bidirectional cross-view contrastive constraints, the model jointly optimizes the alignment from the social view to the interaction view and from the interaction view to the social view. As a result, semantic discrepancies caused by view differences are alleviated, and the robustness and stability of user representations are improved.

3.3.3 Adaptive Gating Fusion Mechanism

Although bidirectional cross-view contrastive learning enhances the consistency between the interaction-view representation q_u and the social-view representation p_u , the importance of the two types of view information differs across users. For users with sufficient interaction records, historical behaviors can more directly reflect their preference characteristics. In contrast, for users with sparse interactions, social relationships may provide more effective supplementary information. Therefore, an adaptive gating fusion mechanism is designed to dynamically integrate user representations from the interaction view and the social view.

First, the user representations from the two views are concatenated and then fed into a gating layer to generate a user-level fusion weight:

$$g_u = \sigma(W_g [p_u \parallel q_u] + b_g), \quad (25)$$

where $[\cdot \parallel \cdot]$ denotes vector concatenation, $W_g \in \mathbb{R}^{1 \times 2d}$ and $b_g \in \mathbb{R}$ denote the weight matrix and bias of the gating layer, respectively, $\sigma(\cdot)$ denotes the sigmoid activation function, and $g_u \in (0, 1)$.

Next, the two view-specific representations are adaptively fused according to the gating weight to obtain the final representation of user u :

$$h_u = g_u q_u + (1 - g_u) p_u. \quad (26)$$

According to Eq. (26), when g_u is relatively large, the model relies more on the interaction-view representation q_u . When g_u is relatively small, the model relies more on the social-view representation p_u . Therefore, the gating fusion mechanism can adaptively balance interaction behavior information and social relationship information according to user characteristics, thereby generating more comprehensive and robust user representations.

3.4 Joint Optimization Objective

Different optimization objectives are adopted in the two training stages. During the pre-training stage, representation learning is performed only on the user–item interaction view, and the Bayesian Personalized Ranking (BPR) loss [52] is employed to optimize user and item representations. During the fine-tuning stage, bidirectional cross-view contrastive learning and the gating fusion mechanism are further incorporated. The ranking loss and contrastive learning loss are jointly optimized to enhance the consistency between different view representations, thereby improving recommendation performance and model robustness.

During the pre-training stage, let q_u denote the interaction-view representation of user u , and let q_i and q_j denote the representations of items i and j , respectively, where i represents a positive-feedback item and j represents a negative-feedback item. The predicted scores of user u for the positive and negative items are defined as

$$\hat{y}_{ui}^{pre} = q_u^\top q_i, \quad \hat{y}_{uj}^{pre} = q_u^\top q_j. \quad (27)$$

Based on the predicted scores, the ranking optimization objective in the pre-training stage is defined as

$$\mathcal{L}_{pre} = \frac{1}{|B|} \sum_{(u,i,j) \in B} \text{softplus}(\hat{y}_{uj}^{pre} - \hat{y}_{ui}^{pre}) + \lambda_{reg} \mathcal{L}_{reg}, \quad (28)$$

where B denotes the set of training triplets in a batch, $|B|$ denotes the batch size, $\text{softplus}(x) = \ln(1 + e^x)$, $\mathcal{L}_{reg}$ denotes the L_2 regularization term, and λ_{reg} is the regularization coefficient.

During the fine-tuning stage, the interaction-view representation q_u and the social-view representation p_u are first learned from the denoised interaction graph and denoised social graph, respectively. The cross-view-aligned representations are then adaptively fused to obtain the final user representation h_u . Since item nodes do not participate in social-view propagation, the item representation remains the interaction-view representation q_i . Accordingly, the predicted score of user u for positive-feedback item i is defined as

$$\hat{y}_{ui}^{fine} = h_u^\top q_i. \quad (29)$$

To improve the discriminative ability of the model for hard negative samples, a semi-hard negative mining strategy is further introduced during the fine-tuning stage. For each user u , K_{neg} candidate negative samples are first randomly sampled to construct the candidate set C_u . The predicted scores of the candidate negative samples are calculated as

$$\hat{y}_{uj^{(t)}}^{fine} = h_u^\top q_{j^{(t)}}, \quad j^{(t)} \in C_u. \quad (30)$$

Subsequently, the item with the highest predicted score in the candidate set is selected as the semi-hard negative sample:

$$j_u^{hard} = \arg \max_{j \in C_u} \hat{y}_{uj}^{fine}. \quad (31)$$

To balance training stability and hard negative mining effectiveness, the semi-hard negative samples are further combined with randomly sampled negative samples to obtain the final negative sample $\tilde{j}$ for ranking optimization. Therefore, the joint optimization objective in the fine-tuning

stage is defined as

$$\mathcal{L}_{fine} = \frac{1}{|B|} \sum_{(u,i,j) \in B} \text{softplus}(\hat{y}_{uj}^{fine} - \hat{y}_{ui}^{fine}) + \lambda_{cl} \mathcal{L}_{cl} + \lambda_{reg} \mathcal{L}_{reg}, \quad (32)$$

where $\mathcal{L}_{cl}$ denotes the bidirectional cross-view contrastive learning loss, λ_{cl} denotes its weight coefficient, and λ_{reg} denotes the regularization coefficient.

Through the joint optimization strategy above, the pre-training stage mainly learns stable user–item preference structures from interaction ranking signals. The fine-tuning stage further refines user preference representations by incorporating social auxiliary information, cross-view consistency constraints, and semi-hard negative mining, thereby improving ranking performance and recommendation robustness.

4. Experiments

This section presents empirical studies to assess the recommendation quality achieved by our proposed DGDC-SR framework.

4.1 Datasets

Table1: Statistics of the datasets.

Dataset	LastFM	Ciao	Douban
# of Users	1892	7375	2848
# of Items	17632	105114	39586
# of User–Item Interactions	92834	284086	894887
User–Item Density	0.278%	0.037%	0.793%
# of User–User Relations	25434	57544	35770
User–User Density	0.711%	0.106%	0.441%

To assess the performance of the proposed approach, we conduct experiments using three publicly accessible social recommendation benchmarks, namely LastFM, Ciao, and Douban. The statistical characteristics of these datasets are reported in Table 1, followed by a brief description of each dataset. The datasets are publicly available at <https://grouplens.org/datasets/hetrec-2011/> (LastFM), <https://www.cse.msu.edu/~tangjili/datasetcode/truststudy.htm> (Ciao), and <https://www.douban.com> (Douban).

LastFM. LastFM records implicit user–item interactions involving music consumption, along with social connections between users. It is commonly adopted in recommendation research for examining users’ preference patterns and the contribution of social information.

Ciao. Ciao is a representative benchmark for evaluating social recommendation methods. The dataset consists of user ratings for items on a 1–5 scale, as well as social links that describe relationships between users.

Douban. Douban is derived from an online social platform and covers several content categories, such as movies, books, and music. It includes both user–item interaction records and user–user social connections. Users provide explicit item ratings ranging from 1 to 5.

4.2 Baselines

To validate the effectiveness of DGDC-SR, we compare it

with nine representative recommendation methods, including traditional social recommendation models, graph neural network-based methods, and contrastive learning-based

methods. Table 2 summarizes the key characteristics of these baseline methods, which are briefly introduced as follows.

Table 2: Comparison of the key characteristics of DGDC-SR and baseline methods.

Method	Interaction Graph Denoising	Social Graph Denoising	Distinct Dynamic Denoising Strategies for Dual Graphs	Contrastive Learning	Bidirectional Cross-View Alignment
SBPR	×	×	×	×	×
DiffNet	×	×	×	×	×
LightGCN	×	×	×	×	×
MGDCF	×	×	×	×	×
SimGCL	×	×	×	✓	×
ESRF	×	✓	×	×	×
DSL	×	✓	×	✓	×
SocialLGN	×	×	×	×	×
CLDS	✓	✓	×	✓	×
DGDC-SR	✓	✓	✓	✓	✓

1) SBPR [53]: SBPR incorporates social relationship information into the BPR [52] framework and is one of the earliest recommendation models to integrate social information into ranking optimization.

2) DiffNet [38]: DiffNet models user–item interactions and user–user social relations using graph neural networks. It learns user representations through multi-layer information diffusion to capture the propagation of social influence and improve recommendation performance.

3) LightGCN [40]: LightGCN is a lightweight graph neural network-based recommendation model. It simplifies graph convolution by removing feature transformation and nonlinear activation, thereby enabling effective user and item representation learning.

4) MGDCF [54]: MGDCF models user–item interactions as a Markov graph diffusion process and learns representations by balancing neighborhood smoothing and self-information preservation.

5) SimGCL [47]: SimGCL is a contrastive learning-based recommendation method that replaces conventional structure-perturbation-based graph augmentation with random embedding perturbations, thereby reducing semantic distortion and improving representation learning.

6) ESRF [55]: ESRF constructs an adversarial learning framework based on graph neural networks. It generates and denoises latent social relations and further incorporates an attention mechanism for information propagation.

7) DSL [20]: DSL alleviates noise in social graphs through cross-view denoising contrastive learning and optimizes the recommendation model within a multi-task learning framework.

8) SocialLGN [39]: SocialLGN performs lightweight graph convolution on both the user–item interaction graph and the user–user social graph. User representations from the two views are then fused for recommendation.

9) CLDS [56]: CLDS performs pre-training denoising by constructing a social preference network and further improves representation learning through self-supervised contrastive learning. It also employs a dual-stage graph neural network to

alleviate noise in the interaction graph.

4.3 Evaluation Metrics

We employ three prevalent evaluation metrics, namely Precision@ K , Recall@ K and NDCG@ K , to measure the performance of our social recommendation model. We configure the truncation length as $K \in \{10, 20\}$ in our experiments to assess model performance with varied recommendation list sizes. For each user, all candidate items are ranked according to their predicted scores. Items that appear in the training set are excluded, and the top- K remaining items are selected as the recommendation list. These items are then compared with the ground-truth items in the test set to compute the evaluation metrics. Precision@ K quantifies the fraction of recommended items that are relevant, whereas Recall@ K indicates the extent to which relevant items are retrieved within the top- K list. NDCG@ K assesses the quality of the resulting ranking by giving greater importance to relevant items placed at higher ranks. For each metric, the value is first calculated for individual test users and subsequently averaged across all users in the test set.

4.4 Parameter Settings

To ensure a fair comparison, the common parameter settings are kept as consistent as possible with those used by CLDS. The model is trained using the Adam optimizer. The initial learning rate is set to 1×10^{-3} , the embedding dimension is set to 64, the number of graph propagation layers is set to 3, and the maximum number of training epochs is set to 10,000. An early-stopping strategy is adopted to reduce overfitting. Several key hyperparameters related to graph denoising and contrastive learning are tuned according to the characteristics of each dataset. These hyperparameters include the interaction graph pruning ratio ρ , the social-edge reliability fusion coefficients (α, β, η) , and the contrastive learning loss weight λ_{cl} . All remaining hyperparameters are shared across the three datasets to maintain comparability.

4.5 Performance Comparison

As can be seen from Table 3, DGDC-SR achieves the best performance on most evaluation metrics across the three datasets, indicating that the proposed model can effectively improve social recommendation performance. Compared with the best baseline model for each metric, DGDC-SR

shows consistent improvements in Precision@K, Recall@K, and NDCG@K. The improvement in Recall@K indicates that the model can better capture users' potential interests. The improvement in NDCG@K demonstrates that the model can

achieve better ranking quality in recommendation lists. The improvement in Precision@K further confirms the enhanced accuracy of the recommendation results.

Table 3: Overall recommendation performance.

Dataset	Metric	SBPR	DiffNet	LightGCN	MGDCF	SimGCL	ESRF	DSL	SocialLGN	CLDS	Ours	Impro.
LastFM	Precision@10	0.1398	0.1727	0.1961	0.1969	0.0526	0.1953	0.1974	0.1972	0.1979	0.2036	2.88%
	Precision@20	0.1010	0.1215	0.1358	0.1378	0.0362	0.1373	0.1399	0.1368	0.1388	0.1425	1.86%
	Recall@10	0.1442	0.1779	0.2003	0.2015	0.0529	0.1979	0.1995	0.2026	0.2030	0.2077	2.32%
	Recall@20	0.2070	0.2488	0.2769	0.2809	0.0722	0.2785	0.2823	0.2794	0.2834	0.2902	2.40%
	NDCG@10	0.1749	0.2219	0.2536	0.2549	0.0635	0.2373	0.2383	0.2566	0.2593	0.2639	1.77%
	NDCG@20	0.1978	0.2472	0.2788	0.2817	0.0699	0.2802	0.2812	0.2883	0.2868	0.2914	1.08%
Ciao	Precision@10	0.0179	0.0238	0.0271	0.0277	0.0233	0.0224	0.0281	0.0276	0.0287	0.0292	1.74%
	Precision@20	0.0141	0.0182	0.0202	0.0209	0.0170	0.0179	0.0208	0.0205	0.0217	0.0223	2.76%
	Recall@10	0.0259	0.0341	0.0410	0.0401	0.0357	0.0347	0.0356	0.0430	0.0413	0.0434	0.93%
	Recall@20	0.0412	0.0527	0.0591	0.0590	0.0504	0.0550	0.0560	0.0618	0.0621	0.0626	0.81%
	NDCG@10	0.0266	0.0359	0.0437	0.0427	0.0376	0.0345	0.0346	0.0441	0.0436	0.0447	1.36%
	NDCG@20	0.0307	0.0403	0.0478	0.0471	0.0409	0.0402	0.0405	0.0486	0.0488	0.0495	1.43%
Douban	Precision@10	0.1922	0.1722	0.2238	0.2467	0.0188	0.2067	0.2365	0.2345	0.2534	0.2596	2.45%
	Precision@20	0.1678	0.1560	0.1965	0.2117	0.0218	0.1840	0.2054	0.2035	0.2192	0.2248	2.55%
	Recall@10	0.0454	0.0431	0.0508	0.0581	0.0140	0.0513	0.0524	0.0628	0.0616	0.0646	2.87%
	Recall@20	0.0760	0.0798	0.0857	0.0954	0.0263	0.0885	0.0890	0.1007	0.1010	0.1028	1.78%
	NDCG@10	0.2108	0.1851	0.2437	0.2715	0.0194	0.2234	0.2335	0.2579	0.2774	0.2841	2.42%
	NDCG@20	0.1951	0.1792	0.2269	0.2492	0.0256	0.2120	0.2275	0.2405	0.2562	0.2630	2.65%

From the results on different datasets, DGDC-SR achieves stable performance improvements on LastFM, Ciao, and Douban. These datasets exhibit different characteristics in terms of user-item interactions and user-user social relations. LastFM contains relatively rich social information, Ciao suffers from sparse interaction and social connections, while Douban contains denser interaction information but relatively limited social relations. Under these diverse data distributions, DGDC-SR consistently maintains performance advantages over existing methods, demonstrating the effectiveness and general applicability of the proposed dual-graph denoising and cross-view alignment strategies in different social recommendation scenarios.

4.6 Hyperparameter Analysis

In this section, we analyze several key hyperparameters of the proposed model on the three datasets. These hyperparameters include the interaction graph pruning ratio ρ , the social-edge reliability fusion coefficients (α, β, η) , and the contrastive learning loss weight λ_{cl} .

1) First, we investigate the effect of the interaction graph pruning ratio ρ on model performance. As shown in Figure 3(a), the model performance in terms of Recall@10 and NDCG@10 first increases and then decreases as ρ increases. This indicates that properly removing low-quality interactions can suppress noise propagation and improve representation learning. However, an excessively large ρ may mistakenly remove useful interactions and weaken the model's ability to

learn user preferences, leading to a decline in recommendation performance. In our experiments, the model achieves the best performance when $\rho = 0.1$.

2) Second, we examine the effects of the three reliability signals in the social graph dynamic denoising module by varying the gradient-based reliability coefficient α , the preference-consistency coefficient β , and the structural-consistency coefficient η . Representative results are shown in Figure 3(b). Different coefficient combinations lead to different recommendation results, indicating that the three signals contribute differently to social-edge reliability estimation. Assigning an excessively large coefficient to a single signal does not further improve performance, suggesting that no individual signal is sufficient to characterize social-edge reliability. In contrast, jointly incorporating gradient-based stability, cross-view preference consistency, and local structural consistency provides a more balanced evaluation of social edges. The best performance is obtained when $(\alpha, \beta, \eta) = (0.4, 0.4, 0.2)$.

3) Finally, we analyze the effect of the contrastive learning loss weight λ_{cl} . As shown in Figure 3(c), model performance first improves and then declines as λ_{cl} increases. A very small λ_{cl} provides insufficient cross-view alignment supervision, whereas an excessively large value causes the contrastive objective to dominate the recommendation objective. The model performs better when λ_{cl} is between 10^{-4} and 10^{-3} . Accordingly, we set $\lambda_{cl} = 10^{-3}$ for LastFM and Douban and $\lambda_{cl} = 5 \times 10^{-4}$ for Ciao.

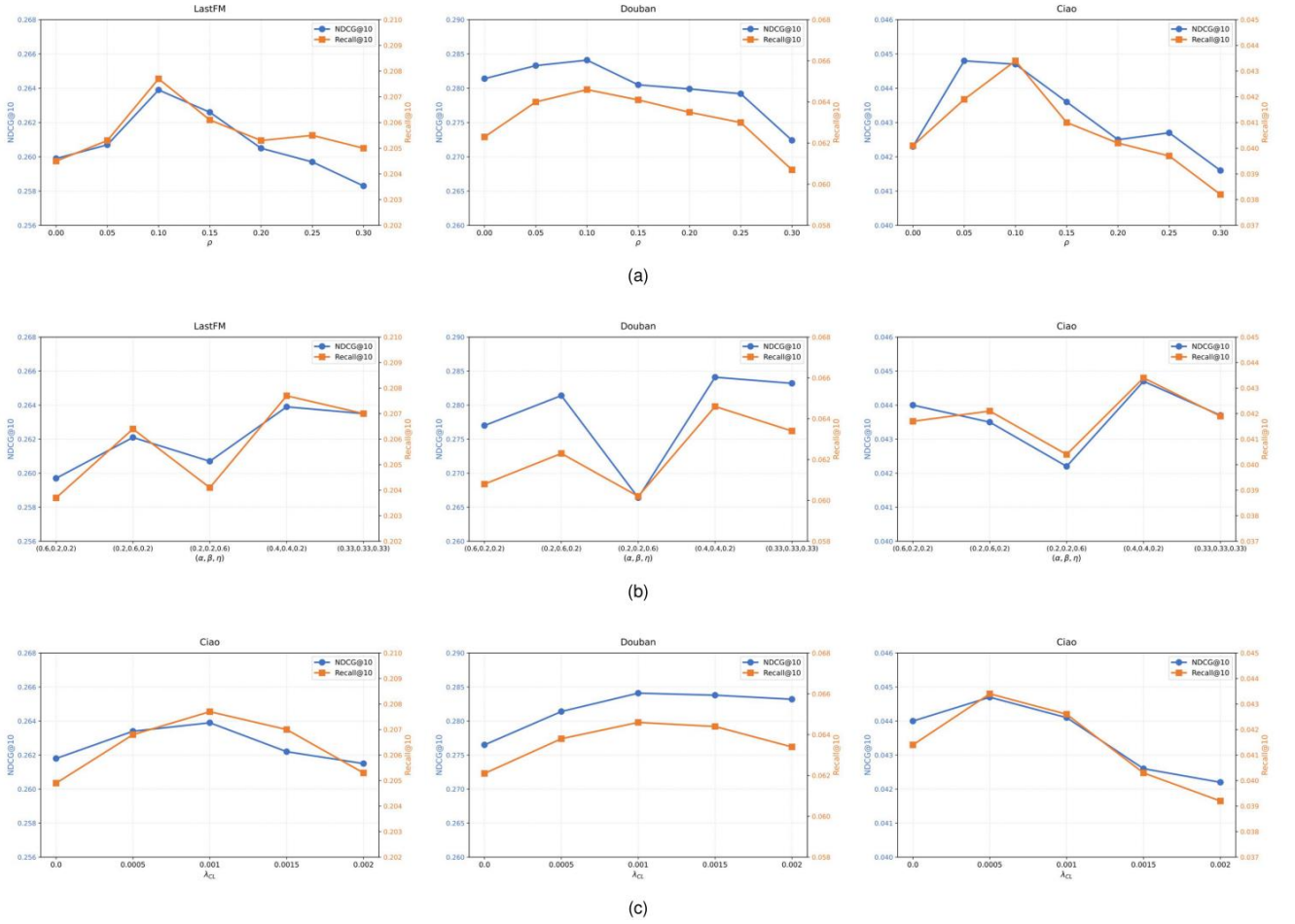


Figure 3: Hyperparameter analysis of the proposed model on the LastFM, Douban, and Ciao datasets.

4.7 Ablation Study

To evaluate the contribution of each key component, we conduct a series of ablation experiments. All variants use the same experimental settings, and only the target component is removed or replaced.

For the dual-graph dynamic denoising module, we consider three variants:

- **w/o Inter-D:** The interaction graph dynamic denoising module is removed, and representation propagation is performed on the original user–item interaction graph.
- **w/o Social-D:** The social graph dynamic denoising module is removed, and representation propagation is performed on the original user–user social graph.
- **w/o Dual-D:** Both the interaction graph dynamic denoising module and the social graph dynamic denoising module are removed.

For the cross-view representation learning module, we consider four variants:

- **w/o Bi-CL:** The bidirectional cross-view contrastive learning objective is removed, and only the recommendation objective is retained during fine-tuning.
- **S→I CL:** Only the social-to-interaction contrastive

learning objective is retained.

- **I→S CL:** Only the interaction-to-social contrastive learning objective is retained.
- **w/o Gate:** The adaptive gating fusion mechanism is replaced with simple average fusion.

Based on the results in Table 4, the following observations can be made.

1) Effect of dual-graph dynamic denoising. Removing either interaction graph dynamic denoising or social graph dynamic denoising decreases recommendation performance. This result indicates that low-quality information in both the user–item interaction graph and the user–user social graph can interfere with representation learning during graph propagation. Removing both denoising modules causes a larger overall performance decrease than removing either module individually, showing that interaction graph denoising and social graph denoising provide complementary benefits. These results verify that the dual-graph dynamic denoising module reduces the influence of low-quality connections and improves user representation learning.

2) Effect of bidirectional cross-view contrastive learning. Removing bidirectional cross-view contrastive learning decreases recommendation performance, demonstrating the importance of aligning user representations between the interaction and social views. In addition, both S→I CL and

I→S CL perform worse than the full bidirectional variant. This result indicates that a single-direction objective provides limited cross-view supervision, whereas bidirectional contrastive learning aligns the two representation spaces from both directions and yields more effective user representations.

3) Effect of adaptive gating fusion. Replacing adaptive gating fusion with simple average fusion decreases performance. A fixed averaging strategy assigns the same relative importance to the two views for all users and therefore cannot capture differences in users' reliance on interaction and social

information. In contrast, the adaptive gating mechanism adjusts the contribution of each view according to the corresponding user representations, resulting in more effective fused representations.

Overall, removing or replacing any key component of DGDC-SR leads to lower recommendation performance. This result confirms that dual-graph dynamic denoising, bidirectional cross-view contrastive learning, and adaptive gating fusion all contribute to the effectiveness of the complete model.

Table 4: Ablation study of the key components of DGDC-SR.

Dataset	Metric	DGDC-SR	w/o Inter-D	w/o Social-D	w/o Dual-D	w/o Bi-CL	S→I CL	I→S CL	w/o Gate
LastFM	Precision@10	0.2036	0.1981	0.1999	0.1961	0.2019	0.2021	0.2022	0.2032
	Precision@20	0.1425	0.1383	0.1407	0.1377	0.1403	0.1415	0.1413	0.1422
	Recall@10	0.2077	0.2046	0.2053	0.2024	0.2049	0.2061	0.2065	0.2075
	Recall@20	0.2902	0.2850	0.2866	0.2834	0.2885	0.2893	0.2891	0.2899
	NDCG@10	0.2639	0.2605	0.2614	0.2587	0.2621	0.2630	0.2629	0.2631
	NDCG@20	0.2914	0.2898	0.2902	0.2875	0.2904	0.2899	0.2898	0.2904
Ciao	Precision@10	0.0292	0.0283	0.0281	0.0274	0.0287	0.0289	0.0289	0.0288
	Precision@20	0.0223	0.0214	0.0208	0.0193	0.0215	0.0218	0.0220	0.0220
	Recall@10	0.0431	0.0414	0.0414	0.0406	0.0420	0.0424	0.0423	0.0426
	Recall@20	0.0626	0.0619	0.0613	0.0609	0.0619	0.0620	0.0620	0.0619
	NDCG@10	0.0447	0.0425	0.0418	0.0412	0.0438	0.0441	0.0443	0.0444
	NDCG@20	0.0495	0.0489	0.0487	0.0478	0.0489	0.0492	0.0491	0.0491
Douban	Precision@10	0.2596	0.2585	0.2589	0.2581	0.2571	0.2586	0.2589	0.2594
	Precision@20	0.2248	0.2229	0.2231	0.2224	0.2217	0.2232	0.2230	0.2245
	Recall@10	0.0646	0.0631	0.0633	0.0626	0.0628	0.0635	0.0634	0.0644
	Recall@20	0.1028	0.1010	0.1015	0.1003	0.1016	0.1022	0.1025	0.1027
	NDCG@10	0.2841	0.2797	0.2798	0.2785	0.2813	0.2823	0.2824	0.2836
	NDCG@20	0.2630	0.2596	0.2610	0.2588	0.2604	0.2616	0.2616	0.2628

5. Conclusion

This paper proposes a two-stage social recommendation model based on graph denoising and contrastive learning. To address the noise commonly found in user-item interaction graphs and user-user social graphs, the model performs dynamic denoising on both interaction and social relations in the pre-training stage. This process effectively suppresses the interference caused by noisy connections during representation propagation, thereby improving the quality of user preference modeling. During the fine-tuning stage, a bidirectional cross-view contrastive learning mechanism is employed to align user representations across different views. An adaptive fusion mechanism is then applied to integrate user representations from the interaction view and the social view, enabling the model to learn more comprehensive user preference representations. Experiments conducted on three mainstream datasets including LastFM, Ciao and Douban demonstrate that our designed model achieves superior performance over various baseline approaches across all adopted evaluation indicators. These results demonstrate the effectiveness of the dual-graph dynamic denoising strategy and the cross-view representation alignment strategy. Furthermore, ablation studies confirm that each key module contributes to the overall performance improvement.

Future research will explore the following directions. First, the negative sample construction strategy can be further optimized by introducing more stable hard negative mining methods or multi-negative learning mechanisms to improve contrastive learning and ranking optimization. Second, temporal information from user-item interactions can be incorporated to capture the dynamic evolution of user interests. Finally, large language models or multimodal

information can be explored for evaluating the reliability of social edges, which may further improve recommendation performance in complex real-world scenarios.

Declarations

Conflict of interest The authors confirm that no competing interests exist in relation to this manuscript's publication.

Data availability All datasets adopted and investigated in this research are publicly open, and relevant acquisition links are provided in the article.

References

- [1] X. Li, L. Sun and M. Ling, "A survey of graph neural network based recommendation in social networks." *Neurocomputing*, vol. 549, p. 126441, 2023, doi: <https://doi.org/10.1016/j.neucom.2023.126441>
- [2] K. Sharma, Y. C. Lee and S. Nambi, "A survey of graph neural networks for social recommender systems." *ACM Computing Surveys*, vol. 56no. 10, p. 1–34, 2024, doi: <https://doi.org/10.1145/3661821>
- [3] H. Ma, H. Yang and M. R. Lyu, "SoRec: Social recommendation using probabilistic matrix factorization." in *Proceedings of the 17th ACM conference on information and knowledge management*, p. 931–40, 2008, doi: <https://doi.org/10.1145/1458082.1458205>
- [4] M. Jamali and M. Ester, "A matrix factorization technique with trust propagation for recommendation in social networks." in *Proceedings of the fourth ACM conference on recommender systems*, p. 135–42, 2010, doi: <https://doi.org/10.1145/1864708.1864736>

- [5] H. Ma, D. Zhou and C. Liu, "Recommender systems with social regularization." in Proceedings of the fourth ACM international conference on web search and data mining, p. 287–96, 2011, doi: <https://doi.org/10.1145/1935826.1935877>
- [6] B. Wu, L. Zhong and L. Yao, "EAGCN: An efficient adaptive graph convolutional network for item recommendation in social internet of things." IEEE Internet of Things Journal, vol. 9no. 17, p. 16386–401, 2022, doi: <https://doi.org/10.1109/JIOT.2022.3151400>
- [7] L. Yang, Z. Liu and Y. Dou, "ConsisRec: Enhancing GNN for social recommendation via consistent neighbor aggregation." in Proceedings of the 44th ACM SIGIR conference on research and development in information retrieval, p. 2141–5, 2021, doi: <https://doi.org/10.1145/3404835.3463028>
- [8] J. Li and H. Wang, "Graph diffusive self-supervised learning for social recommendation." in Proceedings of the 47th ACM SIGIR conference on research and development in information retrieval, p. 2442–6, 2024, doi: <https://doi.org/10.1145/3626772.3657962>
- [9] W. Fan, Y. Ma and Q. Li, "Graph neural networks for social recommendation." in Proceedings of the world wide web conference, p. 417–26, 2019, doi: <https://doi.org/10.1145/3308558.3313488>
- [10] C. Tian, Y. Xie and Y. Li, "Learning to denoise unreliable interactions for graph collaborative filtering." in Proceedings of the 45th ACM SIGIR conference on research and development in information retrieval, p. 122–32, 2022, doi: <https://doi.org/10.1145/3477495.3531889>
- [11] Y. Qi, Q. Zhang and X. Lin, "Seeing beyond noise: Joint graph structure evaluation and denoising for multimodal recommendation." in Proceedings of the AAAI conference on artificial intelligence, p. 12461–9, 2025, doi: <https://doi.org/10.1609/aaai.v39i12.33358>
- [12] W. Ma, Y. Wang and Y. Zhu, "MADM: A model-agnostic denoising module for graph-based social recommendation." in Proceedings of the 17th ACM international conference on web search and data mining, p. 501–9, 2024, doi: <https://doi.org/10.1145/3616855.3635784>
- [13] Z. Hu, S. Nakagawa and Y. Zhuang, "Hierarchical denoising for robust social recommendation." IEEE Transactions on Knowledge and Data Engineering, vol. 37no. 2, p. 739–53, 2024, doi: <https://doi.org/10.1109/TKDE.2024.3508778>
- [14] W. Chen, Y. Zhang and H. Li, "Dual-domain collaborative denoising for social recommendation." IEEE Transactions on Computational Social Systems, 2025, doi: <https://doi.org/10.1109/TCSS.2025.3529706>
- [15] X. Chen, C. He and Q. Guan, "Boost social recommendation via adaptive denoising network." in Companion proceedings of the ACM web conference, p. 577–80, 2024, doi: <https://doi.org/10.1145/3589335.3651473>
- [16] B. Zu, Y. Liu and Y. Li, "DBA-SR: A denoised and bias-aware framework for social recommendation." IEEE Transactions on Computational Social Systems, 2026, doi: <https://doi.org/10.1109/TCSS.2026.3666811>
- [17] A. Li and B. Yang, "Dual graph denoising model for social recommendation." in Proceedings of the ACM web conference, p. 347–56, 2025, doi: <https://doi.org/10.1145/3696410.3714874>
- [18] Y. Sun, Z. Sun and Y. Du, "Self-supervised denoising through independent cascade graph augmentation for robust social recommendation." in Proceedings of the 30th ACM SIGKDD conference on knowledge discovery and data mining, p. 2806–17, 2024, doi: <https://doi.org/10.1145/3637528.3671958>
- [19] J. Gao, Y. Zhang and Y. Zhang, "Graph diffusion social recommendation." IEEE Transactions on Consumer Electronics, 2025, doi: <https://doi.org/10.1109/TCE.2025.3593495>
- [20] T. Wang, L. Xia and C. Huang, "Denoised self-augmented learning for social recommendation." 2023, doi: <https://doi.org/10.48550/arXiv.2305.12685>
- [21] L. Wu, J. Li and P. Sun, "DiffNet++: A neural influence and interest diffusion network for social recommendation." IEEE Transactions on Knowledge and Data Engineering, vol. 34no. 10, p. 4753–66, 2020, doi: <https://doi.org/10.1109/TKDE.2020.3048414>
- [22] Y. Peng, C. Gao and Y. Zhang, "Denoising alignment with large language model for recommendation." ACM Transactions on Information Systems, vol. 43no. 2, p. 1–35, 2025, doi: <https://doi.org/10.1145/3696662>
- [23] R. Jalan, T. Prakash and N. Pedanekar, "LLM-BRec: Personalizing session-based social recommendation with LLM-BERT fusion framework." in Proceedings of the second workshop on generative information retrieval, 2024
- [24] T. Mane and S. Kumbhar, "Optimized social recommendations using graphs and language model integration." in IEEE intelligent mobile computing (MobileCloud), p. 26–35, 2025, doi: <https://doi.org/10.1109/MobileCloud66020.2025.00008>
- [25] J. Chen, X. Xin and X. Liang, "GDSRec: Graph-based decentralized collaborative filtering for social recommendation." IEEE Transactions on Knowledge and Data Engineering, vol. 35no. 5, p. 4813–24, 2022, doi: <https://doi.org/10.1109/TKDE.2022.3153284>
- [26] Y. Yang, L. Wu and Z. Wang, "Graph bottlenecked social recommendation." in Proceedings of the 30th ACM SIGKDD conference on knowledge discovery and data mining, p. 3853–62, 2024, doi: <https://doi.org/10.1145/3637528.3671807>
- [27] Y. Quan, J. Ding and C. Gao, "Robust preference-guided denoising for graph based social recommendation." in Proceedings of the ACM web conference, p. 1097–108, 2023, doi: <https://doi.org/10.1145/3543507.3583374>
- [28] L. Qin, Y. Liu and C. Mu, "A social recommendation model based on social semantic mining and denoising." Journal of Information and Intelligence, 2025, doi: <https://doi.org/10.1016/j.jiixd.2025.04.003>
- [29] J. W. Taylor, "Short-term load forecasting with exponentially weighted methods." IEEE Transactions on Power Systems, vol. 27no. 1, p. 458–64, 2012, doi: <https://doi.org/10.1109/TPWRS.2011.2161780>
- [30] P. V. Marsden and N. E. Friedkin, "Network studies of social influence." Sociological Methods & Research, vol. 22no. 1, p. 127–51, 1993.
- [31] M. McPherson, L. Smith-Lovin and J. M. Cook, "Birds of a feather: Homophily in social networks." Annual Review of Sociology, vol. 27no. 1, p. 415–44, 2001, doi: <https://doi.org/10.1146/annurev.soc.27.1.415>
- [32] I. Konstas, V. Stathopoulos and J. M. Jose, "On social networks and collaborative recommendation." in Proceedings of the 32nd ACM SIGIR conference on

- research and development in information retrieval, p. 195–202, 2009, doi: <https://doi.org/10.1145/1571941.1571977>
- [33] L. Wu, Y. Ge and Q. Liu, “Modeling the evolution of users’ preferences and social links in social networking services.” *IEEE Transactions on Knowledge and Data Engineering*, vol. 29no. 6, p. 1240–53, 2017, doi: <https://doi.org/10.1109/TKDE.2017.2663422>
- [34] M. Gao, J. Zhang and J. Yu, “Recommender systems based on generative adversarial networks: A problem-driven perspective.” *Information Sciences*, vol. 546, p. 1166–85, 2021, doi: <https://doi.org/10.1016/j.ins.2020.09.013>
- [35] W. Fan, Y. Ma and Q. Li, “A graph neural network framework for social recommendations.” *IEEE Transactions on Knowledge and Data Engineering*, vol. 34no. 5, p. 2033–47, 2020, doi: <https://doi.org/10.1109/TKDE.2020.3008732>
- [36] B. Fu, W. Zhang and G. Hu, “Dual side deep context-aware modulation for social recommendation.” in *Proceedings of the web conference*, p. 2524–34, 2021, doi: <https://doi.org/10.1145/3442381.3449940>
- [37] W. Wang, Y. Xu and F. Feng, “Diffusion recommender model.” in *Proceedings of the 46th ACM SIGIR conference on research and development in information retrieval*, p. 832–41, 2023, doi: <https://doi.org/10.1145/3539618.3591663>
- [38] L. Wu, P. Sun and Y. Fu, “A neural influence diffusion model for social recommendation.” in *Proceedings of the 42nd ACM SIGIR conference on research and development in information retrieval*, p. 235–44, 2019, doi: <https://doi.org/10.1145/3331184.3331214>
- [39] J. Liao, W. Zhou and F. Luo, “SocialLGN: Light graph convolution network for social recommendation.” *Information Sciences*, vol. 589, p. 595–607, 2022, doi: <https://doi.org/10.1016/j.ins.2022.01.001>
- [40] X. He, K. Deng and X. Wang, “LightGCN: Simplifying and powering graph convolution network for recommendation.” in *Proceedings of the 43rd ACM SIGIR conference on research and development in information retrieval*, p. 639–48, 2020, doi: <https://doi.org/10.1145/3397271.3401063>
- [41] H. Xu, C. Huang and Y. Xu, “Global context enhanced social recommendation with hierarchical graph neural networks.” in *Proceedings of the IEEE international conference on data mining*, p. 701–10, 2020, doi: <https://doi.org/10.1109/ICDM50108.2020.00079>
- [42] W. Song, Z. Xiao and Y. Wang, “Session-based social recommendation via dynamic graph attention networks.” in *Proceedings of the twelfth ACM international conference on web search and data mining*, p. 555–63, 2019, doi: <https://doi.org/10.1145/3289600.3290989>
- [43] P. Veličković, W. Fedus, W. L. Hamilton, P. Liò, Y. Bengio and R. D. Hjelm, “Deep graph infomax.” in *International conference on learning representations*, 2018, doi: <https://doi.org/10.48550/arXiv.1809.10341>
- [44] F.-Y. Sun, J. Hoffman, V. Verma and J. Tang, “InfoGraph: Unsupervised and semi-supervised graph-level representation learning via mutual information maximization.” in *International conference on learning representations*, 2019, doi: <https://doi.org/10.48550/arXiv.1908.01000>
- [45] Y. You, T. Chen and Y. Sui, “Graph contrastive learning with augmentations.” *Advances in Neural Information Processing Systems*, vol. 33, p. 5812–23, 2020.
- [46] K. Hassani and A. H. Khasahmadi, “Contrastive multi-view representation learning on graphs.” in *Proceedings of the international conference on machine learning*, p. 4116–26, 2020.
- [47] J. Yu, H. Yin and X. Xia, “Are graph augmentations necessary? Simple graph contrastive learning for recommendation.” in *Proceedings of the 45th ACM SIGIR conference on research and development in information retrieval*, p. 1294–303, 2022, doi: <https://doi.org/10.1145/3477495.3531937>
- [48] J. Yu, X. Xia and T. Chen, “XSimGCL: Towards extremely simple graph contrastive learning for recommendation.” *IEEE Transactions on Knowledge and Data Engineering*, vol. 36no. 2, p. 913–26, 2023, doi: <https://doi.org/10.1109/TKDE.2023.3288135>
- [49] Z. Lin, C. Tian, Y. Hou and W. X. Zhao, “Improving graph collaborative filtering with neighborhood-enriched contrastive learning.” in *Proceedings of the ACM web conference*, p. 2320–9, 2022, doi: <https://doi.org/10.1145/3485447.3512104>
- [50] C. Zhou, J. Ma, J. Zhang, J. Zhou and H. Yang, “Contrastive learning for debiased candidate generation in large-scale recommender systems.” in *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery and data mining*, p. 3985–95, 2021, doi: <https://doi.org/10.1145/3447548.3467102>
- [51] X. Wang, W. Lu, M. Ester, C. Wang and C. Chen, “Social recommendation with strong and weak ties.” in *Proceedings of the 25th ACM international conference on information and knowledge management*, p. 5–14, 2016, doi: <https://doi.org/10.1145/2983323.2983701>
- [52] S. Rendle, C. Freudenthaler and Z. Gantner, “BPR: Bayesian personalized ranking from implicit feedback.” 2012, doi: <https://doi.org/10.48550/arXiv.1205.2618>
- [53] T. Zhao, J. McAuley and I. King, “Leveraging social connections to improve personalized ranking for collaborative filtering.” in *Proceedings of the 23rd ACM international conference on information and knowledge management*, p. 261–70, 2014, doi: <https://doi.org/10.1145/2661829.2661998>
- [54] J. Hu, B. Hooi and S. Qian, “MGDCF: Distance learning via markov graph diffusion for neural collaborative filtering.” *IEEE Transactions on Knowledge and Data Engineering*, vol. 36no. 7, p. 3281–96, 2024, doi: <https://doi.org/10.1109/TKDE.2023.3348537>
- [55] J. Yu, H. Yin and J. Li, “Enhancing social recommendation with adversarial graph convolutional networks.” *IEEE Transactions on Knowledge and Data Engineering*, vol. 34no. 8, p. 3727–39, 2020, doi: <https://doi.org/10.1109/TKDE.2020.3033673>
- [56] G. F. Ma, X. H. Yang and H. Long, “Robust social recommendation based on contrastive learning and dual-stage graph neural network.” *Neurocomputing*, vol. 584, p. 127597, 2024, doi: <https://doi.org/10.1016/j.neucom.2024.127597>