Graph Bottlenecked Social Recommendation

Yonghui Yang
Key Laboratory of Knowledge Engineering with Big Data, Hefei University of Technology
yyh.hfut@gmail.com

Le Wu*
Key Laboratory of Knowledge Engineering with Big Data, Hefei University of Technology
lewu.ustc@gmail.com

Zihan Wang
Key Laboratory of Knowledge Engineering with Big Data, Hefei University of Technology
zhwang.hfut@gmail.com

Zhuangzhuang He
Key Laboratory of Knowledge Engineering with Big Data, Hefei University of Technology
hyicheng223@gmail.com

Richang Hong
Key Laboratory of Knowledge Engineering with Big Data, Hefei University of Technology
hongrc.hfut@gmail.com

Meng Wang
Key Laboratory of Knowledge Engineering with Big Data, Hefei University of Technology
eric.mengwang@gmail.com

ABSTRACT

With the emergence of social networks, social recommendation has become an essential technique for personalized services. Recently, graph-based social recommendations have shown promising results by capturing the high-order social influence. Most empirical studies of graph-based social recommendations directly take the observed social networks into formulation, and produce user preferences based on social homogeneity. Despite the effectiveness, we argue that social networks in the real-world are inevitably noisy (existing redundant social relations), which may obstruct precise user preference characterization. Nevertheless, identifying and removing redundant social relations is challenging due to a lack of labels. In this paper, we focus on learning the denoised social structure to facilitate recommendation tasks from an information bottleneck perspective. Specifically, we propose a novel *Graph Bottlenecked Social Recommendation (GBSR)* framework to tackle the social noise issue. *GBSR* is a model-agnostic social denoising framework, that aims to maximize the mutual information between the denoised social graph and recommendation labels, meanwhile minimizing it between the denoised social graph and the original one. This enables *GBSR* to learn the minimal yet sufficient social structure, effectively reducing redundant social relations and enhancing social recommendations. Technically, *GBSR* consists of two elaborate components, preference-guided social graph refinement, and HSIC-based bottleneck learning. Extensive experimental results demonstrate the superiority of the proposed *GBSR* , including high performances and good generality combined with various backbones. Our code is available at: https://github.com/yimutianyang/KDD24-GBSR.

Le Wu is the Corresponding author.

KDD '24, August 25–29, 2024, Barcelona, Spain

ACM ISBN 979-8-4007-0490-1/24/08
https://doi.org/10.1145/3637528.3671807

CCS CONCEPTS

• **Human-centered computing → Collaborative and social computing**.

KEYWORDS

Robust Social Recommendation, Social Denoising, Information Bottleneck

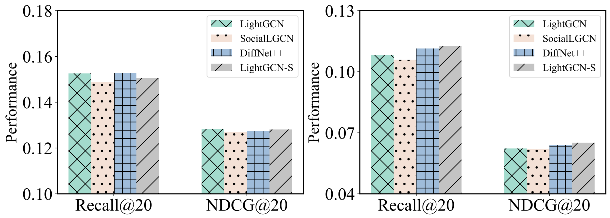
ACM Reference Format:
Yonghui Yang, Le Wu, Zihan Wang, Zhuangzhuang He, Richang Hong, and Meng Wang. 2024. Graph Bottlenecked Social Recommendation. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24), August 25–29, 2024, Barcelona, Spain.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3637528.3671807

1 INTRODUCTION

Learning informative user and item representations is the key to building modern recommender systems. Classic collaborative filtering paradigm factorizes user-item interaction matrix to learn user and item representations, which is widely researched but usually limited by sparse interactions. With the proliferation of social media, social recommendation has become an important technique to provide personalized suggestions [44]. Both user-item interactions [38, 39] and user-user social relations [54, 55] are available on social platforms, prompting the development of various social recommendation methods designed to exploit these behavior patterns [23, 26].

Following the social homophily [30] and social influence theory [29], many efforts are devoted to characterizing social relation effects on user preferences. Early works mainly focus on exploiting first-order social relations, i.e., social regularization that assumes socially connected users share similar preference [19], and social enhancement that incorporates user-trusted friends' feedback as auxiliary for the target user [15]. Recently, witnessed the power of graph neural networks (GNNs) on machine learning [5, 6, 22, 57, 60], graph-based recommendations have attracted more and more attention [4, 16, 56, 63]. Graph-based social recommendations [8, 55, 58] achieve impressive progress in improving recommendation performances by formulating users' high-order interest propagation and social influence diffusion with GNNs.

Despite the effectiveness, current graph-based social recommendations rarely notice the social noise problem, i.e., social graphs



(a) Douban-Book (b) Yelp

Figure 1: Performance comparisons between LightGCN and SOTA graph-based social recommendation methods.

are inevitably noisy with redundant social relations. Those redundant relations are caused by unreliable social relations and low preference-affinity social relations [35, 43]. Consequently, directly using the observed social graph may hinder precise user preference characterization, leading to sub-optimal recommendation results. We conduct an empirical study to illustrate the social noise problem. As shown in Figure 1, we compare LightGCN with current SOTA graph-based recommendation methods, including SocialLGCN [25] and DiffNet++ [54]. To avoid the effect of the message-passing mechanism of different methods, we additionally implement the extension of LightGCN, called LightGCN-S which additionally performs social neighbor aggregation for user representation learning. We can find that compared with LightGCN, graph-based social recommendations do not present significant strength on both metrics, even worse on the Douban-Book dataset. This indicates that social networks are usually noisy, it's necessary to filter redundant social relations to enhance the robustness of social recommendations. However, identifying and removing redundant social relations is non-trivial due to a lack of ground-truth labels. Besides, how can guarantee the recommendation accuracy while removing social relations?

In this paper, we focus on learning the denoised social graph structure to facilitate recommendation tasks from an information bottleneck perspective. Specifically, we propose a novel *Graph Bottlenecked Social Recommendation (GBSR)* framework to tackle the social noise problem. Let $\mathcal{G}^S = \{U, \mathbf{S}\}$ denote the user-user social graph and $\mathbf{R}$ denote the user-item interaction matrix, where U is userset and $\mathbf{S}$ is social structure matrix. The optimal denoised social graph structure $\mathbf{S}'$ should satisfy: *the minimal from* $\mathbf{S}$ *yet efficient for infer* $\mathbf{R}$. To achieve this goal, we first introduce user preference signals to guide the social graph denoising process, then optimize the learning process via the Information Bottleneck (IB) principle. Specifically, *GBSR* maximize the mutual information between the denoised social graph structure $\mathbf{S}'$ and interaction matrix $\mathbf{R}$, meanwhile minimizing it between the denoised social graph structure $\mathbf{S}'$ and the original $\mathbf{S}$. Therefore, the learning objective is formulated as: $Max : I(\mathbf{R}; \mathbf{S}') - \beta I(\mathbf{S}'; \mathbf{S})$.

Nevertheless, optimizing the objective of *GBSR* for social recommendation is still challenging due to the following two challenges. For the maximization of $I(\mathbf{R}; \mathbf{S}')$, social graph and sparse interaction matrix are two non-Euclidean data, which are hard to compare directly. For the minimization of $I(\mathbf{S}'; \mathbf{S})$, estimating the upper bound of MI is an intractable problem. Although some works [1, 7] leverage variational techniques to estimate the upper bound, they heavily rely on the prior assumption. To address the above two challenges, *GBSR* is implemented as follows. First, regarding the hard-comparable issue of $I(\mathbf{R}; \mathbf{S}')$, we take all nodes into intermediary and derive the lower bound of $I(\mathbf{R}; \mathbf{S}')$ for maximization. Second, we introduce the Hilbert-Schmidt independence criterion (HSIC) [28] to replace the minimization of $I(\mathbf{S}'; \mathbf{S})$. HSIC [12] is a statistic measure of variable dependency, minimizing HSIC approximate the minimization of mutual information. Our contributions are summarized as follows:

- In this paper, we revisit the social denoising recommendation from an information theory perspective, and propose a novel *Graph Bottlenecked Social Recommendation (GBSR)* framework to tackle the noise issue.
- Technically, we derive the lower bound of $I(\mathbf{R}; \mathbf{S}')$ for maximization, and introduce the Hilbert-Schmidt independence criterion (HSIC) to approximate the minimization of $I(\mathbf{S}'; \mathbf{S})$.
- Empirical studies on three benchmarks clearly demonstrate the effectiveness and generality of the proposed *GBSR* , i.e., *GBSR* achives over 17.06%, 10%, and 11.27% improvements of NDCG@20 compared with the strongest baseline.

2 PRELIMINARIES

2.1 Problem Statement

There are two kinds of entities in fundamental social recommendation scenarios: a userset U ($|U| = M$) and an itemset V ($|V| = N$). Users have two kinds of behaviors, user-user social relations and user-item interactions. We use matrix $\mathbf{S} \in \mathbb{R}^{M \times M}$ to describe user-user social structure, where each element $s_{ab} = 1$ if user b follows user a, otherwise $s_{ab} = 0$. Similar, we use matrix $\mathbf{R} \in \mathbb{R}^{M \times N}$ to describe user-item interactions, where each element $\mathbf{r}_{ai} = 1$ if user a interacted with item i, otherwise $\mathbf{r}_{ai} = 0$. Given user a, item i, and social relation matrix $\mathbf{S}$ as input, graph-based social recommenders aim to infer the probability user a will interact with item i: $\hat{r_{ai}} = \mathcal{G}_\theta(a, i, \mathbf{S})$, where $\mathcal{G}_\theta$ denotes GNN formulation. Thus, the optimization objective of graph-based social recommendation is defined as follows:

$$\theta^* = \arg\min_{\theta} \mathbb{E}_{(a,i,r_{ai})\sim\mathcal{P}} \mathcal{L}_r(r_{ai}; \mathcal{G}_\theta(a, i, \mathbf{S})), \tag{1}$$

where $\mathcal{P}$ denote distribution of training data, and θ denote GNN parameters. However, user social networks are usually noisy with redundant relations [35], directly using $\mathbf{S}$ to infer interaction probability may decrease the recommendation accuracy. In this work, we focus on learning robust social structure $\mathbf{S}'$ to facilitate recommendation performance:

$$\mathbf{S}' = \mathcal{F}_\phi(U, \mathbf{S}), \tag{2}$$

where $\mathcal{F}_\phi$ denotes social denoising function with the parameters ϕ. Consequently, the final optimization of graph-noised social recommendation is described as follows:

$$\theta^*, \phi^* = \arg\min_{\theta,\phi} \mathbb{E}_{(a,i,r_{ai})\sim\mathcal{P}} \mathcal{L}_r(r_{ai}; \mathcal{G}_\theta(a, i, \mathcal{F}_\phi(U, \mathbf{S}))). \tag{3}$$

2.2 Information Bottleneck Principle

Information Bottleneck (IB) is a representation learning principle in machine learning, which seeks a trade-off between data fit and reducing irrelevant information [45, 46]. Given input data X, Z is the hidden representation, and Y is the downstream task label, which follows the Markov Chain $< X \rightarrow Z \rightarrow Y >$. IB principle describes that an optimal representation should maintain the minimal sufficient information for the downstream tasks [37, 46]:

$$Z^* = \arg\max_{Z} I(Y;Z) - \beta I(X;Z), \tag{4}$$

where $I(Y;Z)$ denotes the mutual information between the hidden representation Z and label Y, $I(X;Z)$ denotes the mutual information between the hidden representation Z and input data X two variables, β is the coefficient to balance these two parts. IB principle has been widely applied in machine learning tasks, such as model robustness [53, 59], fairness [13], and explainability [2]. In this work, we introduce the IB principle to robust social denoising learning, which aims to seek the minimal yet sufficient social structure for recommendation tasks.

3 THE PROPOSED *GBSR* FRAMEWORK

In this section, we introduce our proposed *Graph Bottlenecked Social Recommendation (GBSR)* framework for social denoising based recommendation. Essentially, *GBSR* aims to learn the minimal yet efficient social structure to facilitate recommendation tasks, which is guaranteed by the information bottleneck principle. Next, we first give the overall optimization objective of *GBSR* , then introduce how to implement each component of *GBSR* in detail. Finally, we instantiate *GBSR* with LightGCN-S backbone.

3.1 Overview of *GBSR*

As shown in Figure 2, we present the overall objective of our proposed *GBSR* framework for the social recommendation. Instead of directly using the original social structure $\mathbf{S}$, we aim to learn a denoised yet informative social structure $\mathbf{S}'$ to enhance recommendation. Due to the lack of available prior for social denoising, we introduce user preference signals to guide social graph denoising. To guarantee the trade-off between social denoising and recommendation tasks, we optimize *GBSR* via graph information bottleneck principle. Thus, the goal of *GBSR* is: $Max : I(\mathbf{R};\mathbf{S}') - \beta I(\mathbf{S}';\mathbf{S})$. Due to the intractability of $I(\mathbf{R};\mathbf{S}')$, we take all nodes into an intermediary for calculation. Thus, we obtain the final optimization objective of *GBSR* :

$$Max : I(\mathbf{R};U,V,\mathbf{S}') - \beta I(\mathbf{S}';\mathbf{S}), \tag{5}$$

where the first term is encouraging that the denoised social graph preserves the essential information to facilitate recommendation tasks. The second term is the compression of the original social graph, aiming to filter redundant social relations.

3.2 Preference-guided Social Denoising

To achieve the above objective of *GBSR* , we first need to refine the denoised social graph. The challenge is that although the social graph has noisy relations, there are no available labels to guide the denoising process. Based on social homogeneity social-connected individuals have more similar behavior similarity, we inject user preference signals into the social denoising process, i.e., users with similar preferences are more likely to have social relations.

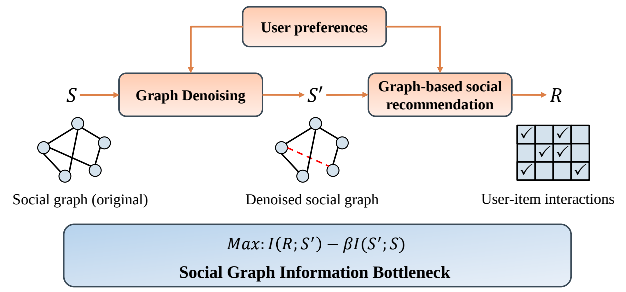


Figure 2: Overview of our proposed *GBSR* framework.

Formally, we formulate the social denoising process as a graph edge dropout problem. Given the original social graph structure $\mathbf{S}$, the denoised one is defined as:

$$\mathbf{S}' = \mathcal{F}_{\phi}(U,\mathbf{S}) = \{s_{ab} \odot \rho_{ab}\}, \tag{6}$$

in which $\rho_{ab} \sim Bern(w_{ab}) + \epsilon$ denotes that each edge $< u_a, u_b >$ will be dropped with the probability $1 - w_{ab} + \epsilon$. Here, we add the parameter $\epsilon > 0$ to represent the observation bias. Due to lacking prior information, we introduce task-relevant user preferences to refine social structure. Let $\mathbf{E}^U \in \mathbb{R}^{M\times d}$ denote user preference representations learned from the observed interactions, such as Matrix Factorization [31] and LightGCN [16]. For each observed social relation $< a, b >$, the link confidence is calculated as:

$$w_{ab} = (g(\mathbf{e}_a, \mathbf{e}_b)), \tag{7}$$

where $\mathbf{e}_a$ and $\mathbf{e}_b$ denotes user a and user b preference representations, respectively. $g()$ is the fusion function, we employ MLPs to realize it. However, $\mathbf{S}'$ is not differentiable with the parameter ρ of Bernoulli distribution, so we use the popular concrete relaxation method [20] to replace:

$$Bern(w_{ab}) = sigmoid(log(\delta/(1-\delta) + w_{ab})/t), \tag{8}$$

where $\delta \sim U(0,1)$, and $t \in \mathbb{R}^+$ is the temperature parameter (we set t=0.2 in our experiments). After re-parameterization, the discrete Bernoulli distribution is transferred to a differentiable function.

3.3 Maximization of $I(\mathbf{R};U,V,\mathbf{S}')$

Given the denoised social graph $\mathbf{S}'$, we first present how to maximize the mutual information $I(\mathbf{R};U,V,\mathbf{S}')$, which ensures the denoised social graph satisfy recommendation tasks. Specifically, we derivate the lower bound of $I(\mathbf{R};U,V,\mathbf{S}')$ as follows:

$$\begin{aligned} I(\mathbf{R};U,V,\mathbf{S}') &\overset{(a)}{=} H(\mathbf{R}) - H(\mathbf{R}|U,V,\mathbf{S}') \\ &\overset{(b)}{\geq} \sum_{a=0}^{M-1}\sum_{i=0}^{N-1}\sum_{r=0}^{1} p(r,a,i,\mathbf{S}')log(p(r|a,i,\mathbf{S}') \\ &\overset{(c)}{\geq} \sum_{(a,i,j)\in\mathcal{D}} log(p(r_{ai}=1|a,i,\mathbf{S}')) + log(p(r_{ai}=0|a,j,\mathbf{S}')) \\ &\overset{(d)}{=} \sum_{(a,i,j)\in\mathcal{D}} log(\sigma(\mathcal{G}(a,i,\mathbf{S}'))) - log(\sigma(\mathcal{G}(a,j,\mathbf{S}'))) \\ &\overset{(e)}{\geq} \sum_{(a,i,j)\in\mathcal{D}} log(\sigma(\mathcal{G}(a,i,\mathbf{S}') - \mathcal{G}(a,j,\mathbf{S}'))), \end{aligned} \tag{9}$$

where $\mathcal{G}(\cdot)$ is any graph-based social recommender as we mentioned in the preliminaries, $\sigma(\cdot)$is the sigmoid activation, $\mathcal{D} = \{(a,i,j)|r_{ai} = 1 \wedge r_{aj} = 0\}$ is all training data. Next, we introduce each derivation step as follows: (a) is the definition of mutual information; (b) is the non-negative property of $H(\mathbf{R})$; (c) is that $p(r|a,i,\mathbf{S}') \leq 1$, and we split all samples into observed interactions and non-observed interactions; (d) $\sigma(\mathcal{G}(a,i,\mathbf{S}'))$ is the variational approximation of $p(r_{ai} = 1|a,i,\mathbf{S}')$; (e) is due to $log(\sigma(x)) - log(\sigma(y)) \geq log(\sigma(x-y))$.

According to the above derivation, we can find that the popular BPR ranking loss [36] is the lower bound of mutual information $I(\mathbf{R};U,V,\mathbf{S}')$. Therefore, we employ BPR loss as the objective of mutual information maximization.

3.4 Minimization of $I(\mathbf{S}';\mathbf{S})$

Next, we introduce how to minimize $I(\mathbf{S}',\mathbf{S})$, which aims to reduce the redundant social relations in the original graph. Estimating the upper bound of mutual information is an intractable problem. Although some works [1, 7] leverage variational techniques to estimate the upper bound, but heavily rely on the prior assumption. Therefore, we introduce Hilbert-Schmidt Independence Criterion (HSIC [12]) as the approximation of the minimization of $I(\mathbf{R};\mathbf{S}')$.

HSIC brief. HSIC serves as a statistical measure of dependency [12], which is formulated as the Hilbert-Schmidt norm, assessing the cross-covariance operator between distributions within the Reproducing Kernel Hilbert Space (RKHS). Mathematically, given two variables X and Y, HSIC(X,Y) is defined as follows:

$$\begin{aligned} HSIC(X,Y) &= ||C_{XY}||^2_{hs} \\ &= \mathbb{E}_{X,X',Y,Y'}[K_X(X,X')K_Y(Y,Y')] \\ &+ \mathbb{E}_{X,X'}[K_X(X,X')]\mathbb{E}_{Y,Y'}[K_Y(Y,Y')] \\ &- 2\mathbb{E}_{XY}[\mathbb{E}_{X'}[K_X(X,X')]\mathbb{E}_{Y'}[K_Y(Y,Y')]], \end{aligned} \tag{10}$$

where K_X and K_Y are two kernel functions for variables X and Y, X' and Y' are two independent copies of X and Y. Given the sampled instances $(x_i,y_i)^n_{i=1}$ from the batch training data, the $HSIC(X,Y)$ can be estimated as:

$$H\hat{S}IC(X,Y) = (n-1)^{-2}Tr(K_XHK_YH), \tag{11}$$

where K_X and K_Y are used kernel matrices [12], with elements $K_{X_{ij}} = K_X(x_i,x_j)$ and $K_{Y_{ij}} = K_Y(y_i,y_j)$, $H = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T$ is the centering matrix, and $Tr(\cdot)$ denotes the trace of matrix. In practice, we adopt the widely used radial basis function (RBF) [48] as the kernel function:

$$K(x_i,x_j) = exp(-\frac{||x_i - x_j||^2}{2\sigma^2}), \tag{12}$$

where σ is the parameter that controls the sharpness of RBF.

HSIC-based bottleneck learning. Given the original and denoised social graph structures $\mathbf{S}$ and $\mathbf{S}'$, we minimize $HSIC(\mathbf{S}';\mathbf{S})$ to replace the minimization of $I(\mathbf{S}';\mathbf{S})$. However, social graphs are non-Euclidean data, making it difficult to measure dependency. In practice, we adopt Monte Carlo sampling [40] on all the node representations for calculation:

$$Min: HSIC(\mathbf{S}',\mathbf{S}) = H\hat{S}IC(\mathbf{E'}_{\mathcal{B}},\mathbf{E}_{\mathcal{B}}), \tag{13}$$

where $\mathcal{B}$ denotes the batch sampling users, $\mathbf{E}'$ and $\mathbf{E}$ denote node representations, which are learned from recommenders $\mathcal{G}_{\theta,\phi}(U,V,\mathbf{S}')$ and $\mathcal{G}_{\theta}(U,V,\mathbf{S})$. Thus, we can reduce the redundant social relations via the HSIC-based bottleneck regularization:

$$\mathcal{L}_{ib} = H\hat{S}IC(\mathbf{E'}_{\mathcal{B}},\mathbf{E}_{\mathcal{B}}). \tag{14}$$

3.5 Instantiating the *GBSR* Framework

In this section, we instantiate our proposed *GBSR* with specific graph-based social recommender $\mathcal{G}_\theta(U,V,\mathbf{S})$. To avoid the effect of different message-passing mechanisms, we implement LightGCN-S as the backbone model (we also realize *GBSR* with other backbones, refer to the generality analysis). Firstly, we formulate the available data and denoised social structure as a graph $\mathcal{G} = \{U \cup V, \mathbf{A}\}$, where $U \cup V$ denotes the set of nodes, and $\mathbf{A}$ is the adjacent matrix defined as follows:

$$\mathbf{A} = \begin{bmatrix} \mathbf{S} & \mathbf{R} \\ \mathbf{R}^T & \mathbf{0}^{N\times N} \end{bmatrix}. \tag{15}$$

Given the initialized node embeddings $\mathbf{E}^0 \in \mathbb{R}^{(M+N)\times d}$, LightGCN-S updates node embeddings through multiple graph convolutions:

$$\mathbf{E}^{l+1} = \mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{-\frac{1}{2}}\mathbf{E}^l, \tag{16}$$

where $\mathbf{D}$ is the degree matrix of graph $\mathcal{G}$, $\mathbf{E}^{l+1}$ and $\mathbf{E}^l$ denote node embeddings in $l+1^{th}$ and l^{th} graph convolution layer, respectively. When stacking L graph convolution layers, the final node representations can be obtained with a readout operation:

$$\mathbf{E} = Readout(\mathbf{E}^0,\mathbf{E}^1,...,\mathbf{E}^L). \tag{17}$$

After obtaining the learned node representations through GCNs, LightGCN-S infers the propensity that user a interacts with item i by an inner product: $\hat{r}_{ai} =< e_a, e_i >$. All the above process are summarized as $\hat{r}_{ai} = \mathcal{G}_\theta(a,i,\mathbf{S})$.

Next, we give the illustration of graph-denoised social recommendation. We first use the initialized node embeddings to obtain user preference representations $\mathbf{P} = \mathbf{E}^0[:M]$, then achieve the denoised social structure $\mathbf{S}'$ based on preference-guided social structure learning (section 3.2). Given the learned denoised social structure $\mathbf{S}'$, we establish graph-denoised social recommender $\hat{r}_{ai} = \mathcal{G}_{\theta,\phi}(a,i,\mathbf{S}')$. Then, we select the pairwise ranking loss [36] to optimize model parameters:

$$\mathcal{L}_{rec} = \sum_{a=0}^{M-1}\sum_{(i,j)\in D_a} -log\sigma(\hat{r}_{ai} - \hat{r}_{aj}) + \lambda||\mathbf{E}^0||^2, \tag{18}$$

where $\sigma(\cdot)$ is the sigmoid activation function, λ is the regularization coefficient. $D_a = \{(i,j)|i \in R_a \wedge j \notin R_a\}$ denotes the pairwise training data for user a. R_a represents user a interacted items on the training data. Combined with the HSIC-based bottleneck regularizer, we obtain the final optimization objective:

$$\arg\min_{\theta,\phi} \mathcal{L} = \mathcal{L}_{rec} + \beta\mathcal{L}_{ib}, \tag{19}$$

The overall learning process of *GBSR* is illustrated in Algorithm 1.

3.6 Model Discussion

In this section, we analyze the proposed *GBSR* from model complexity and model generalization.

Algorithm 1: The algorithm of *GBSR*

Data: Userset U, Itemset V, User-item interactions $\mathbf{R}$, User-user social relations $\mathbf{S}$, and observation bias ϵ

Result: Optimal graph-denoised social recommender $\mathcal{G}^*_{\theta,\phi}(\cdot)$

Initialize recommender $\mathcal{G}_{\theta,\phi}$ with random weights;
while *not converged* **do**
    Sample a batch training data $\mathcal{D}$;
    Compute social edge dropout probability via Eq.(7)-Eq.(8);
    Refine the denoised social structure $\mathbf{S}'$ via Eq.(6);
    Obtain node representations $\mathbf{E}'$ via $\mathcal{G}_{\theta,\phi}(U, V, \mathbf{S}')$;
    Obtain node representations $\mathbf{E}$ via $\mathcal{G}_{\theta}(U, V, \mathbf{S})$;
    Compute recommendation task loss $\mathcal{L}_{rec}$ via Eq.(18);
    Compute HSIC bottleneck loss $\mathcal{L}_{ib}$ via Eq.(14);
    Update model parameters according to Eq.(19);
end
Return the optimal $\mathcal{G}^*_{\theta,\phi}(\cdot)$;

3.6.1 Space Complexity. As illustrated in Algorithm 1, the parameters of *GBSR* are composed of two parts: graph-based social recommender parameters θ and social denoising parameters ϕ. Among them, $\theta = \mathbf{E}^0$ are the general parameters equipped for backbone models (such as LightGCN-S). ϕ are the parameters of MLPs, which are used to calculate the social edge confidence. Because ϕ are the shared parameters for all social edges, the additional parameters of *GBSR* are ignorable compared with backbone models.

3.6.2 Time Complexity. Compared with the backbone model (such as LightGCN-S), the additional time cost is social graph denoising and HSIC-bottleneck optimization. Social graph denoising is conducted on the observed social relations, which performs a sparse matrix. Besides, the time complexity of the HSIC-bottleneck regularizer lies in the number of the sampled nodes (refer to Eq.(13)). In practice, we adopt a mini-batch training strategy to reduce the time cost of bottleneck learning, and the additional time cost of *GBSR* is affordable. Besides, as we remove redundant social relations, the denoised yet informative social graph makes *GBSR* convergence much faster than the backbone model. Experiments also verify the efficiency of *GBSR*.

3.6.3 Model Generalization. The proposed *GBSR* is designed for social denoising under graph-based social recommendation scenarios. It does not depend on specific graph-based social recommenders, such as DiffNet++ [54] and SocialLGN [25]. Our proposed *GBSR* is a flexible denoising framework to enhance social recommendations, we also conduct experiments on four backbones to demonstrate the generalization. Besides the backbone model, the idea of introducing the information bottleneck principle to graph denoising can also be generalized for different recommendation scenarios.

4 EXPERIMENTS

In this section, we conduct extensive experiments on three real-world datasets to validate the effectiveness of our proposed *GBSR* .

Table 1: The statistics of three datasets.

Dataset	Douban-Book	Yelp	Epinions
Users	13,024	19,593	18,202
Items	22,347	21,266	47,449
Interactions	792,062	450,884	298,173
Social Relations	169,150	864,157	381,559
Interaction Density	0.048%	0.034%	0.035%
Relation Density	0.268%	0.206%	0.115%

We first introduce experimental settings, followed by recommendation performance comparisons. Finally, we give a detailed model investigation, including training efficiency, visualization of the denoised social graph, and parameter sensitivities.

4.1 Experimental Settings

4.1.1 Datasets. We conduct empirical studies on three public datasets to verify the effectiveness of our proposed *GBSR* , including Yelp, Epinions, and Dianping [62]. All datasets involve user-user social links and user-item interactions. For the Yelp dataset, we follow the released version in [65]. For epinions and dianping datasets, we filter ratings less than 3 and keep the remaining ratings as positive feedback. After that, we sample 80% interactions as training data, and the remaining 20% as test data. The detailed statistics of all datasets are summarized in Table 1.

4.1.2 Baselines and Evaluation Metrics. To evaluate the effectiveness of our proposed *GBSR* , we select state-of-the-art baselines for comparisons. Specifically, these baselines can be divided into two groups: graph-based social recommendation methods [8, 16, 54] and social graph denoising methods [35, 66, 69], which are list as follows:

- **LightGCN** [16]: is the SOTA graph-based collaborative filtering method, which simplifies GCNs by removing the redundant feature transformation and non-linear activation components for ID-based recommendation.
- **LightGCN-S**: We extend LightGCN to graph-based social recommendation, that each user's neighbors include their interacted items and linked social users. LightGCN-S is a basic and lightweight model, considering our proposed *GBSR* is a model-agnostic social graph denoising method, we select LightGCN-S as the backbone model.
- **GraphRec** [8]: is a classic graph-based social recommendation method, it incorporates user opinions and user two kinds of graphs for preference learning.
- **DiffNet++** [54]: is the SOTA graph-based social recommendation method, it recursively formulates user interest propagation and social influence diffusion process with a hierarchical attention mechanism.
- **SocialLGN** [25]: propagates user representations on both user-item interactions graph and user-user social graph with light graph convolutional layers, and fuses them for recommendation.
- **Rule-based**: We follow [35] and remove unreliable social relations based on the similarity of the user-interacted items.

Table 2: Overall performance comparisons on three benchmarks. The best performance is highlighted in bold and the second is highlighted by underlines. Impro. indicates the relative improvement of our proposed *GBSR* compared to the best baseline.

Models	Douban-Book				Yelp				Epinions			
	R@10	N@10	R@20	N@20	R@10	N@10	R@20	N@20	R@10	N@10	R@20	N@20
LightGCN	0.1039	0.1195	0.1526	0.1283	0.0698	0.0507	0.1081	0.0623	0.0432	0.0314	0.0675	0.0385
GraphRec	0.0971	0.1145	0.1453	0.1237	0.0672	0.0485	0.1077	0.0607	0.0436	0.0315	0.0681	0.0387
DiffNet++	0.1010	0.1184	0.1489	0.1270	0.0707	0.0516	0.1114	0.0640	0.0468	0.0329	0.0727	0.0406
SocialLGN	0.1034	0.1182	0.1527	0.1274	0.0681	0.0507	0.1059	0.0620	0.0416	0.0307	0.0634	0.0371
LightGCN-S	0.1021	0.1187	0.1506	0.1281	0.0714	0.0529	0.1126	0.0651	0.0477	0.0347	0.0716	0.0417
Rule-based	0.1033	0.1192	0.1518	0.1289	0.0705	0.0526	0.1126	0.0652	0.0465	0.0340	0.0716	0.0414
ESRF	0.1042	0.1199	0.1534	0.1301	0.0718	0.0526	0.1123	0.0645	0.0462	0.0329	0.0727	0.0406
GDMSR	0.1026	0.1001	0.1538	0.1245	0.0739	0.0535	0.1148	0.0658	0.0461	0.0326	0.0721	0.0414
GBSR	**0.1189**	**0.1451**	**0.1694**	**0.1523**	**0.0805**	**0.0592**	**0.1243**	**0.0724**	**0.0529**	**0.0385**	**0.0793**	**0.0464**
Impro.	**14.11%**	**21.02%**	**10.14%**	**17.06%**	**8.93%**	**10.65%**	**8.28%**	**10.00%**	**10.90%**	**10.95%**	**9.08%**	**11.27%**

- **ESRF** [66]: proposes adversarial graph convolutional networks to enhance social recommendation, it generate alternative social neighbors and further perform neighbor denoising with adversarial training.
- **GDMSR** [35]: designs the robust preference-guided social denoising to enhance graph-based social recommendation, it only remains the informative social relations according to preference confidences.

As we focus on implicit recommendation scenarios, we employ two widely used ranking metrics: Recall@N and NDCG@N [14, 41]. Specifically, Recall@N measures the percentage of the recalled positive samples for the Top-N ranking lists. Furthermore, NDCG@N assigns higher scores for those items in the top-ranked positions. In the evaluation stage, we adopt a full-ranking strategy that views all non-interacted items as candidates to avoid biased evaluation [24, 68]. For each model, we repeat experiments in 5 times and report the average values.

4.1.3 Parameter Settings. We implement our proposed *GBSR* and backbone with Tensorflow [1]. For all baselines, we follow the original settings and carefully fine-tune parameters for fair comparisons. For latent embedding based methods, we initialize their embeddings with a Gaussian distribution with a mean value of 0 and a standard variance of 0.01, and fix the embedding size to 64. For model optimization, we use Adam optimizer with a learning rate of 0.001 and a batch size of 2048. We follow the mainstream ranking-based methods [36], and randomly select 1 non-interacted item as the negative sample for pairwise ranking optimization. We search the GCN layer in $[1, 2, 3, 4]$, the regularization parameter λ in $[0.0001, 0.001, 0.01]$. For the observation bias, we set $\epsilon = 0.5$ for all datasets. For information bottleneck constraint coefficient β, we use grid-search with different scales over three datasets, and report detailed analysis in experiments.

4.2 Recommendation Performances

4.2.1 Overall Comparisons with Baselines. As shown in Table 2, we compare our proposed *GBSR* with state-of-the-art methods on three benchmarks. For a fair comparison, all denoising methods are conducted on the LightGCN-S backbone. Given the empirical studies, we have the following observations:

- Compared with LightGCN, graph-based social recommendation methods present slight improvements under most of the datasets, i.e., DiffNet++ obtains a 2.24% improvement on the NDCG@20 metric for Yelp dataset. However, this is not always the case, all social graph recommendations show a performance degradation on the Douban-Book dataset. While supported by social graphs, it is noteworthy that graph-based social recommendation methods do not consistently outperform LightGCN in terms of performance. These demonstrate that directly using social graphs may decrease recommendation performance, it's necessary to remove redundant social relations to enhance recommendation.
- Compared with directly using original social graphs, social denoising methods present better performances in most cases. This indicates that social noise is ubiquitous in real-world recommendation scenarios. All social denoising methods are implemented on LightGCN-S backbone, we find that GDMSR is the strongest baseline, which benefits from preference-guided social denoising and self-correcting curriculum learning. However, these social denoising methods don't present large-margin improvements compared with the backbone model. The reason is that simple rule or assumption based denoising methods lack of theoretical guarantee, it's hard to seek an effective trade-off between social denoising and recommendation accuracy.
- Our proposed *GBSR* consistently outperforms all baselines under all experimental settings. Specifically, *GBSR* improves the strongest baseline $w.r.t$ NDCG@20 by 17.06%, 10% and 11.27% on Douban-Book, Yelp, and Epinions datasets, respectively. Compared with the backbone model, *GBSR* achieves impressive superiority over three benchmarks. These indicate that our proposed *GBSR* can significantly improve graph-based social recommendations, demonstrating the effectiveness of graph bottleneck learning to reduce redundant social relations. Compared with other social denoising methods, our *GBSR* can better obtain the trade-off between removing social relations and recommendation tasks.

[1]https://www.tensorflow.org

Table 3: Performance comparisons of *GBSR* on different backbones.

Models	Douban-Book		Yelp		Epinions	
	Recall@20	NDCG@20	Recall@20	NDCG@20	Recall@20	NDCG@20
GraphRec	0.1453	0.1237	0.1077	0.0607	0.0681	0.0387
+GBSR	0.1510(+3.92%)	0.1306(+5.58%)	0.1136(+5.48%)	0.0662(+9.06%)	0.0706(+3.67%)	0.0403(+4.13%)
DiffNet++	0.1489	0.1270	0.1114	0.0640	0.0727	0.0406
+GBSR	0.1545(+3.76%)	0.1334(+5.04%)	0.1224(+9.87%)	0.0721(+12.66%)	0.0790+(+8.67%)	0.0451(+11.08%)
SocialLGN	0.1537	0.1274	0.1059	0.0620	0.0634	0.0371
+GBSR	0.1591(+3.51%)	0.1353(+6.20%)	0.1152(+8.78%)	0.0675(+8.87%)	0.0675(+6.47%)	0.0399(+7.55%)
LightGCN-S	0.1506	0.1281	0.1126	0.0651	0.0716	0.0417
+GBSR	0.1694(+12.48%)	0.1523(+18.89%)	0.1243(+10.39%)	0.0724(+11.21%)	0.0793(+10.75%)	0.0464(+11.27%)

4.2.2 Ablation study. We conduct ablation studies on three datasets to explore the effectiveness of each component of the proposed *GBSR* framework. As shown in Table 4, we compare *GBSR* with corresponding variants on Top-20 recommendation performances. *GBSR*-w/o HSIC denotes that remove the HSIC-based bottleneck regularization of *GBSR* , we only keep preference-guided social denoising module. From Table 4, we can find that GBSR-w/o HSIC performs worse in all cases, even worse than the backbone model. This indicates that simple social structure learning without HSIC-based bottleneck regularization is useless for recommendation tasks. Furthermore, under the constraint of the information bottleneck principle, the learned social structure is meaningful, which can effectively improve social recommendations on three datasets.

4.2.3 Generality study of GBSR *.* As we mentioned in the model discussion, the proposed *GBSR* is a model-agnostic social denoising framework. To better illustrate the generality of *GBSR* , we conduct experiments of *GBSR* on several graph-based social recommendation backbones. As shown in Table 3, we implement *GBSR* under four backbones, including GraphRec [8], DiffNet++[54], SocialLGN [25], and LightGCN-S, and report their performances of Top-20 recommendation task. From Table 3, we observe that *GBSR* consistently outperforms each backbone by a large margin. For example, on the Yelp dataset, *GBSR* achieves 9.06%, 12.66%, 8.87%, and 11.21% improvements of NDCG@20 compared with GraphRec, DiffNet++, SocialLGN, and LightGCN-S, respectively. Similarly, *GBSR* also obtains 5.48%, 9.87%, 8.78%, and 10.39% improvements on the Recall@20 metric. Extensive experimental results show that our proposed *GBSR* has a good generalization ability, which can easily coupled with current graph-based social recommendation methods and further enhancement.

4.3 Investigation of *GBSR*

In this section, we further analyze *GBSR* from the following aspects: training efficiency, visualization of the denoised social graphs, and hyper-parameter sensitivity analysis.

4.3.1 Training efficiency of GBSR. To analyze the training efficiency of *GBSR* , we compare the convergence speed of *GBSR* and corresponding backbone (LightGCN-S). As shown in Figure 3, we compare the convergence process of both models. As the space limit, we only present the convergence process on Douban-Book and Yelp datasets. We set gcn layer to 3 and keep all experimental settings the same. According to these figures, we can observe that

Table 4: Ablation study on three datasets.

Models	Douban-Book		Yelp		Epinions	
	R@20	N@20	R@20	N@20	R@20	N@20
LightGCN-S	0.1506	0.1281	0.1126	0.0651	0.0716	0.0417
GBSR-w/o HSIC	0.1482	0.1259	0.1119	0.0644	0.0688	0.0388
GBSR	0.1694	0.1523	0.1243	0.0724	0.0793	0.0464

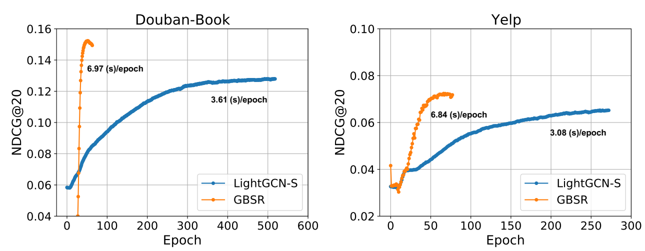
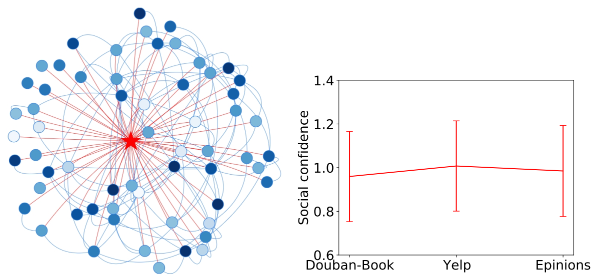


Figure 3: Convergence curves of *GBSR* and LightGCN-S on Douban-Book and Yelp datasets.

GBSR converges much faster than the backbone model. Particularly, *GBSR* reaches the best performances at the 82^{th}, the 67^{th} epoch on Douban-Book and Yelp datasets. In contrast, LightGCN-S obtains the best results on 509^{th}, and 261^{th} epoch, respectively. Empirical evidence shows that *GBSR* convergence 2-3 times faster than LightGCN-S.

4.3.2 Visualization and statistics of the denoised social graphs. Here we first present the visualization of the denoised social graph. As shown in Figure 4(a), we present the sampled ego-network from Douban-Book datasets. The red node denotes the center user of this ego-network, and the blue nodes denote social neighbors. The depth of the node color denotes the probability of edge dropping, where the darker the color, the lower the dropping probability. We can observe that user social neighbors perform different confidences of social relations. Besides, we analyze the statistics of the denoised social graphs. As shown in Figure 4(b), we plot the mean and variance values of social relation confidence on three datasets. We can observe that Douban-Book presents the lowest mean value of social confidence, which means that it has the most social noise over the three datasets. This also explains the results of Figure 1 that graph-based social recommendations show a performance decrease



(a) Ego-network (from Douban-Book) (b) Errorbar of social relation confidence

Figure 4: Visualization and statistics of the denoised social graphs.

compared with LightGCN on the Douban-Book dataset. These results demonstrate that the proposed *GBSR* can effectively refine the observed social graph via information bottleneck, which provides informative social structures to enhance social recommendations.

4.3.3 Parameter Sensitivity Analysis. In this part, we analyze the impact of different hyper-parameters of *GBSR* . There are two key parameters, bottleneck loss coefficient β and RBF sharpness parameter σ^2. As both parameters determine the scale of bottleneck loss, we combine them to analyze the influence of recommendation results. As shown in Figure 5, we conduct careful grid-search of (β, σ^2) on three datasets. We can observe that *GBSR* reaches the best performance when $\beta = 40, \sigma^2 = 2.5$ on Douban-Book, $\beta = 2.0, \sigma^2 = 0.25$ on Yelp, and $\beta = 3.0, \sigma^2 = 0.25$, respectively.

5 RELATED WORKS

5.1 Graph-based Social Recommendation

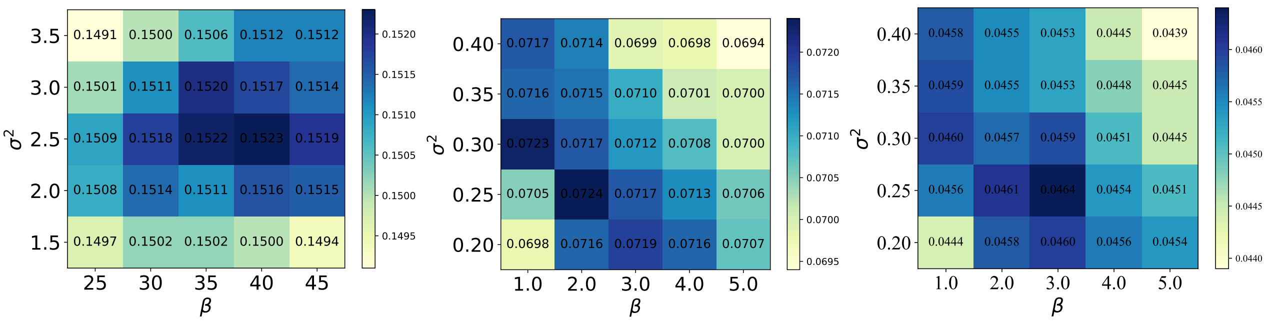
With the emergence of social media, social recommendation has been an important technique and has attracted more and more research attention [19, 23, 26, 27, 44]. Following the social homophily [30] and social influence theory [29], social recommendations are devoted to characterizing social relation effects on user preferences. Early efforts exploit social relations in a shallow form, such as co-factorization methods [23, 26] and regularization-based methods [19, 21, 27]. For example, SoRec [26] jointly co-factorize the interaction and social matrices and then project interaction and social contexts into the same semantic space. [27] designs a social regularization term that assumes two socially connected users should be closer in preference space. Recently, with the great success of graph neural networks [22, 47], graph-based social recommendations have been widely researched and achieved impressive process [8, 25, 54, 55, 62, 67]. By formulating user-user social relations as a graph, graph-based social recommendations inject high-order social influences into user preference learning, vibrant the representation ability. For example, DiffNet models the high-order social influence diffusion process to enhance user representation [55], and DiffNet++ further improves it by combining both social influence diffusion and user-item interest propagation with a hierarchical attention mechanism [54]. Inspired by the architecture of LightGCN, [25] proposes SocialLGN to model user interaction and social behaviors. Instead of learning social graphs on Euclidean space, some works attempt to introduce hyperbolic learning for graph-based social recommendations [50, 62]. Despite the effectiveness of modeling high-order social influence to improve recommendation, these works are built on the clean social relation assumption. However, social graphs are inevitably noisy with redundant relations, and these graph-based social recommendation methods are usually far from satisfactory. Instead of directly using the original social graph, in this work, we propose a graph noising framework to improve social recommendation.

5.2 Recommendation Denoising

Recommendation denoising works mainly focus on implicit feedback, which aims to refine implicit feedback to build robust recommender systems [11, 17, 51, 52, 61]. Most efforts are devoted to removing noise feedback, which is easily vulnerable to users' unconscious behaviors and various biases. For example, [51] proposes to drop noisy feedback based on the observation that noisy feedback has higher training loss, [52] devises a bi-level optimization method to implement recommendation denoising. Besides, graph augmentation methods are proposed to realize recommendation denoising [9, 61]. Different from the above feedback-based denoising works, we focus on social denoising for recommendations. Social graphs are inevitably noisy with redundant relations, including unreliable relations and low preference-affinity relations [35, 43]. Early works employ statistics to identify unstable social relations [27, 33], or model different user influences with attention mechanism [42, 54]. Besides, fine-grained social leveraging [10] and adversarial learning based methods have been proposed [64, 66]. Recently, GDMSR [35] proposes a distilled social graph based on progressive preference-guided social denoising. Nevertheless, the above methods still face the challenge of lacking ground-truth. Whether rule-based or assumption-based social denoising is hard to guarantee the trade-off between social denoising and social recommendation. Distinguished by these denoising methods, we address the social denoising recommendation from a novel information bottleneck perspective, which seeks the denoised yet informative social structure to enhance recommendations.

5.3 Information Bottleneck and Applications

Information Bottleneck (IB) is an effective representation learning principle in machine learning tasks, that the optimal representation should satisfy the minimal yet efficient manner [45, 46]. In the era of deep learning, calculating high-dimensional variables' mutual information (MI) is the key challenge for IB. The general solution is estimating the upper/lower bounds instead of directly calculating mutual information [1, 7]. Specifically, VIB [1] leverages the variational technique to estimate the bounds of mutual information. Besides, MINE [3], InfoNCE [32] are proposed to estimate the lower bound of MI. In contrast, a few attempts propose to estimate the upper bound of MI [7, 18]. Besides optimizing the bounds of MI, HSIC-based methods [28, 53] are proposed to implement IB learning, which employs the Hilbert-Schmidt Independence Criterion (HSIC) to replace mutual information for optimization. HSIC measures the independence of two variables, which can approximate the mutual



(a) Douban-Book dataset **(b) Yelp dataset** **(c) Epinions dataset**

Figure 5: Performance comparisons under different parameters (β, σ^2)**.**

information objective [12]. IB principle has been successfully applied to many applications, such as image classification [49], text understanding [34], and graph learning [59]. In this work, we introduce the HSIC-based bottleneck to the graph-denoised social recommendation, aiming to filtering redundant social relations for robust recommendation.

6 CONCLUSION

In this paper, we investigate graph-denoised social recommendations and propose a novel *Graph Bottlenecked Social Recommendation (GBSR)* framework. Specifically, *GBSR* aims to learn the denoised yet informative social structure for recommendation tasks. To achieve this goal, we first design preference-guided social denoising, then optimize the denoising process via the information bottleneck principle. Particularly, we derive the lower bound of mutual information maximization and introduce HSIC regularization to replace mutual information minimization. Extensive experiments conducted on three benchmarks demonstrate the effectiveness of our proposed *GBSR* framework, i.e., over 10% improvements on Top-20 Recommendation. Moreover, *GBSR* is a model-agnostic framework, which can be flexibly coupled with various graph-based social recommenders. In the future, we will explore more potential of leveraging the IB principle to recommendation tasks, i.e., self-supervised recommendation, fairness-aware recommendation, and LLM-enhanced recommendation.

ACKNOWLEDGMENTS

This work was supported in part by grants from the National Key Research and Development Program of China(Grant No.2021ZD0111802), and the National Natural Science Foundation of China(Grant No. U23B2031, 721881011).

REFERENCES

[1] Alexander A Alemi, Ian Fischer, Joshua V Dillon, and Kevin Murphy. 2017. Deep variational information bottleneck. In *ICLR*.
[2] Seojin Bang, Pengtao Xie, Heewook Lee, Wei Wu, and Eric Xing. 2021. Explaining a black-box by using a deep variational information bottleneck approach. In *AAAI*, Vol. 35. 11396–11404.
[3] Mohamed Ishmael Belghazi, Aristide Baratin, Sai Rajeshwar, Sherjil Ozair, Yoshua Bengio, Aaron Courville, and Devon Hjelm. 2018. Mutual information neural estimation. In *ICML*. PMLR, 531–540.
[4] Miaomiao Cai, Lei Chen, Yifan Wang, Haoyue Bai, Peijie Sun, Le Wu, Min Zhang, and Meng Wang. 2024. Popularity-Aware Alignment and Contrast for Mitigating Popularity Bias. *arXiv preprint arXiv:2405.20718* (2024).
[5] Miaomiao Cai, Min Hou, Lei Chen, Le Wu, Haoyue Bai, Yong Li, and Meng Wang. 2024. Mitigating Recommendation Biases via Group-Alignment and Global-Uniformity in Representation Learning. *ACM Transactions on Intelligent Systems and Technology* (2024).
[6] Lei Chen, Le Wu, Kun Zhang, Richang Hong, Defu Lian, Zhiqiang Zhang, Jun Zhou, and Meng Wang. 2023. Improving recommendation fairness via data augmentation. In *WWW*. 1012–1020.
[7] Pengyu Cheng, Weituo Hao, Shuyang Dai, Jiachang Liu, Zhe Gan, and Lawrence Carin. 2020. Club: A contrastive log-ratio upper bound of mutual information. In *ICML*. PMLR, 1779–1788.
[8] Wenqi Fan, Yao Ma, Qing Li, Yuan He, Eric Zhao, Jiliang Tang, and Dawei Yin. 2019. Graph neural networks for social recommendation. In *WWW*. 417–426.
[9] Ziwei Fan, Ke Xu, Zhang Dong, Hao Peng, Jiawei Zhang, and Philip S Yu. 2023. Graph collaborative signals denoising and augmentation for recommendation. In *SIGIR*. 2037–2041.
[10] Bairan Fu, Wenming Zhang, Guangneng Hu, Xinyu Dai, Shujian Huang, and Jiajun Chen. 2021. Dual side deep context-aware modulation for social recommendation. In *WWW*. 2524–2534.
[11] Yunjun Gao, Yuntao Du, Yujia Hu, Lu Chen, Xinjun Zhu, Ziquan Fang, and Baihua Zheng. 2022. Self-guided learning to denoise for robust recommendation. In *SIGIR*. 1412–1422.
[12] Arthur Gretton, Olivier Bousquet, Alex Smola, and Bernhard Schölkopf. 2005. Measuring statistical dependence with Hilbert-Schmidt norms. In *International conference on algorithmic learning theory*. Springer, 63–77.
[13] Adam Gronowski, William Paul, Fady Alajaji, Bahman Gharesifard, and Philippe Burlina. 2023. Classification utility, fairness, and compactness via tunable information bottleneck and Rényi measures. *IEEE Transactions on Information Forensics and Security* (2023).
[14] Asela Gunawardana and Guy Shani. 2009. A survey of accuracy evaluation metrics of recommendation tasks. *Journal of Machine Learning Research* 10, 12 (2009).
[15] Guibing Guo, Jie Zhang, and Neil Yorke-Smith. 2015. Trustsvd: Collaborative filtering with both the explicit and implicit influence of user trust and of item ratings. In *AAAI*, Vol. 29.
[16] Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, Yongdong Zhang, and Meng Wang. 2020. Lightgcn: Simplifying and powering graph convolution network for recommendation. In *SIGIR*. 639–648.
[17] Zhuangzhuang He, Yifan Wang, Yonghui Yang, Peijie Sun, Le Wu, Haoyue Bai, Jinqi Gong, Richang Hong, and Min Zhang. 2024. Double Correction Framework for Denoising Recommendation. *arXiv preprint arXiv:2405.11272* (2024).
[18] Michal Hledík, Thomas R Sokolowski, and Gašper Tkačik. 2019. A tight upper bound on mutual information. In *2019 IEEE ITW*. IEEE, 1–5.
[19] Mohsen Jamali and Martin Ester. 2010. A matrix factorization technique with trust propagation for recommendation in social networks. In *Recsys*. 135–142.
[20] Eric Jang, Shixiang Gu, and Ben Poole. 2017. Categorical reparameterization with gumbel-softmax. *ICLR* (2017).
[21] Meng Jiang, Peng Cui, Fei Wang, Wenwu Zhu, and Shiqiang Yang. 2014. Scalable recommendation with social contextual information. *IEEE TKDE* 26, 11 (2014), 2789–2802.
[22] Thomas N Kipf and Max Welling. 2017. Semi-supervised classification with graph convolutional networks. In *ICLR*.
[23] Ioannis Konstas, Vassilios Stathopoulos, and Joemon M Jose. 2009. On social networks and collaborative recommendation. In *SIGIR*. 195–202.
[24] Walid Krichene and Steffen Rendle. 2020. On sampled metrics for item recommendation. In *SIGKDD*. 1748–1757.

[25] Jie Liao, Wei Zhou, Fengji Luo, Junhao Wen, Min Gao, Xiuhua Li, and Jun Zeng. 2022. SocialLGN: Light graph convolution network for social recommendation. *Information Sciences* 589 (2022), 595–607.

[26] Hao Ma, Haixuan Yang, Michael R Lyu, and Irwin King. 2008. Sorec: social recommendation using probabilistic matrix factorization. In *CIKM*. 931–940.

[27] Hao Ma, Dengyong Zhou, Chao Liu, Michael R Lyu, and Irwin King. 2011. Recommender systems with social regularization. In *WSDM*. 287–296.

[28] Wan-Duo Kurt Ma, JP Lewis, and W Bastiaan Kleijn. 2020. The HSIC bottleneck: Deep learning without back-propagation. In *AAAI*, Vol. 34. 5085–5092.

[29] Peter V Marsden and Noah E Friedkin. 1993. Network studies of social influence. *Sociological Methods & Research* 22, 1 (1993), 127–151.

[30] Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology* 27, 1 (2001), 415–444.

[31] Andriy Mnih and Russ R Salakhutdinov. 2007. Probabilistic matrix factorization. *NeurIPS* 20 (2007).

[32] Aaron van den Oord, Yazhe Li, and Oriol Vinyals. 2018. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748* (2018).

[33] Yiteng Pan, Fazhi He, and Haiping Yu. 2020. A correlative denoising autoencoder to model social influence for top-N recommender system. *Frontiers of Computer science* 14 (2020), 1–13.

[34] Mandela Patrick, Po-Yao Huang, Yuki Asano, Florian Metze, Alexander Hauptmann, Joao Henriques, and Andrea Vedaldi. 2021. Support-set bottlenecks for video-text representation learning. *ICLR* (2021).

[35] Yuhan Quan, Jingtao Ding, Chen Gao, Lingling Yi, Depeng Jin, and Yong Li. 2023. Robust Preference-Guided Denoising for Graph based Social Recommendation. In *WWW*. 1097–1108.

[36] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *UAI*. 452–461.

[37] Andrew M Saxe, Yamini Bansal, Joel Dapello, Madhu Advani, Artemy Kolchinsky, Brendan D Tracey, and David D Cox. 2019. On the information bottleneck theory of deep learning. *Journal of Statistical Mechanics: Theory and Experiment* 2019, 12 (2019), 124020.

[38] Pengyang Shao, Le Wu, Lei Chen, Kun Zhang, and Meng Wang. 2022. FairCF: Fairness-aware collaborative filtering. *Science China Information Sciences* 65, 12 (2022), 222102.

[39] Pengyang Shao, Le Wu, Kun Zhang, Defu Lian, Richang Hong, Yong Li, and Meng Wang. 2024. Average User-side Counterfactual Fairness for Collaborative Filtering. *ACM TOIS* (2024).

[40] Alexander Shapiro. 2003. Monte Carlo sampling methods. *Handbooks in operations research and management science* 10 (2003), 353–425.

[41] Harald Steck. 2013. Evaluation of recommendations: rating-prediction and ranking. In *RecSys*. 213–220.

[42] Peijie Sun, Le Wu, and Meng Wang. 2018. Attentive recurrent social recommendation. In *SIGIR*. 185–194.

[43] Youchen Sun. 2023. Denoising Explicit Social Signals for Robust Recommendation. In *RecSys*. 1344–1348.

[44] Jiliang Tang, Xia Hu, and Huan Liu. 2013. Social recommendation: a review. *Social Network Analysis and Mining* 3 (2013), 1113–1133.

[45] Naftali Tishby, Fernando C Pereira, and William Bialek. 2000. The information bottleneck method. *arXiv preprint physics/0004057* (2000).

[46] Naftali Tishby and Noga Zaslavsky. 2015. Deep learning and the information bottleneck principle. In *2015 IEEE ITW*. IEEE, 1–5.

[47] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2018. Graph attention networks. In *ICLR*.

[48] Jean-Philippe Vert, Koji Tsuda, and Bernhard Schölkopf. 2004. A primer on kernel methods. *Kernel methods in computational biology* 47 (2004), 35–70.

[49] Bowen Wang, Liangzhi Li, Yuta Nakashima, and Hajime Nagahara. 2023. Learning Bottleneck Concepts in Image Classification. In *CVPR*. 10962–10971.

[50] Hao Wang, Defu Lian, Hanghang Tong, Qi Liu, Zhenya Huang, and Enhong Chen. 2021. Hypersorec: Exploiting hyperbolic user and item representations with multiple aspects for social-aware recommendation. *TOIS* 40, 2 (2021), 1–28.

[51] Wenjie Wang, Fuli Feng, Xiangnan He, Liqiang Nie, and Tat-Seng Chua. 2021. Denoising implicit feedback for recommendation. In *WSDM*. 373–381.

[52] Zongwei Wang, Min Gao, Wentao Li, Junliang Yu, Linxin Guo, and Hongzhi Yin. 2023. Efficient bi-level optimization for recommendation denoising. In *SIGKDD*. 2502–2511.

[53] Zifeng Wang, Tong Jian, Aria Masoomi, Stratis Ioannidis, and Jennifer Dy. 2021. Revisiting hilbert-schmidt information bottleneck for adversarial robustness. *NeurIPS* 34 (2021), 586–597.

[54] Le Wu, Junwei Li, Peijie Sun, Richang Hong, Yong Ge, and Meng Wang. 2020. Diffnet++: A neural influence and interest diffusion network for social recommendation. *TKDE* (2020).

[55] Le Wu, Peijie Sun, Yanjie Fu, Richang Hong, Xiting Wang, and Meng Wang. 2019. A neural influence diffusion model for social recommendation. In *SIGIR*. 235–244.

[56] Le Wu, Yonghui Yang, Lei Chen, Defu Lian, Richang Hong, and Meng Wang. 2020. Learning to transfer graph embeddings for inductive graph based recommendation. In *SIGIR*. 1211–1220.

[57] Le Wu, Yonghui Yang, Kun Zhang, Richang Hong, Yanjie Fu, and Meng Wang. 2020. Joint item recommendation and attribute inference: An adaptive graph convolutional network approach. In *SIGIR*. 679–688.

[58] Shiwen Wu, Fei Sun, Wentao Zhang, Xu Xie, and Bin Cui. 2022. Graph neural networks in recommender systems: a survey. *Comput. Surveys* 55, 5 (2022), 1–37.

[59] Tailin Wu, Hongyu Ren, Pan Li, and Jure Leskovec. 2020. Graph information bottleneck. *NeurIPS* 33 (2020), 20437–20448.

[60] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and S Yu Philip. 2020. A comprehensive survey on graph neural networks. *IEEE TNNLS* 32, 1 (2020), 4–24.

[61] Yonghui Yang, Le Wu, Richang Hong, Kun Zhang, and Meng Wang. 2021. Enhanced graph learning for collaborative filtering via mutual information maximization. In *SIGIR*. 71–80.

[62] Yonghui Yang, Le Wu, Kun Zhang, Richang Hong, Hailin Zhou, Zhiqiang Zhang, Jun Zhou, and Meng Wang. 2023. Hyperbolic Graph Learning for Social Recommendation. *IEEE TKDE* (2023).

[63] Yonghui Yang, Zhengwei Wu, Le Wu, Kun Zhang, Richang Hong, Zhiqiang Zhang, Jun Zhou, and Meng Wang. 2023. Generative-contrastive graph learning for recommendation. In *SIGIR*. 1117–1126.

[64] Junliang Yu, Min Gao, Hongzhi Yin, Jundong Li, Chongming Gao, and Qinyong Wang. 2019. Generating reliable friends via adversarial training to improve social recommendation. In *ICDE*. IEEE, 768–777.

[65] Junliang Yu, Hongzhi Yin, Min Gao, Xin Xia, Xiangliang Zhang, and Nguyen Quoc Viet Hung. 2021. Socially-aware self-supervised tri-training for recommendation. In *SIGKDD*. 2084–2092.

[66] Junliang Yu, Hongzhi Yin, Jundong Li, Min Gao, Zi Huang, and Lizhen Cui. 2020. Enhancing social recommendation with adversarial graph convolutional networks. *IEEE TKDE* 34, 8 (2020), 3727–3739.

[67] Junliang Yu, Hongzhi Yin, Jundong Li, Qinyong Wang, Nguyen Quoc Viet Hung, and Xiangliang Zhang. 2021. Self-supervised multi-channel hypergraph convolutional network for social recommendation. In *WWW*. 413–424.

[68] Wayne Xin Zhao, Junhua Chen, Pengfei Wang, Qi Gu, and Ji-Rong Wen. 2020. Revisiting alternative experimental settings for evaluating top-n item recommendation algorithms. In *CIKM*. 2329–2332.

[69] Cheng Zheng, Bo Zong, Wei Cheng, Dongjin Song, Jingchao Ni, Wenchao Yu, Haifeng Chen, and Wei Wang. 2020. Robust graph representation learning via neural sparsification. In *ICML*. PMLR, 11458–11468.