Graph Collaborative Signals Denoising and Augmentation for Recommendation

Ziwei Fan, Ke Xu
University of Illinois Chicago
USA
{zfan20,kxu25}@uic.edu

Zhang Dong
Amazon Ads
USA
andydong@amazon.com

Hao Peng*
Beihang University
China
penghao@act.buaa.edu.cn

Jiawei Zhang
University of California, Davis
USA
jiawei@ifmlab.org

Philip S. Yu
University of Illinois Chicago
USA
psyu@uic.edu

https://github.com/zfan20/GraphDA

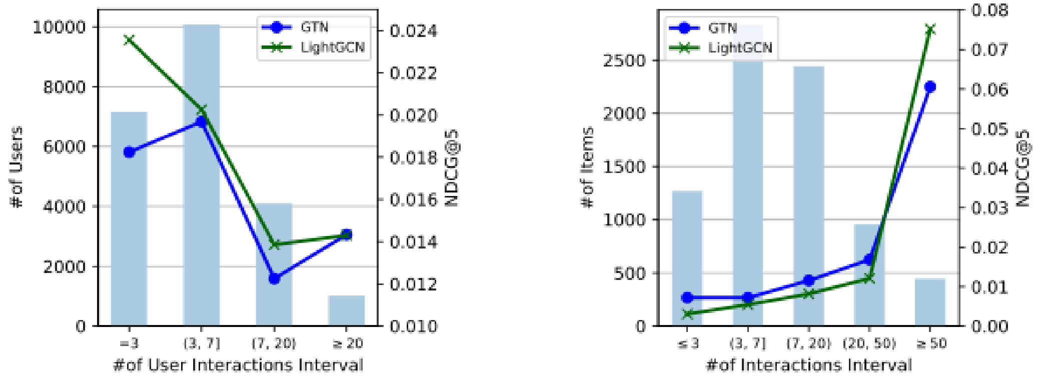
（SIGIR-2023)

Leibniz-Institut
für Sozialwissenschaften

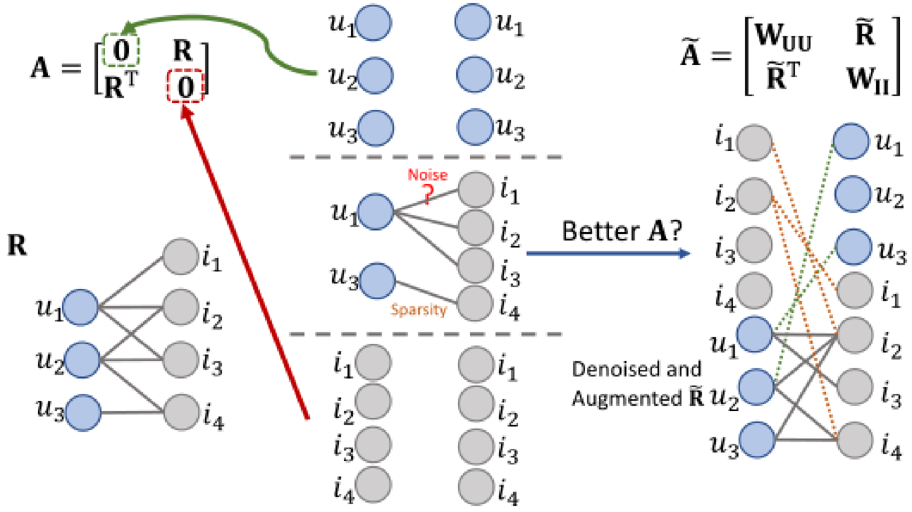
Reported by Ke Gan

1. Introduction
2. Approach
3. Experiments

Introduction



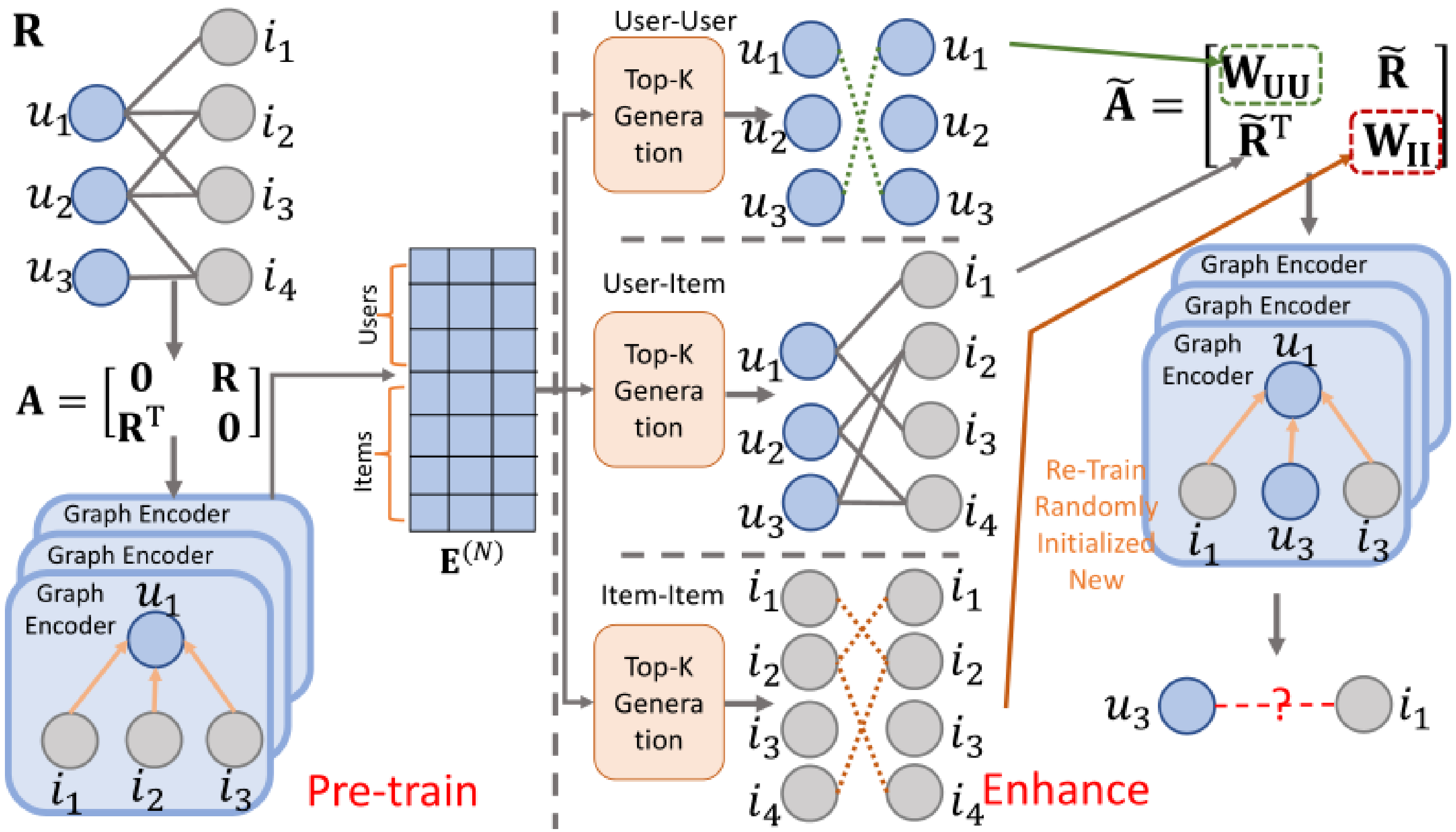
(a) Users (b) Items



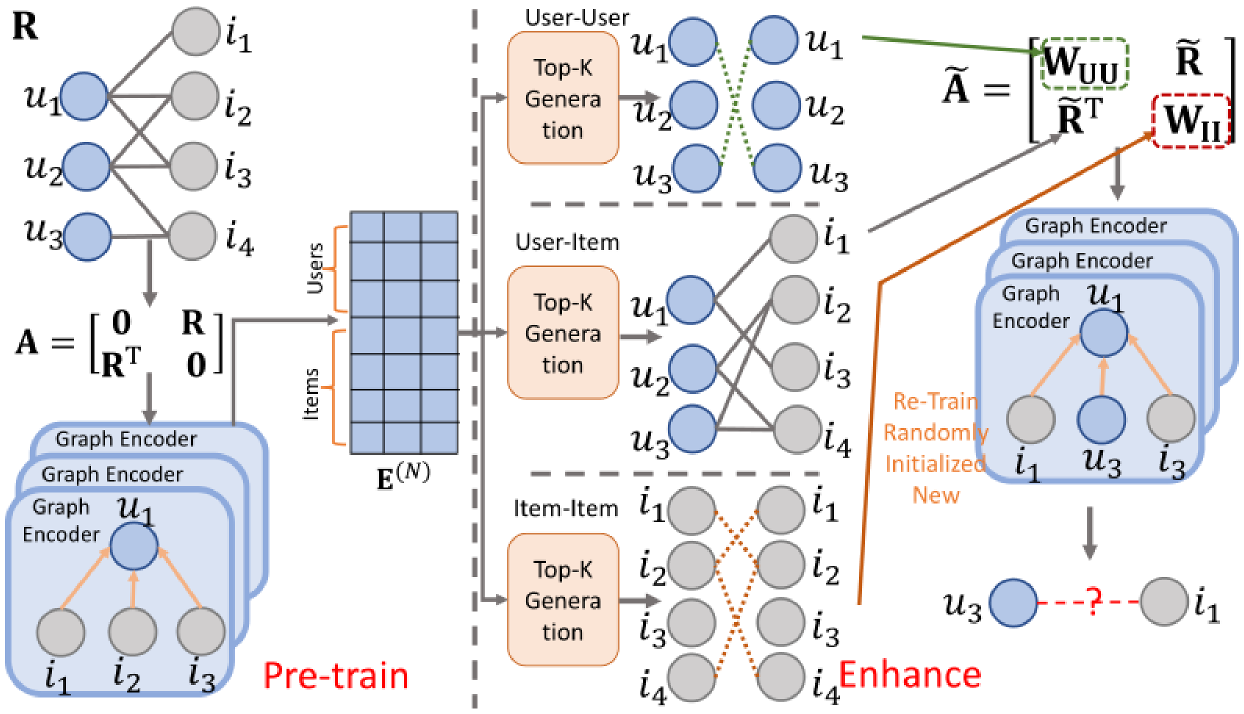
adjacency matrix can be noisy for users/items with abundant interactions and insufficient for users/items with scarce interactions

adjacency matrix ignores user-user and item-item correlations

Approach


R
u_1 u_2 u_3
i_1 i_2 i_3 i_4
$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{R} \\ \mathbf{R}^T & \mathbf{0} \end{bmatrix}$
Graph Encoder
Graph Encoder
Graph Encoder
Pre-train
Users
Items
$\mathbf{E}^{(N)}$
User-User
Top-K Generation
User-Item
Top-K Generation
Item-Item
Top-K Generation
$\tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{W}_{UU} & \tilde{\mathbf{R}} \\ \tilde{\mathbf{R}}^T & \mathbf{W}_{II} \end{bmatrix}$
Re-Train
Randomly
Initialized
New
Enhance
u_3 - - -?- - i_1

Approach



$$\mathbf{E} \in \mathbb{R}^{(|\mathcal{U}|+|\mathcal{I}|)\times d}$$

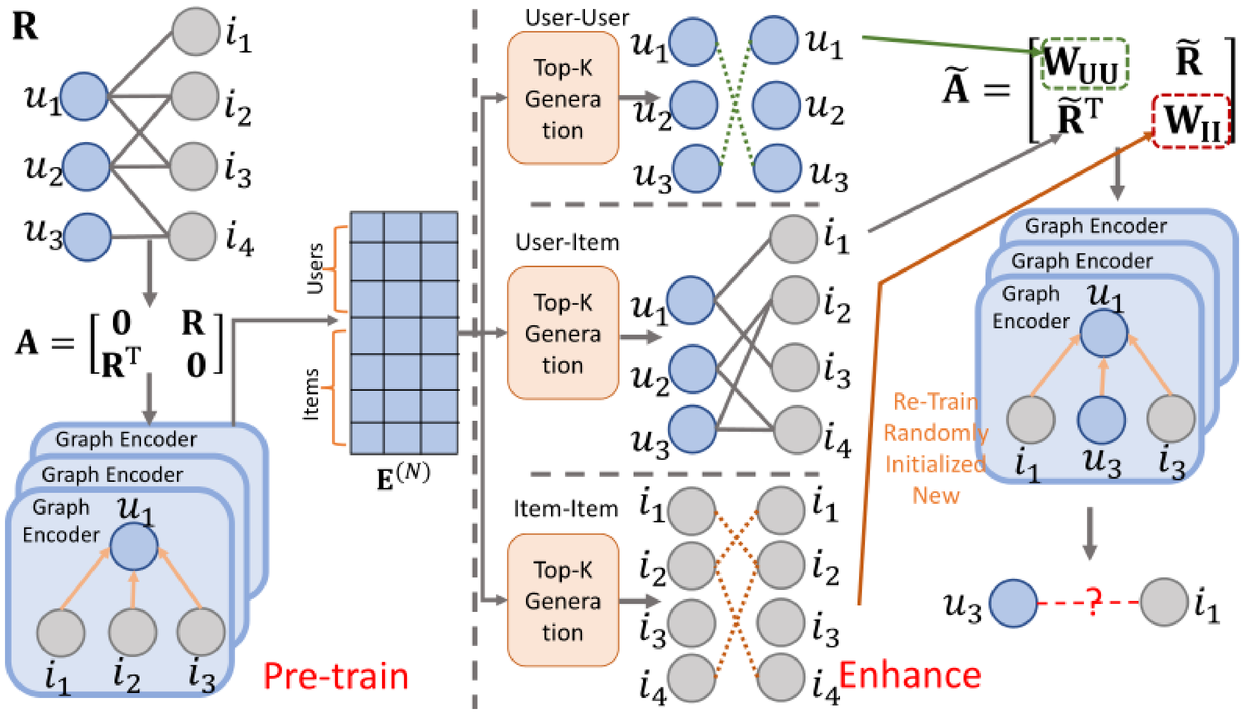
$$\mathbf{E}^{(N)} = \text{Encoder}(\mathbf{A}, \mathbf{E}) = (\mathbf{L})^{N-1}\mathbf{E}^{(0)}, \tag{1}$$

$$P(i|u, \mathbf{A}) = \sigma\left(\mathbf{e}_u^{\top}\mathbf{e}_i\right) \quad \text{where} \quad i \in \mathcal{I} \setminus \mathcal{I}_u^{+}, \tag{2}$$

$$\mathcal{L} = -\sum_{(u,i^+,i^- \in \mathbf{R})} \log \sigma(\mathbf{e}_u^{\top}\mathbf{e}_{i^+} - \mathbf{e}_u^{\top}\mathbf{e}_{i^-}), \tag{3}$$

matrix factorization [29].

Approach



$$\underset{\{i_1, i_2, \ldots, i_{U_k} \in \mathcal{I}\}}{\arg\max} \quad \mathbf{e}_u^\top \mathbf{E}_{\mathcal{I}}^{(N)}, \tag{4}$$

We adopt the union of generated user-item interactions

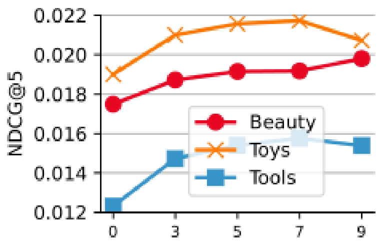
$$\underset{\{u_1, u_2, \ldots, u_{UU_k} \in \mathcal{U}\}}{\arg\max} \quad \mathbf{e}_u^\top \mathbf{E}_{\mathcal{U}}^{(N)}, \tag{5}$$

$$\widetilde{\mathbf{A}} = \begin{bmatrix} \mathbf{W}_{UU} & \widetilde{\mathbf{R}} \\ \widetilde{\mathbf{R}}^\top & \mathbf{W}_{II} \end{bmatrix}.$$

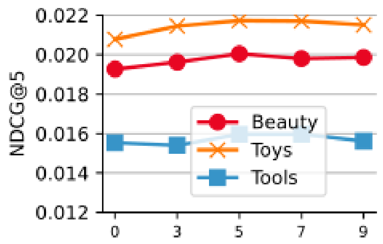
Experiment

Dataset	Beauty				Toys				Tools				Office			
Metric	H@10	N@10	H@20	N@20	H@10	N@10	H@20	N@20	H@10	N@10	H@20	N@20	H@10	N@10	H@20	N@20
NGCF	0.0447	0.0232	0.0724	0.0299	0.0461	0.0251	0.0672	0.0306	0.0329	0.0179	0.0480	0.0216	0.0261	0.0159	0.0453	0.0208
UltraGCN	0.0451	0.0234	0.0728	0.0304	0.0464	0.0250	0.0675	0.0308	0.0331	0.0179	0.0481	0.0217	0.0302	0.0171	0.0471	0.0210
GTN	0.0446	0.0230	0.0680	0.0289	0.0453	0.0248	0.0661	0.0301	0.0337	0.0184	0.0484	0.0221	0.0283	0.0161	0.0453	0.0204
LightGCN	0.0471	0.0244	0.0730	0.0309	0.0512	0.0273	0.0716	0.0325	0.0334	0.0182	0.0482	0.0219	0.0355	0.0197	0.0522	0.0238
Enhanced-UI	0.0486	0.0252	0.0755	0.0317	0.0530	0.0276	0.0765	0.0335	0.0364	0.0195	0.0527	0.0236	0.0363	0.0208	**0.0565**	0.0259
Improv.	+3.2%	+3.3%	+3.1%	+2.9%	+3.5%	+1.1%	+6.8%	+3.3%	+8.0%	+6.0%	+4.3%	+6.6%	+2.3%	+5.6%	+8.2%	+8.8%
GraphDA	**0.0514**	**0.0264**	**0.0804**	**0.0336**	**0.0549**	**0.0289**	**0.0795**	**0.0347**	**0.0373**	**0.0205**	**0.0532**	**0.0245**	**0.0383**	**0.0225**	0.0561	**0.0270**
Improv.	+9.1%	+8.2%	+8.7%	+7.9%	+7.2%	+5.9%	+11.1%	+6.9%	+10.7%	+11.4%	+5.4%	+10.8%	+7.9%	+14.2%	+7.5%	+13.4%

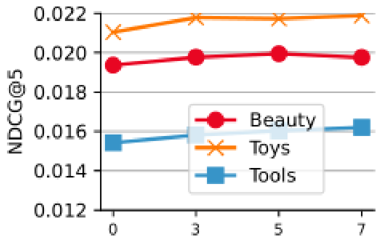
Experiment



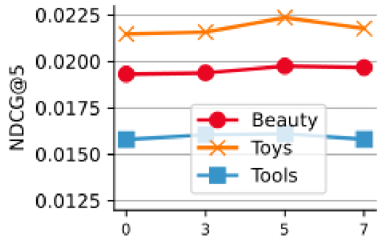
(a) Different values of U_k with best $I_k > 0$.



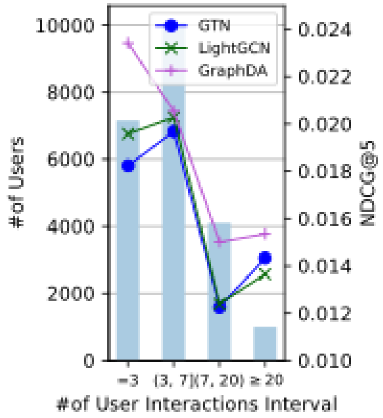
(b) Different values of I_k with best $U_k > 0$.



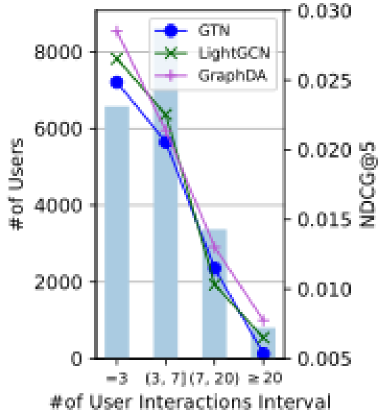
(c) Different values of UU_k



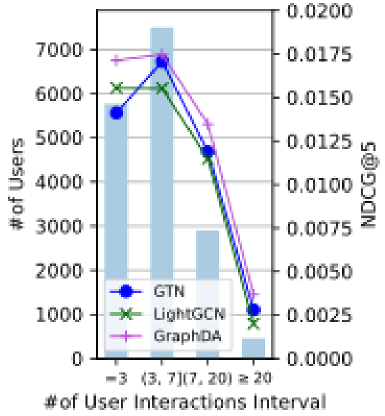
(d) Different values of II_k



(a) Beauty



(b) Toys



(c) Tools

Thank you!