Robust Graph Based Social Recommendation Through Contrastive Multi-View Learning

Fei Xiong[1], Tao Zhang[1], Shirui Pan[2], Guixun Luo[3*], Liang Wang[4]

[1]School of Electronic and Information Engineering, Beijing Jiaotong University
[2]School of Information and Communication Technology, Griffith University
[3]School of Computer Science and Technology, Beijing Jiaotong University
[4]School of Computer Science, Northwestern Polytechnical University
xiongf@bjtu.edu.cn, taoakamoria@foxmail.com, s.pan@griffith.edu.au, gxluo@bjtu.edu.cn, liangwang0123@gmail.com

Abstract

Social recommendation leverages the social connections between users to mitigate the issue of data sparsity and enhance recommendation quality. Although existing related works show their effectiveness, there remain two critical questions: i) The patterns of preference interactions among users are varied and heterogeneous. Current models struggle to accurately capture preference shifts from user interactions in noisy social environments. ii) Existing methods handle the integration of auxiliary information coarsely, potentially introducing noise and leading to biases in user preferences. To address the limitations above, we introduce a novel framework named Robust Graph Based Social Recommendation Through Contrastive Multi-View Learning (RGCML). This framework leverages denoised social relations and global intents as dual auxiliary information sources to provide comprehensive characterization of users. Firstly, RGCML employs the concept of opinion dynamics to simulate how user preferences evolve due to noisy social relations. Then, it utilizes a specifically designed information fusion module to extract critical contextual information from multiple semantic perspectives, thereby achieving personalized information fusion. Finally, it adopts the designed global-local contrastive learning paradigm that untangles and discriminates user preferences from global intents, further addressing the noise problem and enhancing the quality of user representations. Extensive experiments conducted on three real-world datasets demonstrate the superior performance of RGCML compared to several state-of-the-art (SOTA) baselines.

Introduction

In the era of information explosion, recommender systems are capable of filtering complex information and play an important role in uncovering user preferences and providing personalized services (Sharma et al. 2024; Wu et al. 2022). Collaborative filtering (CF) is a prominent approach that infers user preferences from historical interactions to recommend items that align with their interests (Su and Khoshgoftaar 2009). Recently, CF models based on graph neural networks (GNNs) have achieved exceptional performance and become the dominant paradigm for recommender systems (He et al. 2020; Mao et al. 2021; Wang et al. 2019). The success can be attributed to the message propagation and neighbor aggregation mechanisms of GNN models, enabling them to capture high-order connectivity and obtain more relevant representations of users and items. However, the challenge of data sparsity, characterized by the limited number of historical interactions, significantly hinders the progress of recommender systems. According to the assumption of social homophily (McPherson, Smith-Lovin, and Cook 2001; Jiang et al. 2024), users with social connections are more likely to share similar interaction preferences, so mining user social relations can provide helpful user-side information to alleviate the data sparsity issue. Therefore, many scholars consider social relations as supplementary information and incorporate them into recommender systems to enhance the quality of recommendations (Fan et al. 2019; Yang et al. 2021).

In addition, to further alleviate the data sparsity problem, self-supervised learning is widely employed in recommender systems, enabling the exploitation of unlabelled data and improving model robustness and generalization (Yu et al. 2023). The basic idea of self-supervised learning is to generate self-supervised signals by constructing auxiliary views on raw data. For example, SGL (Wu et al. 2021) constructs augmented views by perturbing the graph structure and provides effective self-supervised signals to improve recommendation performance.

Despite the convincing results of existing models, we contend that several shortcomings persist.

First, blindly incorporating social information may introduce noise and degrade model performance. For example, a mother and her child might be socially connected, yet the mother may prefer household items while the child prefers sport items, highlighting the potential for heterogeneity in preferences even among connected individuals. Without proper handling, directly applying GNNs to acquire user social embeddings may introduce noise. Although some models have begun to address the issue of social noise and attempt to alleviate it through strategies like self-supervised learning (Wang, Xia, and Huang 2023), accurately simulating the propagation and evolution of preferences within social networks remains a challenge that requires further investigation.

Second, while existing models alleviate the data sparsity problem by integrating various types of auxiliary informa-

*Corresponding author

tion, there is a lack of scientifically rigorous and universally applicable methods to integrate auxiliary signals from multiple perspectives. This deficiency frequently leads to embeddings that fall short of optimality, thereby constraining the overall effectiveness of recommender systems. For example, GraphRec (Fan et al. 2019) adopts the approach of concatenating embeddings under different semantics and feeding them into MLP to get the final user representations, which may introduce noise. The gating mechanism used by IDVT (Yang et al. 2023) does not adequately account for the differences between users. Although some approaches attempt to utilize attention mechanisms for fusion, these methods can be computationally intensive and resource-demanding. Therefore, there is a need for an efficient and robust personalized fusion strategy.

In this paper, we propose a novel global-local contrastive learning model named Robust Graph Based Social Recommendation Through Contrastive Multi-View Learning (RGCML) that introduces rich user-side auxiliary information to portray user characteristics comprehensively. To capture the changes in user preferences due to social relations, we introduce the concept of the Hegselmann-Krause (HK) model of opinion dynamics from social physics, which filters out unreliable social connections and focuses on learning from reliable ones. To capture finer-grained interaction patterns between users and items, we further leverage user global intents as user-side auxiliary information. Subsequently, RGCML employs a side information fusion module to extract essential features under different semantics, promoting effective representation learning. Ultimately, a global-local contrastive learning module is designed to align the final user embeddings with preference embeddings and intent embeddings, intending to further address the noise problem inherent in information fusion. In summary, our contributions are as follows:

- We propose a novel social recommendation model that denoises social relations based on opinion dynamics and leverages user global intents to provide helpful auxiliary information.
- We propose an effective information fusion module that enables the personalized fusion of auxiliary information tailored to the characteristics of different users. Global-local contrastive learning is utilized to further address the noise problem inherent in information fusion and to obtain reliable user representations.
- We have conducted extensive experiments on three real-world datasets to confirm the improvements achieved by RGCML compared to several SOTA models.

Related Work

Self-Supervised Learning

Given the excellent performance of self-supervised learning in computer vision and natural language processing (He et al. 2019; Devlin et al. 2018), many recommendation models have integrated contrastive learning components (Cai et al. 2023; Chen et al. 2023). These models enhance recommendation accuracy and alleviate data sparsity by generating augmented views and maximizing the similarity between positive pairs while reducing it between negative pairs.

There are several methods to build augmented views, such as SGL (Wu et al. 2021) employing random edge/node dropout and SimGCL (Yu et al. 2022b) introducing perturbations to the node features. In NCL (Lin et al. 2022), representation alignment is performed among the specific user, the structural neighbors, and the semantic-centric node. DCCF (Ren et al. 2023) utilizes graph contrastive learning to disentangle user intents from preferences. Our approach refrains from altering the graph structure, thereby preserving critical interaction information.

Social Recommendation

Given the exceptional capability of GNNs to capture complex dependencies among nodes, GNN based social recommendation models have become the dominant paradigm. Such models primarily utilize GNNs to learn users' social and preference features and subsequently apply various fusion methods to derive the final user representations. DiffNet (Wu et al. 2019) considers the impact of social diffusion and performs embedding fusion through element-wise addition. DESIGN (Tao et al. 2022) utilizes an approach that integrates knowledge distillation to enhance the learning process. Additionally, some models incorporate the concept of self-supervised learning to further refine their representations. SEPT (Yu et al. 2021a) introduces a tri-training self-supervised framework. MHCN (Yu et al. 2021b) introduces a multi-channel hypergraph convolution network with self-supervised learning. DSL (Wang, Xia, and Huang 2023) uses dual semantic information for social denoising. SMIN (Long et al. 2021) uses metapath-guided node connections to investigate how the graph topology changes with additional self-supervision signals. We propose a social denoising approach based on opinion dynamics and introduce a personalized information fusion module that can fully take into account the differences among users.

Methodology

Preliminaries

We first introduce definitions and notations used in this paper. In social recommendation, two essential graphs are involved: interaction graph $\mathcal{G}_r$ and social graph $\mathcal{G}_s$. We represent the interaction matrix between the user set $\mathcal{U} = (u_1, \ldots, u_M)$ and item set $\mathcal{I} = (v_1, \ldots, v_N)$ as $\mathcal{R} \in \mathbb{R}^{M \times N}$, where M and N represent the number of users and items. The entry $\mathcal{R}_{i,j} \in \mathcal{R}$ is set to 1 if user u_i has interacted with item v_j before, or else, $\mathcal{R}_{i,j} = 0$. In addition, we denote the social relation matrix as $A_s \in \mathbb{R}^{M \times M}$ in a similar way. The detailed model architecture is shown in Figure 1.

Social Relation Modeling based on Opinion Dynamics

To effectively model interactions between users in the presence of social noise, we introduce the concept of opinion dynamics (Weng et al. 2023). It can simulate the process by which individual opinions evolve within a group. According to the Hegselmann-Krause (HK) model (Hegselmann and

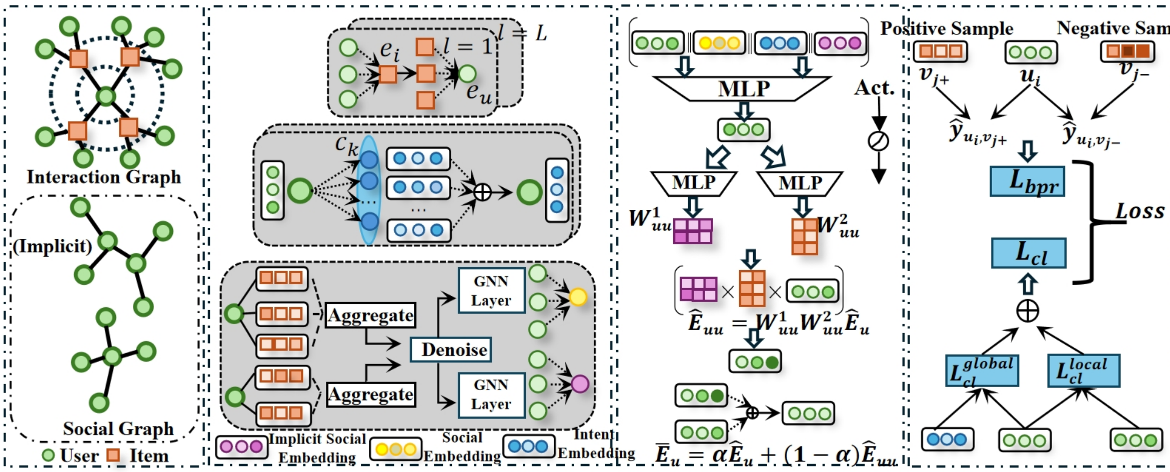

Figure 1: Overall framework of the proposed RGCML model.

Krause 2002), an individual's opinion shifts toward the average of all opinions within the confidence interval. Following the HK rule, the influence weight of user opinions can be represented as:

$$w_{ij}(t) = \begin{cases} \frac{1}{\#I(u_i,x(t))}, u_j \in I(u_i, x(t)) \\ 0, u_j \notin I(u_i, x(t)) \end{cases} \tag{1}$$

where $I(u_i, x(t)) = \{u_j : \|x_i(t) - x_j(t)\| \leq \varepsilon\}$, $I(u_i, x(t))$ denotes the set of neighbors of user u_i, $\#$ denotes the size of the set, ε is the tolerance threshold, and w_{ij} represents the weight of the influence of u_j on u_i. The updating process for user opinions can be represented as follows: $x_i(t+1) = \sum_{u_j \in I(u_i,x(t))} w_{ij}(t) x_j(t)$. In recommender systems, user historical interactions directly reflect their opinions, which can be used to accurately characterize user relevance. Our method considers the average of all item embeddings with which the user has interacted as the user's original opinion. It can be described as:

$$e_u^{in} = \frac{1}{|N_u|} \sum_{j \in N_u} e_j \tag{2}$$

where N_u denotes the set of items that interact with the specific user, e_u^{in} is the obtained user opinion, and e_j is the item embedding. In addition to this, the impact of the one-hop neighbors on the social graph is taken into account as well, so that the denoising process contains both structural information and original opinion information:

$$E_u^{si} = A_s E_u^{in} \tag{3}$$

Then, RGCML quantifies the similarity between users using cosine similarity,

$$\begin{aligned} \widetilde{r}_{u_1,u_2} &= \mathbb{I}(S_{core}(u_1, u_2) > \varepsilon) \\ &= \mathbb{I}((e_{u_1}^{si} (e_{u_2}^{si})^T / \left\|e_{u_1}^{si}\right\| \left\|e_{u_2}^{si}\right\|) > \varepsilon) \end{aligned} \tag{4}$$

Depending on the threshold $\widetilde{r}$, we can derive the denoised social graph $\widetilde{\mathcal{G}}_s$, where only relaible social connections are retained. Unlike the approach of averaging user views in the HK model, RGCML differentiates the influence of users by the degree of the node:

$$\widehat{E}_{s_1} = \mathcal{L}_s E_u, \mathcal{L}_s = D_s^{-\frac{1}{2}} \tilde{A}_s D_s^{-\frac{1}{2}} \tag{5}$$

where D_s is the corresponding diagonal degree matrix of $\widetilde{A}_s$ (denoised social relation matrix), $\widehat{E}_{s_1}$ is the obtained social embeddings, and $\mathcal{L}_s \in \mathbb{R}^{M \times M}$ is the normalized Laplacian matrix. In addition, when two users have interacted with a same item, we infer an implicit social relation between them and employ a similar way to derive the implicit social embeddings $\widehat{E}_{s_2}$.

Graph Convolution and Multi Global Intent Modeling

Message Propagation. We adopt GNNs as the backbone of RGCML to derive user preference embeddings and item embeddings. The message propagation is defined as:

$$\begin{aligned} e_u^{(l+1)} &= \textstyle\sum_{i \in N_u} \frac{1}{\sqrt{|N_u||N_i|}} e_i^{(l)} \\ e_i^{(l+1)} &= \textstyle\sum_{u \in N_i} \frac{1}{\sqrt{|N_i||N_u|}} e_u^{(l)} \end{aligned} \tag{6}$$

where N_i is the set of users interacting with the specific item, N_u is the set of items that interact with the specific user, and l denotes the layer index.

Global Intent Modeling. In real life, users choose to interact with items for a variety of intents, and these intents are often highly entangled with users' historical interactions. For example, a user may choose to watch the movie Oppenheimer out of the preference for biographical films, or because the movie has a high rating. Ignoring the presence of user global intents often results in suboptimal user representations, which in turn limits the recommendation quality. Therefore, in order to capture finer-grained interaction patterns between users and items, RGCML proposes to take user global intents into account as user-side auxiliary information. After getting the user preference embeddings for each layer, we will get the global intent embeddings accordingly.

Firstly, we initialize K global intent prototypes $c^k \in \mathbb{R}^D, k = 1, \ldots, K$. Then, we apply the softmax function to quantify the correlations between user embeddings and intent prototypes. Take user u as an example:

$$P(c^k|e_u^{(l)}) = \frac{exp({e_u^{(l)}}^T c^k)}{\sum_{k'=1}^K exp({e_u^{(l)}}^T c^{k'})} \tag{7}$$

In this way, the correlations between user preferences and global intents can be obtained to capture finer-grained user interaction patterns. Then, we aggregate the K learnable intent prototypes together to generate the user intent embedding at l-th layer:

$$r_u^{(l)} = \sum_{k=1}^{K} c^k P(c^k|e_u^{(l)}) = \sum_{k=1}^{K} c^k \frac{exp({e_u^{(l)}}^T c^k)}{\sum_{k'=1}^K exp({e_u^{(l)}}^T c^{k'})} \tag{8}$$

where $r_u^{(l)} \in \mathbb{R}^D$ is the user intent embedding at l-th layer.

Embedding Aggregation. To acquire the final embeddings, we adopt the mean-pooling method as follows:

$$R_u = \frac{1}{L}\sum_{l=1} R_u^{(l)}, \widehat{E}_u = \frac{1}{L}\sum_{l=1} E_u^{(l)}, \widehat{E}_i = \frac{1}{L}\sum_{l=1} E_i^{(l)} \tag{9}$$

where R_u is the user global intent embeddings, $\widehat{E}_u$ is the user preference embeddings, $\widehat{E}_i$ is the item embeddings.

Multi-View Information Fusion

The previous sections have introduced rich user-side auxiliary information under multiple perspectives. However, existing researches have not proposed an efficient personalized fusion strategy to integrate the auxiliary information. In RGCML, a generalized fusion module is used to efficiently fuse the auxiliary information and address the noise problem. Firstly, RGCML extracts the salient knowledge from different perspectives in order to preserve important user features under each semantic, which can be represented as:

$$\begin{aligned} Z_{uu} &= f(M_{uu}; \theta) \\ &= f(\widehat{E}_{s_1}||\widehat{E}_{s_2}||R_u||\widehat{E}_u; \theta) \end{aligned} \tag{10}$$

where $M_{uu} \in \mathbb{R}^{M \times 4D}$ is obtained by concatenating different side embeddings, and $f(\cdot)$ is a three-layer perceptron. M_{uu} contains the two types of social embeddings $\widehat{E}_{s_1}$ and $\widehat{E}_{s_2}$, as well as the global intent embeddings R_u. To transfer the fused information to the same semantic space as $\widehat{E}_u$, M_{uu} also includes the user preference embeddings $\widehat{E}_u$. Subsequently, Z_{uu} will be used as input to generate personalized knowledge transfer matrices. Meanwhile, the idea of low-rank matrix factorization (Xia et al. 2021; Chen et al. 2023) is adopted to efficiently generate personalized knowledge transfer matrices for each user:

$$W_{uu}^1 = f(\widehat{w}^1 Z_{uu} + \widehat{v}^1); W_{uu}^2 = f(\widehat{w}^2 Z_{uu} + \widehat{v}^2) \tag{11}$$

where $W_{uu}^1 \in \mathbb{R}^{M \times D \times D'}$, $W_{uu}^2 \in \mathbb{R}^{M \times D' \times D}$ are two dynamic network weights, and $f(a) = softmax(a + glo(a))$. $glo(\cdot)$ is the introduced global information (mean-pooling is used in the task). We let $D' < D$ to reduce the number of parameters ($D' = 3$ in our task). $\widehat{w}^1$, $\widehat{v}^1$, $\widehat{w}^2$ and $\widehat{v}^2$ are trainable parameters. In the learning process, the differences of users can be fully considered to achieve efficient personalized integration.

Lastly, personalized multi-view information fusion is achieved by using the obtained matrices W_{uu}^1 and W_{uu}^2 such that $\widehat{E}_{uu}$ contains contextual information in both social semantic and global intent semantic:

$$\widehat{E}_{uu} = W_{uu}^1 W_{uu}^2 \widehat{E}_u \tag{12}$$

where $\widehat{E}_{uu}$ consists auxiliary information from multiple views, which can be treated as comprehensively enhanced auxiliary information to improve the representational capability. The final user embeddings are obtained by: $\bar{E}_u = \alpha\widehat{E}_u + (1-\alpha)\widehat{E}_{uu}$, where α controls the weights and $\bar{E}_u \in \mathbb{R}^{M \times D}$ is the final user embeddings.

Global-Local Contrastive Learning

To further address the noise problem during information fusion, we propose a global-local contrastive learning approach. Concretely, we employ the InfoNCE loss which aims at optimizing the similarity between user embeddings across various perspectives of the same individual:

$$\begin{aligned} \mathcal{L}_{cl}^{local} &= \textstyle\sum_{u \in \mathcal{U}} - log \frac{exp(s(\bar{e}_u, \widehat{e}_u)/\tau_1)}{\sum_{v \in \mathcal{U}} exp(s(\bar{e}_u, \widehat{e}_v)/\tau_1)} \\ \mathcal{L}_{cl}^{global} &= \textstyle\sum_{u \in \mathcal{U}} - log \frac{exp(s(\bar{e}_u, r_u)/\tau_2)}{\sum_{v \in \mathcal{U}} exp(s(\bar{e}_u, r_v/\tau_2)} \end{aligned} \tag{13}$$

where τ_1, τ_2 are the temperature coefficients to automatically identify different negative samples, $s(\cdot)$ is the cosine similarity function, and v denotes negative samples.

For the main recommendation task, we use the classical Bayesian Personalized Ranking (BPR) loss function:

$$\mathcal{L}_{bpr} = -\sum_{(u_i, v_{j_+}, v_{j_-})} \ln \sigma(\hat{y}_{u_i, v_{j_+}} - \hat{y}_{u_i, v_{j_-}}), \hat{y}_{u_i, v_j} = e_i^T e_j \tag{14}$$

where $e_i \in \mathbb{R}^D$, $e_j \in \mathbb{R}^D$ are the corresponding user/item embeddings from $\bar{E}_u$ / $\widehat{E}_i$. $\hat{y}_{u_i, v_j}$ is the prediction score for recommendation which is obtained by inner product.

Finally, we integrate the self-supervised loss with primary recommendation loss into a multi-task learning paradigm as follows:

$$\mathcal{L} = \mathcal{L}_{bpr} + \lambda_1 \mathcal{L}_{cl}^{global} + \lambda_2 \mathcal{L}_{cl}^{local} + \lambda_3 \|\Theta\|_2 \tag{15}$$

where λ_1, λ_2, λ_3 are parameters for weight tuning, and Θ represents the set of learnable parameters in our model.

Model Analysis

The influence of the negative node v on the gradient of the global contrastive loss can be calculated as: $\|c_v\|_2 \propto \sqrt{1-(s_u^T g_v)^2} \exp(s_u^T g_v/\tau)$ where $s_u = \bar{e}_u / \|\bar{e}_u\|$ and $g_v = r_v / \|r_v\|$. If the positive node u and negative node v are much more similar, $\|c_v\|_2$ will be larger with the temperature coefficient τ set reasonably well(Wu et al. 2021). In other words, negative node will provide larger gradients thus guiding the optimization process. This will utilize the supervisory signals provided by the auxiliary information, making the node embeddings more discriminative. In this way, the two contrastive learning tasks constitute a joint optimization objective. In the optimization process, $\bar{E}_u$, R_u, and $\widehat{E}_u$ will interact with each other to find an optimal representation that can satisfy the two contrastive tasks, which can retain the correlation information from different perspectives and more comprehensively describe the sample features.

Evaluation

Datasets

Three real-world datasets: Douban, Ciao and Yelp are used in our experiments to evaluate RGCML. We split the interaction records into training, validation, and testing set with a ratio of 7:1:2. Since the goal of RGCML is to improve the performance of Top-N recommendation, we use Recall and Normalized Discounted Cumulative Gain (NDCG) to measure the model performance. We present the statistics of the datasets in Table 1 and use NVIDIA RTX 4090 to accomplish the experiment.

Dataset	Ciao	Yelp	Douban
#User	7368	19539	12638
#Item	78583	21298	22222
#Interactions	227393	450884	598420
Density	$3.92e^{-4}$	$1.08e^{-3}$	$2.13e^{-3}$
#Social Ties	111781	864157	169150

Table 1: Statistics of datasets

Baselines

To comprehensively evaluate the performance of RGCML, we compare RGCML against various baselines. We categorize them into four groups. i) GNN-based models. **LightGCN** (He et al. 2020) simplifies the design of GNN and proposes a light weight graph convolution network to enhance the training efficiency and recommendation quality. ii) GNN-based social models. **DiffNet** (Wu et al. 2019) and **ESRF** (Yu et al. 2022a). iii) Self-supervised models. **SGL** (Wu et al. 2021), **NCL** (Lin et al. 2022), and **DCCF** (Ren et al. 2023). iv) Self-supervised social models. **SEPT** (Yu et al. 2021a) and **MHCN** (Yu et al. 2021b).

Hyperparameter Settings

As for the general settings of all methods, we empirically set the embedding size D to 64. The models are initialized by Xavier method and optimized by the Adam optimizer (Kingma and Ba 2014). For graph-based models, the number of propagation layers is fixed at 2. As for the RGCML-specific parameters, the number of global intent prototypes K is selected from the range of [100,300,500,700,1000], α is selected from $[0.5, 0.6, 0.7, 0.8, 0.9, 0.99]$, λ_1/λ_2 are tuned from [0.01,0.05,0.1,0.5], and temperature coefficients τ_1/τ_2 are tuned from [0.05,0.1,0.2].

Overall Performance Comparison

Table 2 demonstrates that our proposed RGCML model performs better on three datasets compared to different baselines. We have several observations: (1) RGCML achieves significant improvement in the sparser dataset, Ciao (up to 3.5% improvement in Recall@5, and 2.53% improvement in NDCG@5). This shows that RGCML has excellent ability when dealing with data sparsity problems. This enhancement is attributed to the following factors: i) the provision of useful user-side information; ii) the fusion module which enables the effective integration of auxiliary information. (2) Models that utilize self-supervised learning demonstrate superior performance in most scenarios, emphasizing the significance of extracting self-supervised signals from unlabeled data. Our RGCML proposes a global-local contrastive learning paradigm that mitigates the noise problem during information fusion by maximizing the mutual information of different perspectives. (3) Introducing social connections in recommendation does not always improve performance. In the case of Yelp, which has a large number of social connections, MHCN and SEPT are not as effective as SGL. However, RGCML achieves 5.13% and 2.16% improvement on Recall@5 and NDCG@5 compared to the best baseline, which reflects the excellent ability of RGCML to handle social noise.

Ablation Study

We construct several RGCML variants for ablation study. The results are presented in Table 3. We denote w/o-exc (imp) as the variant that removes auxiliary social information (where **imp** denotes implicit social relations), w/o-fus as the variant using an MLP in place of the designed information fusion module, w/o-local (global) as the variant eliminating local (global) contrastive learning, and w/o-both as the variant without any contrastive learning.

w/o-both variant eliminates the contrastive learning module, leading to a substantial decline in recommendation performance. Specifically, NDCG@5 drops by 11.6%, 17.3%, and 11.9% on the Ciao, Yelp, and Douban. The experimental results from the w/o-global and w/o-local variants reveal that removing either type of contrastive learning component results in performance degradation across various metrics on all three datasets. Notably, when local contrastive learning is omitted, the model's performance deterioration becomes particularly pronounced (with an average reduction of 7.7% in the three metrics on Ciao, 3.1% on Yelp, and 11.3% on Douban). The w/o-fus variant also exhibits degraded recommendation efficacy across different datasets, most strikingly with an average decline of 12.1% in the three metrics on Douban, thus underscoring the effectiveness of the proposed multi-perspective fusion module. The results obtained from Ciao and Yelp underscore the importance of refined social information. In contrast, the absence of a notable impact from socialization on the Douban dataset might be attributed to the limited utility of the social relations present in this particular dataset.

Dataset	Metric	DiffNet	ESRF	LightGCN	SGL	SEPT	MHCN	NCL	DCCF	RGCML
Yelp	Recall@5	0.03587	0.03923	0.04102	0.05084	0.04105	0.04239	0.04882	0.04921	**0.05345**
	NDCG@5	0.04117	0.04477	0.04711	0.05834	0.04746	0.05004	0.05501	0.05619	**0.0596**
	Recall@20	0.09429	0.10208	0.10717	0.12415	0.10614	0.11127	0.11955	0.1242	**0.12759**
	NDCG@20	0.0587	0.06361	0.06701	0.07991	0.06676	0.07031	0.07636	0.07875	**0.08197**
Ciao	Recall@5	0.02041	0.02731	0.0263	0.02915	0.02777	0.02891	0.02895	0.02812	**0.03017**
	NDCG@5	0.02505	0.03489	0.03287	0.03715	0.03374	0.0359	0.03567	0.03438	**0.03809**
	Recall@20	0.04982	0.06722	0.06729	0.07298	0.06846	0.07405	0.07197	0.06919	**0.07464**
	NDCG@20	0.0332	0.04564	0.04436	0.04868	0.04531	0.04815	0.04736	0.04589	**0.04971**
Douban	Recall@5	0.07025	0.06778	0.06549	0.08452	0.06709	0.07117	0.07742	0.08249	**0.08503**
	NDCG@5	0.12229	0.11568	0.11672	0.14497	0.11985	0.13033	0.13284	0.14571	**0.1492**
	Recall@20	0.15312	0.15384	0.15062	0.17696	0.15404	0.15646	0.16511	0.17587	**0.18388**
	NDCG@20	0.12957	0.12728	0.12628	0.15281	0.12956	0.13527	0.14162	0.1521	**0.15748**

Table 2: Overall Performance Comparison

Dataset	Yelp		Ciao		Douban	
Metrics	Recall	NDCG	Recall	NDCG	Recall	NDCG
w/o-imp	0.05228	0.05914	0.02969	0.03755	0.08484	0.14915
w/o-exc	0.05255	0.05903	0.02906	0.0377	0.08438	**0.14933**
w/o-fus	0.05175	0.0578	0.02852	0.03678	0.07564	0.1331
w/o-local	0.05155	0.05765	0.02802	0.03569	0.07601	0.13417
w/o-global	0.05185	0.0592	0.02948	0.03719	0.08342	0.1478
w/o-both	0.04372	0.05083	0.02705	0.03413	0.07609	0.13338
RGCML	**0.05345**	**0.0596**	**0.03017**	**0.03809**	**0.08503**	0.1492

Table 3: Ablation study on key components of RGCML (measured by Recall@5 and NDCG@5)

In-Depth Analysis of RGCML

Model Robustness to Data Sparsity. To systematically evaluate the model robustness against data sparsity, we categorize users into four groups based on the number of items they have interacted with. The specific experimental results are shown in Figure 2. The bars represent the improvement between RGCML and each baseline, while the lines represent performance curves. RGCML achieves the best recommendation quality in the vast majority of cases. On Douban and Yelp, RGCML delivers larger performance gains mainly in the case of fewer interactions (less than 15 items). This phenomenon proves the excellent ability of RGCML to cope with the data sparsity problem.

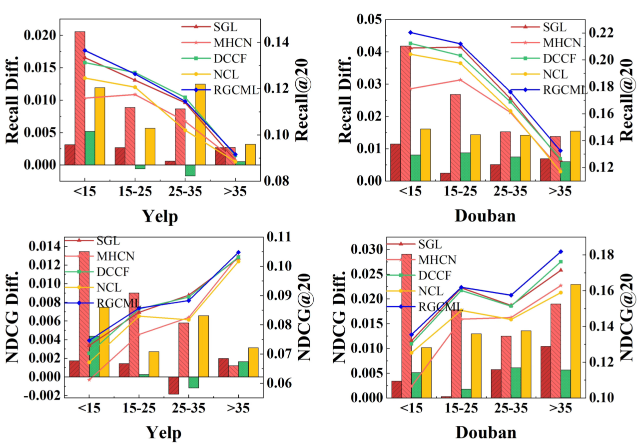

Figure 2: Model robustness study to data sparsity.

Model Robustness to Noise. This section will discuss the robustness of RGCML to the noise problem. Specifically, we will add different percentages (5%, 10%, 15%, 20%) of fake user-item interaction data to the training sets of Ciao and Yelp, respectively, and keep the test sets unchanged to re-compare the performance of RGCML, SGL and NCL. The results are shown in Figure 3.

The performance of RGCML is consistently maintained at a high level in different cases. Especially on the Ciao, it can be clearly found that the performance of RGCML is significantly better than the other models under different noise ratios.

Embedding Visualization Analysis

To gain deeper understanding of the improvements brought by RGCML, we map the representations (randomly sample 2000 users) to 3D space. The first row in Figure 4 shows the results for Ciao and the results for Yelp are shown in the second row. We can see the representations generated by RGCML exhibit uniformity. This indicates that the RGCML model takes into account the variability among users, while the large embedding distribution distance also indicates that it retains more information.

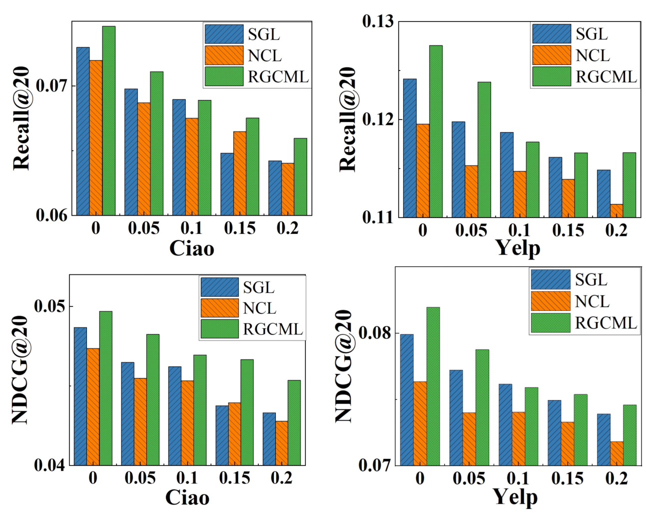

Figure 3: Model robustness study to noise.

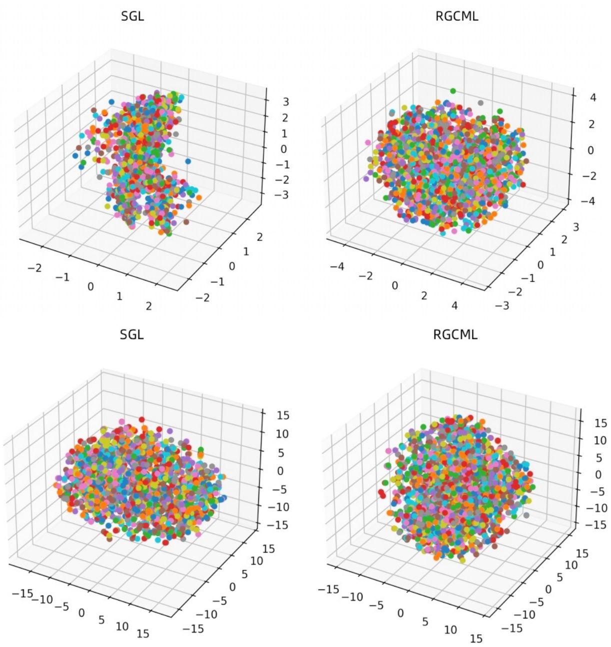

Figure 4: Visualization of user embeddings.

Training Efficiency Study

We further conduct the model efficiency analysis. Figure 5 represents the running time per epoch of each model. We can see that our proposed model achieves competitive results on three datasets. This suggests that RGCML has a high potential application value when dealing with large-scale datasets in real recommendation scenarios.

Hyperparameter Analysis

In this part, we perform parameter sensitivity analysis to show the impact of the two important hyperparameters in RGCML. The results are shown in Figure 6 and Figure 7.

The Number of Global Intents K. The best recommendation results can be achieved on Yelp when K is 500. It can be found that the effect of the model shows a tendency to increase and then decrease with the increase of K. This reflects the effect of global intents on the model, i.e.,

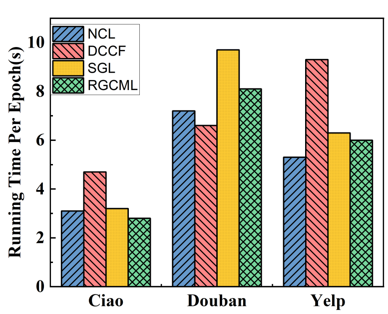

Figure 5: Comparison of computational costs in terms of training time per epoch on the Yelp, Ciao, and Douban.

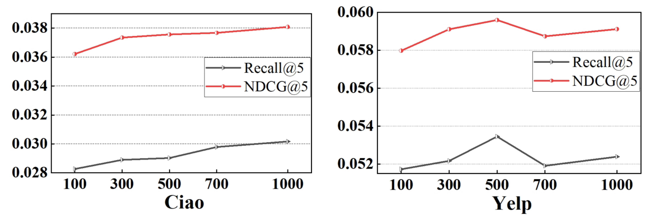

Figure 6: The effect of K on model performance.

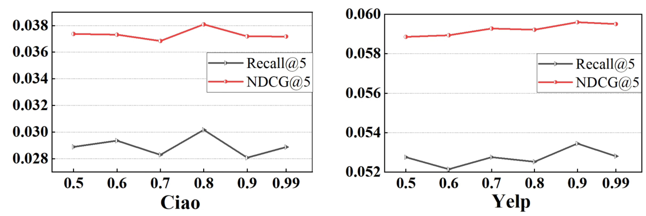

Figure 7: The effect of α on model performance.

RGCML can capture fine-grained user-item interaction patterns, but over-dividing the intents may lead to the introduction of noise. For Ciao, the model effectiveness improves as K increases and reliable recommendation quality can be achieved at $K = 1000$. However, further refinement of the intents may lead to performance degradation based on the experience gained on Yelp.

Fusion Weight α. When the α is 0.8 and 0.9, the best recommendation performance can be achieved on Ciao and Yelp. It is worth mentioning that when α is 0.99, the final user representations will be dominated by preference embeddings. However, the recommendation effect at this time is not as good as the optimal performance. This phenomenon strongly indicates that integrating user-side auxiliary information in RGCML model can substantially optimize the performance of recommender systems.

Conclusion

Despite the commendable performance of existing social recommendation models, various types of noise undermine their recommendation accuracy. To further address these noise issues, we propose an innovative model Robust Graph Based Social Recommendation Through Contrastive Multi-View Learning (RGCML). Firstly, drawing on the concept of opinion dynamics, our model effectively mitigates the problem of social noise by capturing the influence of social relations on preferences. Secondly, we introduce an efficient information fusion module that personalizes the integration of multi-view auxiliary information for different users. Ultimately, through the designed global-local contrastive learning module, we tackle the noise problems encountered during information fusion, enhancing the quality of recommendations. Extensive experiments on three real-world datasets convincingly demonstrate the superiority of our proposed model.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 62472024 and Grant 72004009.

References

Cai, X.; Huang, C.; Xia, L.; and Ren, X. 2023. LightGCL: Simple Yet Effective Graph Contrastive Learning for Recommendation. In *The Eleventh International Conference on Learning Representations*.

Chen, M.; Huang, C.; Xia, L.; Wei, W.; Xu, Y.; and Luo, R. 2023. Heterogeneous Graph Contrastive Learning for Recommendation. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, 544–552. ArXiv:2303.00995 [cs].

Devlin, J.; Chang, M.; Lee, K.; and Toutanova, K. 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *CoRR*, abs/1810.04805.

Fan, W.; Ma, Y.; Li, Q.; He, Y.; Zhao, E.; Tang, J.; and Yin, D. 2019. Graph Neural Networks for Social Recommendation. In *The World Wide Web Conference*, WWW '19, 417–426. New York, NY, USA: Association for Computing Machinery. ISBN 9781450366748.

He, K.; Fan, H.; Wu, Y.; Xie, S.; and Girshick, R. B. 2019. Momentum Contrast for Unsupervised Visual Representation Learning. *CoRR*, abs/1911.05722.

He, X.; Deng, K.; Wang, X.; Li, Y.; Zhang, Y.; and Wang, M. 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*, 639–648. Virtual Event China: ACM. ISBN 978-1-4503-8016-4.

Hegselmann, R.; and Krause, U. 2002. Opinion Dynamics and Bounded Confidence Models, Analysis and Simulation. *Journal of Artificial Societies and Social Simulation*, 5.

Jiang, W.; Gao, X.; Xu, G.; Chen, T.; and Yin, H. 2024. Challenging Low Homophily in Social Recommendation. In *Proceedings of the ACM on Web Conference 2024*, WWW '24, 3476–3484. New York, NY, USA: Association for Computing Machinery. ISBN 9798400701719.

Kingma, D.; and Ba, J. 2014. Adam: A Method for Stochastic Optimization. *Computer Science*.

Lin, Z.; Tian, C.; Hou, Y.; and Zhao, W. X. 2022. Improving Graph Collaborative Filtering with Neighborhood-enriched Contrastive Learning. In *Proceedings of the ACM Web Conference 2022*, 2320–2329. ArXiv:2202.06200 [cs].

Long, X.; Huang, C.; Xu, Y.; Xu, H.; Dai, P.; Xia, L.; and Bo, L. 2021. Social Recommendation with Self-Supervised Metagraph Informax Network. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 1160–1169. ArXiv:2110.03958 [cs].

Mao, K.; Zhu, J.; Xiao, X.; Lu, B.; Wang, Z.; and He, X. 2021. UltraGCN: Ultra Simplification of Graph Convolutional Networks for Recommendation. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, CIKM '21, 1253–1262. New York, NY, USA: Association for Computing Machinery. ISBN 9781450384469.

McPherson, M.; Smith-Lovin, L.; and Cook, J. M. 2001. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27(Volume 27, 2001): 415–444.

Ren, X.; Xia, L.; Zhao, J.; Yin, D.; and Huang, C. 2023. Disentangled Contrastive Collaborative Filtering. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1137–1146. ArXiv:2305.02759 [cs].

Sharma, K.; Lee, Y.-C.; Nambi, S.; Salian, A.; Shah, S.; Kim, S.-W.; and Kumar, S. 2024. A Survey of Graph Neural Networks for Social Recommender Systems. *ACM Comput. Surv.*, 56(10).

Su, X.; and Khoshgoftaar, T. M. 2009. A survey of collaborative filtering techniques. *Adv. in Artif. Intell.*, 2009.

Tao, Y.; Li, Y.; Zhang, S.; Hou, Z.; and Wu, Z. 2022. Revisiting Graph based Social Recommendation: A Distillation Enhanced Social Graph Network. In *Proceedings of the ACM Web Conference 2022*, 2830–2838. Virtual Event, Lyon France: ACM. ISBN 978-1-4503-9096-5.

Wang, T.; Xia, L.; and Huang, C. 2023. Denoised Self-Augmented Learning for Social Recommendation. *arXiv preprint arXiv:2305.12685*.

Wang, X.; He, X.; Wang, M.; Feng, F.; and Chua, T.-S. 2019. Neural Graph Collaborative Filtering. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 165–174. ArXiv:1905.08108 [cs].

Weng, L.; Zhang, Q.; Lin, Z.; Wu, L.; and Zhang, J.-H. 2023. Integrating interactions between target users and opinion leaders for better recommendations: An opinion dynamics approach. *Computer Communications*, 198: 98–107.

Wu, J.; Wang, X.; Feng, F.; He, X.; Chen, L.; Lian, J.; and Xie, X. 2021. Self-supervised Graph Learning for Recommendation. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 726–735. ArXiv:2010.10783 [cs].

Wu, L.; Sun, P.; Fu, Y.; Hong, R.; Wang, X.; and Wang, M. 2019. A Neural Influence Diffusion Model for Social Recommendation. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR'19, 235–244. New York, NY, USA: Association for Computing Machinery. ISBN 9781450361729.

Wu, S.; Sun, F.; Zhang, W.; Xie, X.; and Cui, B. 2022. Graph Neural Networks in Recommender Systems: A Survey. *ACM Comput. Surv.*, 55(5).

Xia, L.; Xu, Y.; Huang, C.; Dai, P.; and Bo, L. 2021. Graph Meta Network for Multi-Behavior Recommendation. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 757–766. Virtual Event Canada: ACM. ISBN 978-1-4503-8037-9.

Yang, D.; Ma, J.; Feng, S.; Zhang, H.; and Zhang, Z. 2023. IDVT: Interest-aware Denoising and View-guided Tuning for Social Recommendation. ArXiv:2308.15926 [cs].

Yang, L.; Liu, Z.; Dou, Y.; Ma, J.; and Yu, P. S. 2021. ConsisRec: Enhancing GNN for Social Recommendation via Consistent Neighbor Aggregation. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2141–2145. ArXiv:2105.02254 [cs].

Yu, J.; Yin, H.; Gao, M.; Xia, X.; Zhang, X.; and Viet Hung, N. Q. 2021a. Socially-Aware Self-Supervised Tri-Training for Recommendation. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 2084–2092. Virtual Event Singapore: ACM. ISBN 978-1-4503-8332-5.

Yu, J.; Yin, H.; Li, J.; Gao, M.; Huang, Z.; and Cui, L. 2022a. Enhancing Social Recommendation With Adversarial Graph Convolutional Networks. *IEEE Transactions on Knowledge and Data Engineering*, 34(8): 3727–3739.

Yu, J.; Yin, H.; Li, J.; Wang, Q.; Hung, N. Q. V.; and Zhang, X. 2021b. Self-Supervised Multi-Channel Hypergraph Convolutional Network for Social Recommendation. In *Proceedings of the Web Conference 2021*, WWW '21, 413–424. New York, NY, USA: Association for Computing Machinery. ISBN 9781450383127.

Yu, J.; Yin, H.; Xia, X.; Chen, T.; Cui, L.; and Nguyen, Q. V. H. 2022b. 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*, SIGIR '22, 1294–1303. New York, NY, USA: Association for Computing Machinery. ISBN 9781450387323.

Yu, J.; Yin, H.; Xia, X.; Chen, T.; Li, J.; and Huang, Z. 2023. Self-supervised learning for recommender systems: A survey. *IEEE Transactions on Knowledge and Data Engineering*.