# Knowledge-refined Denoising Network for Robust Recommendation

Xinjun Zhu
Zhejiang University
xjzhu@zju.edu.cn

Yuntao Du
Zhejiang University
ytdu@zju.edu.cn

Yuren Mao
Zhejiang University
yuren.mao@zju.edu.cn

Lu Chen
Zhejiang University
luchen@zju.edu.cn

Yujia Hu
Zhejiang University
charliehu@zju.edu.cn

Yunjun Gao
Zhejiang University
Alibaba-Zhejiang University Joint Institute of Frontier Technologies
gaoyj@zju.edu.cn

https://github.com/xj-zhu98/KRDN

**（SIGIR-2023)**

gesis
Leibniz-Institut für Sozialwissenschaften

**Reported by Ke Gan**

1. Introduction
2. Approach
3. Experiments

# Introduction

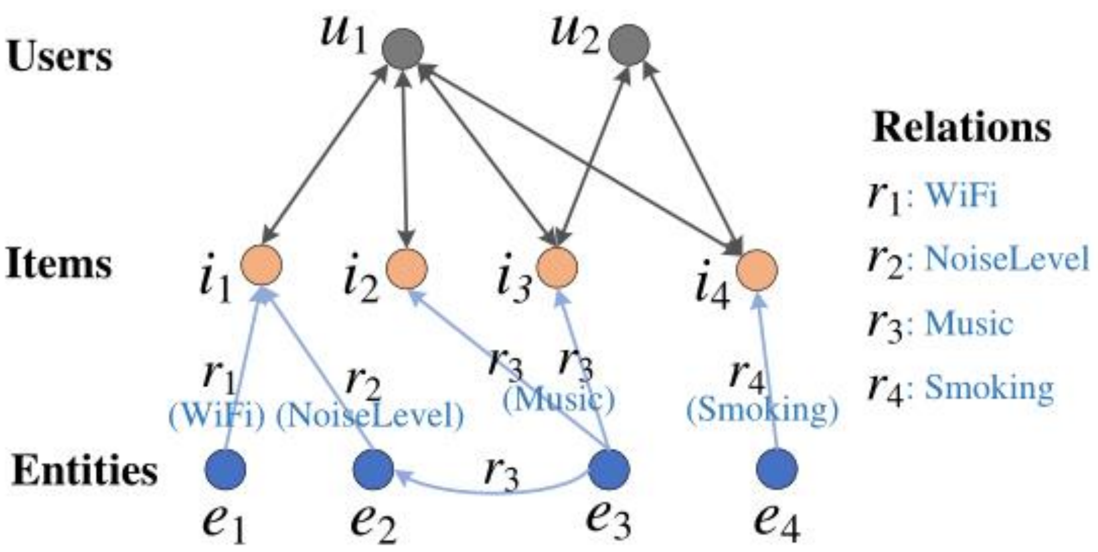



**Figure 1: An example of knowledge-aware recommendation on Yelp2018 dataset.**

## 1.Task-irrelevant Knowledge Propagation

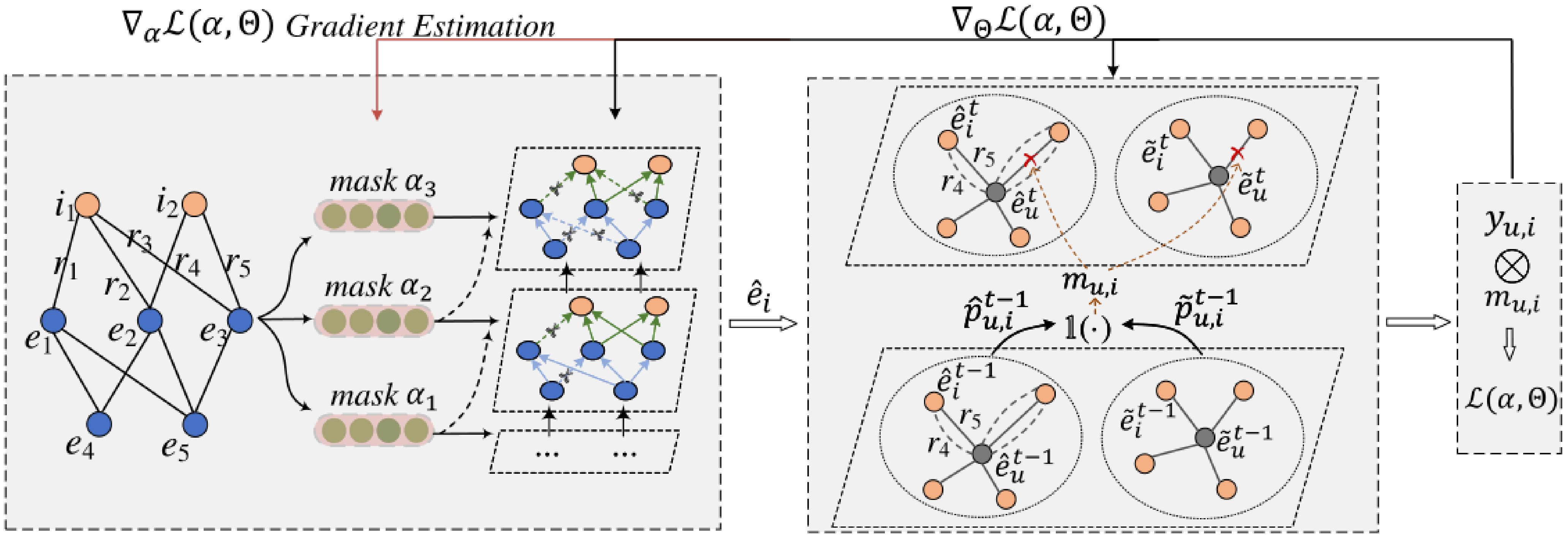
WiFi (r1) is a more marginal attribute linked to i1

## 2.Vulnerable to Interaction Noise

probably a noisy interaction

# Approach

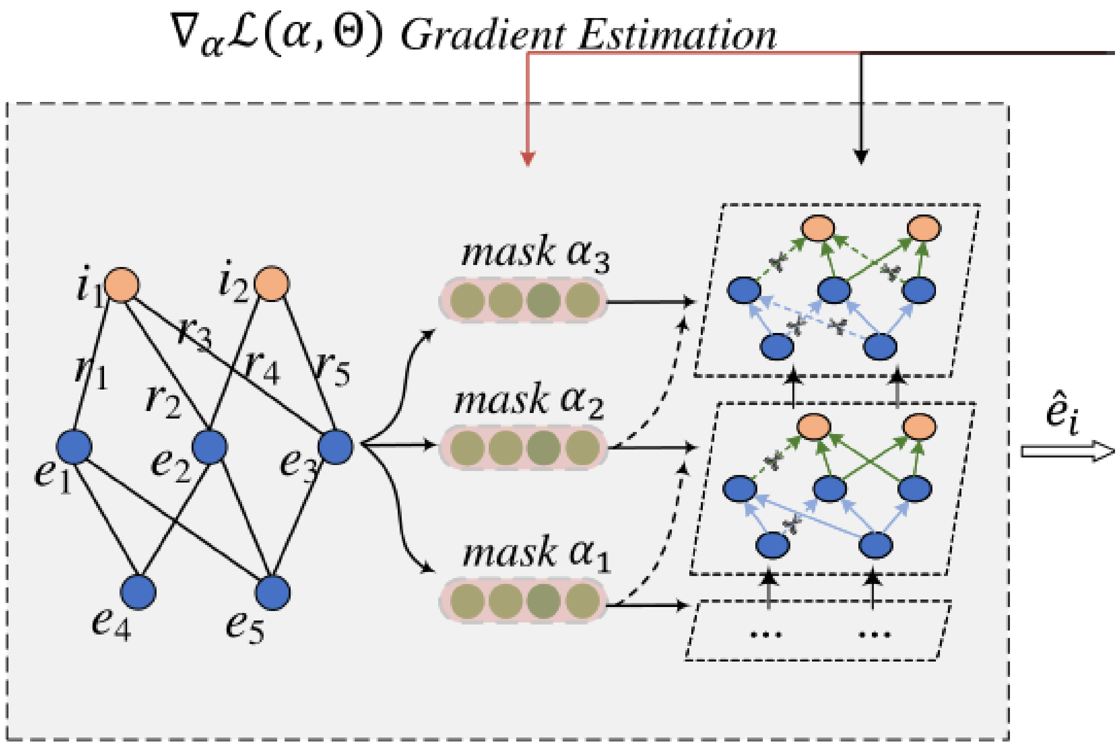





Adaptive Knowledge Refining

Contrastive Denoising Learning

# Approach




*Adaptive Knowledge Refining*

$$\tilde{\mathcal{T}} = \{(h_i, r_i, t_i) | m_i = 1\} \tag{1}$$

$$m_i \sim \mathrm{Bern}(\sigma(\alpha_i)) \tag{2}$$

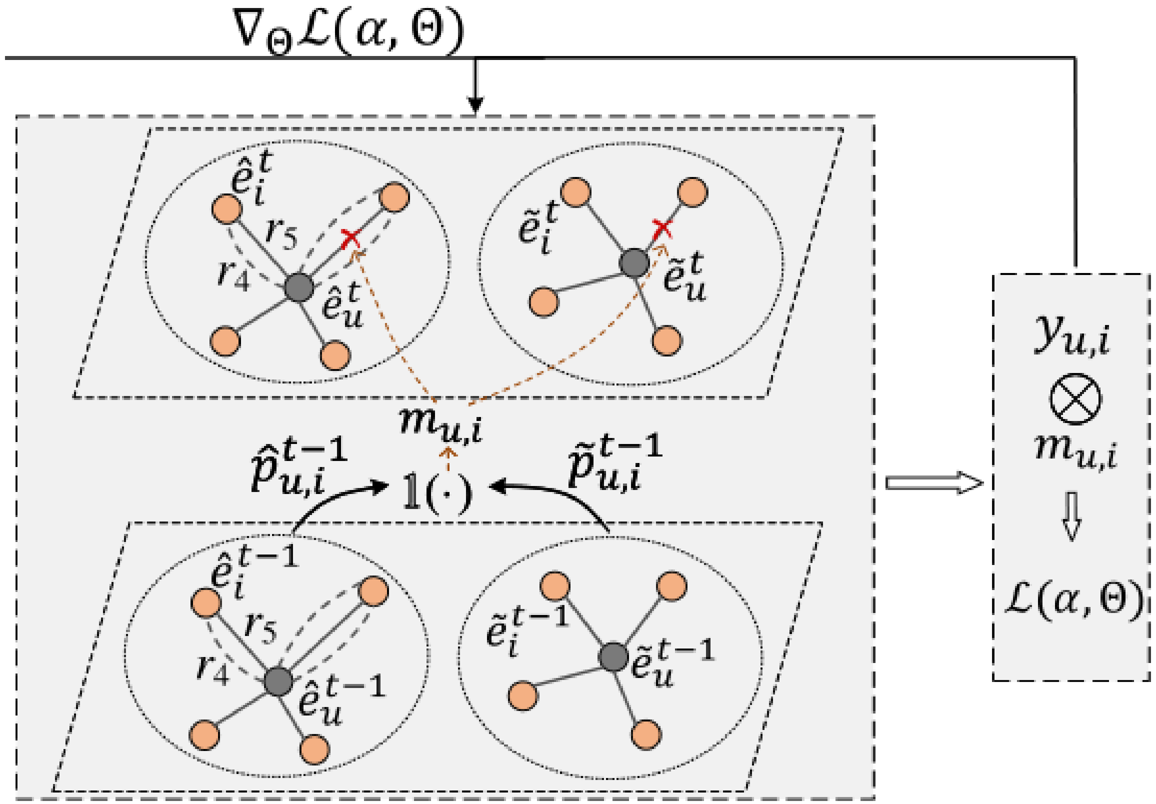
$$\tilde{\mathcal{L}}(\alpha, \Theta) = \mathbb{E}_{M \sim \prod_{i=1}^{K} \mathrm{Bern}(m_i; \alpha_i)} [\mathcal{L}(M, \Theta)] \tag{3}$$

$$\mathbf{e}_h^{(1)} = \frac{1}{|\mathcal{N}_h|} \sum_{(r,t) \in \mathcal{N}_h} \mathrm{ReLU}\left(\mathbf{W}\phi(\mathbf{e}_t^{(0)}, \mathbf{e}_r)\right) \cdot m_{h,t}^{(0)} \tag{4}$$

$$\mathbf{W}\phi\left(\mathbf{e}_t, \mathbf{e}_r\right) = \begin{cases} \mathbf{W}_1\left(\mathbf{e}_t \odot \mathbf{e}_r\right), (h, r, t) \in \mathcal{T}_{1,3} \\ \mathbf{W}_2\left(\mathbf{e}_t \oplus \mathbf{e}_r\right), (h, r, t) \in \mathcal{T}_2 \end{cases} \tag{5}$$

$$\mathbf{e}_h^{(l)} = \frac{1}{|\mathcal{N}_h|} \sum_{(r,t) \in \mathcal{N}_h} \mathrm{ReLU}\left(\mathbf{W}\phi(\mathbf{e}_t^{(l-1)}, \mathbf{e}_r)\right) \cdot m_{h,t}^{(l-1)} \tag{6}$$

# Approach




*Contrastive Denoising Learning*

$$m_{u,i} = \mathbb{1}\left(|\sigma(\tilde{p}_{u,i}) - \sigma(\hat{p}_{u,i})| < \gamma\right) \quad (7)$$

$$\tilde{p}_{u,i} = \frac{\exp\left(s\left(\tilde{\mathbf{e}}_u, \tilde{\mathbf{e}}_i\right)\right)}{\sum_{i' \in \mathcal{N}_{(u)}} \exp\left(s\left(\tilde{\mathbf{e}}_u, \tilde{\mathbf{e}}_{i'}\right)\right)} \quad (8)$$

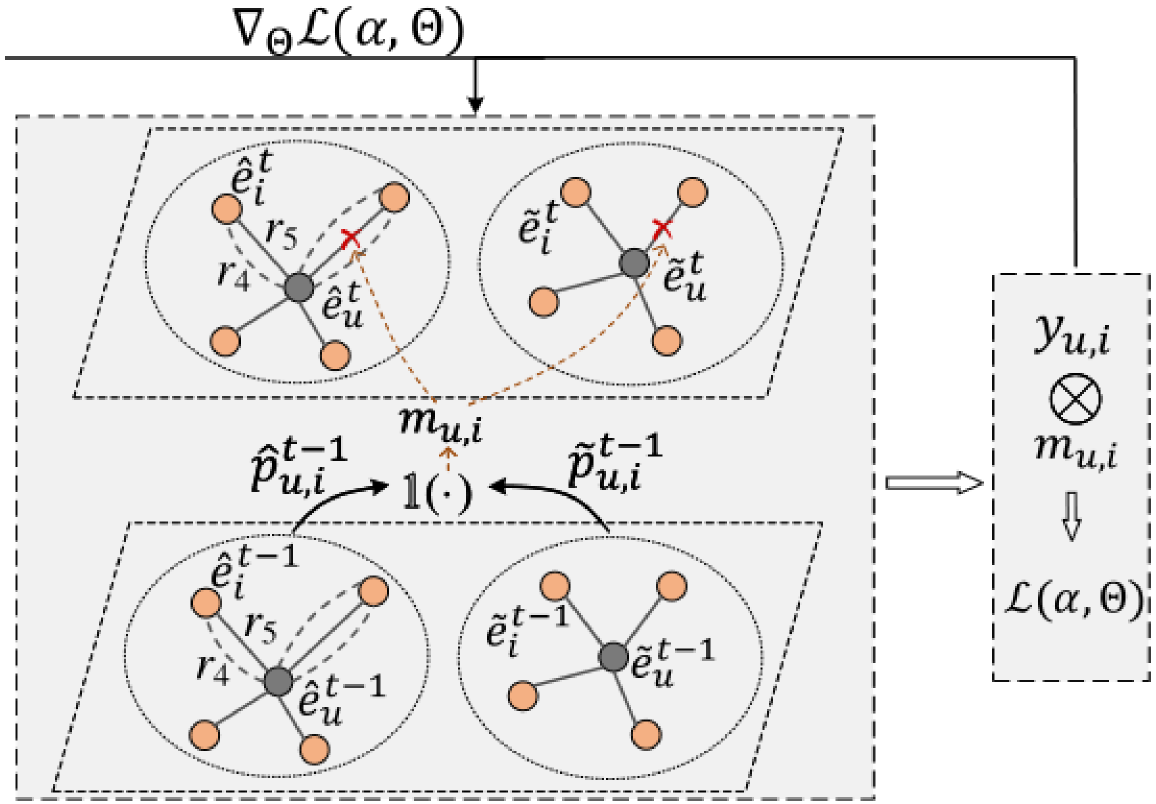
$$\hat{p}_{u,i} = \frac{\exp\left(\frac{1}{|\mathcal{R}_{(i)}|} \sum_{r \in \mathcal{R}_{(i)}} s\left(\mathbf{e}_r^\top \hat{\mathbf{e}}_u, \hat{\mathbf{e}}_i\right)\right)}{\sum_{i' \in \mathcal{N}_{(u)}} \exp\left(\frac{1}{|\mathcal{R}_{(i')}|} \sum_{r \in \mathcal{R}_{(i')}} s\left(\mathbf{e}_r^\top \hat{\mathbf{e}}_u, \hat{\mathbf{e}}_{i'}\right)\right)} \quad (9)$$

item has a relation set noted $\mathcal{R}_{(i)} = \{r | (h, r, t) \in \mathcal{T} \text{ and } h \in \mathcal{I}\}$,

$$\hat{\mathbf{e}}_u = \hat{\mathbf{e}}_u + \sum_{i \in \mathcal{N}_{(u)}} m_{u,i} \hat{p}_{u,i} \hat{\mathbf{e}}_i \quad (10)$$

$\tilde{\mathbf{e}}_u$

# Approach




*Contrastive Denoising Learning*

$$\hat{p}_{u,i} = \frac{\exp\left(\frac{1}{|\mathcal{R}_{(i)}|}\sum_{r\in\mathcal{R}_{(i)}} s\left(\mathbf{e}_r^\top \hat{\mathbf{e}}_u^{(n-1)}, \hat{\mathbf{e}}_i\right)\right)}{\sum_{i'\in\mathcal{N}_{(u)}}\exp\left(\frac{1}{|\mathcal{R}_{(i')}|}\sum_{r\in\mathcal{R}_{(i')}} s\left(\mathbf{e}_r^\top \hat{\mathbf{e}}_u^{(n-1)}, \hat{\mathbf{e}}_{i'}\right)\right)} \quad (11)$$

$$\hat{\mathbf{e}}_u^{(n)} = \frac{\hat{\mathbf{e}}_u^{(n-1)} + \sum_{i\in\mathcal{N}_{(u)}} m_{u,i}\hat{p}_{u,i}\hat{\mathbf{e}}_i}{\left\|\hat{\mathbf{e}}_u^{(n-1)} + \sum_{i\in\mathcal{N}_{(u)}} m_{u,i}\hat{p}_{u,i}\mathbf{e}_i\right\|_2} \quad (12)$$

$$\hat{\mathbf{e}}_v = \sum_{l=0}^{L}\hat{\mathbf{e}}_v^{(l)}, \qquad \tilde{\mathbf{e}}_v = \sum_{l=0}^{L}\tilde{\mathbf{e}}_v^{(l)} \quad (13)$$

$$\hat{y}_{u,i} = \cos(\tilde{\mathbf{e}}_u, \tilde{\mathbf{e}}_i) + \cos(\hat{\mathbf{e}}_u, \hat{\mathbf{e}}_i) \quad (14)$$

$$\mathcal{L} = \sum_{u,i\in\mathcal{D}}\left[m_{u,i}\left(1-\hat{y}_{u,i}\right)_+ + \frac{1}{|\mathcal{N}|}\sum_{j\in\mathcal{N}}\left(\hat{y}_{u,j}-\beta\right)_+\right] \quad (15)$$

$$\nabla_{\alpha}\tilde{\mathcal{L}}(\alpha,\Theta) = \frac{1}{2}\sum_i \left(f(\mathbf{b}) - f(\tilde{\mathbf{b}})\right)\left((-1)^{\tilde{b}_i}\mathbb{1}_{b_i\neq\tilde{b}_i}\sigma(|\alpha_i|)\right) \quad (16)$$

# Experiment

**Table 1: Statistics of the datasets.**

| | | Alibaba-iFashion | Last-FM | Yelp2018 |
|---|---|---|---|---|
| User-Item Interaction | #Users | 114,737 | 23,566 | 45,919 |
| | #Items | 30,040 | 48,123 | 45,538 |
| | #Interactions | 1,781,093 | 3,034,796 | 1,185,068 |
| Knowledge Graph | #Entities | 59,156 | 58,266 | 90,961 |
| | #Total-Triplets | 279,155 | 464,567 | 1,853,704 |

**Table 2: Overall performance comparison. "†" indicates the improvement of the KRDN over the baseline is significant at the level of 0.01. The highest scores are in Bold. R and N refer to Recall and NDCG, respectively.**

| Database | Method | MF | T-CE | SGCN | SGL | SimGCL | CKE | KGNN-LS | KGAT | CKAN | KGIN | MCCLK | KGCL | KRDN | %Imp. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Alibaba-iFashion | R@20 | 0.1095† | 0.1093† | 0.1145† | 0.1232† | 0.1243† | 0.1103† | 0.1039† | 0.1030† | 0.0970† | 0.1147† | 0.1089† | 0.1127† | **0.1372** | 10.38% |
| | N@20 | 0.0670† | 0.0631† | 0.0722† | 0.0771† | 0.0780† | 0.0676† | 0.0557† | 0.0627† | 0.0509† | 0.0716† | 0.0678† | 0.0713† | **0.0879** | 12.69% |
| Yelp2018 | R@20 | 0.0627† | 0.0705† | 0.0768† | 0.0788† | 0.0799† | 0.0653† | 0.0671† | 0.0705† | 0.0646† | 0.0698† | 0.0696† | 0.0748† | **0.0842** | 5.38% |
| | N@20 | 0.0413† | 0.0542† | 0.0547† | 0.0518† | 0.0520† | 0.0423† | 0.0422† | 0.0463† | 0.0441† | 0.0451† | 0.0449† | 0.0491† | **0.0544** | 4.62% |
| Last-FM | R@20 | 0.0724† | 0.0814† | 0.0863† | 0.0879† | 0.0824† | 0.0732† | 0.0880† | 0.0873† | 0.0812† | 0.0978† | 0.0671† | 0.0686† | **0.1023** | 4.60% |
| | N@20 | 0.0617† | 0.0683† | 0.0759† | 0.0775† | 0.0736† | 0.0630† | 0.0642† | 0.0744† | 0.0660† | 0.0848† | 0.0603† | 0.0629† | **0.0946** | 11.56% |
| Polluted Alibaba-iFashion | R@20 | 0.0982† | 0.0990† | 0.1035† | 0.1146† | 0.1161† | 0.0911† | 0.0921† | 0.0902† | 0.0874† | 0.1037† | 0.0981† | 0.1065† | **0.1312** | 13.01% |
| | N@20 | 0.0607† | 0.0584† | 0.0639† | 0.0714† | 0.0722† | 0.0630† | 0.0471† | 0.0542† | 0.0448† | 0.0643† | 0.0613† | 0.0672† | **0.0839** | 16.20% |
| Polluted Yelp2018 | R@20 | 0.0589† | 0.0669† | 0.0697† | 0.0755† | 0.0759† | 0.0634† | 0.0612† | 0.0642† | 0.0609† | 0.0679† | 0.0667† | 0.0718† | **0.0816** | 7.51% |
| | N@20 | 0.0392† | 0.0477† | 0.0480† | 0.0492† | 0.0495† | 0.0412† | 0.0401† | 0.0407† | 0.0416† | 0.0436† | 0.0422† | 0.0472† | **0.0528** | 6.67% |
| Polluted Last-FM | R@20 | 0.0711† | 0.0807† | 0.0858† | 0.0879† | 0.0948† | 0.0849† | 0.0863† | 0.0845† | 0.0805† | 0.0960† | 0.0668† | 0.0731† | **0.1053** | 9.69% |
| | N@20 | 0.0610† | 0.0675† | 0.0741† | 0.0791† | 0.0844† | 0.0735† | 0.0630† | 0.0743† | 0.0658† | 0.0849† | 0.0592† | 0.0695† | **0.0988** | 16.37% |

# Experiment

**Table 3: Impact of knowledge refining & denoising.**

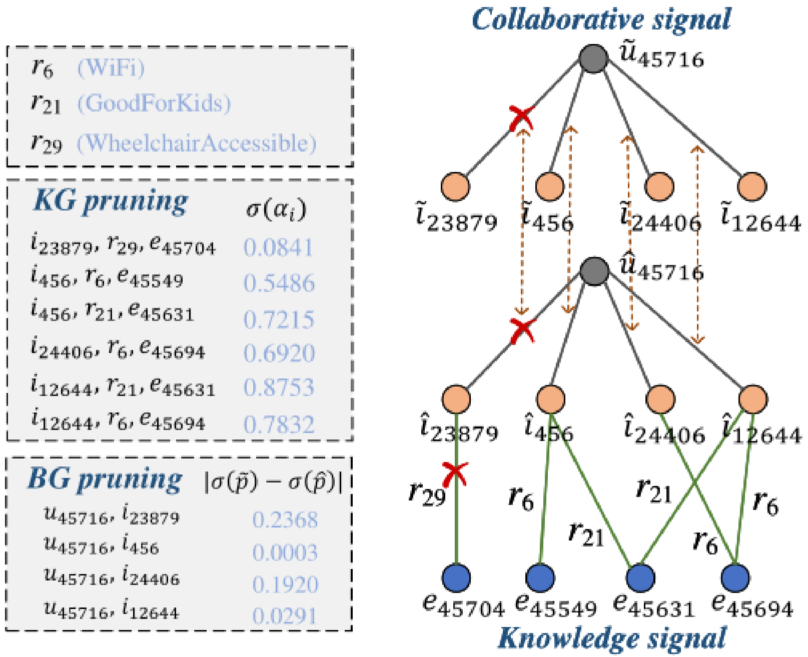
| | Alibaba-iFashion | | Yelp2018 | | Last-FM | |
|---|---|---|---|---|---|---|
| | recall | ndcg | recall | ndcg | recall | ndcg |
| w/o AKR | 0.1317 | 0.0833 | 0.0826 | 0.0538 | 0.1008 | 0.0933 |
| w/o CDL | 0.1240 | 0.0794 | 0.0801 | 0.0521 | 0.0984 | 0.0905 |
| w/o AKR&CDL | 0.1225 | 0.0773 | 0.0789 | 0.0512 | 0.0974 | 0.0903 |

**Table 4: Impact of the number of layers $L$.**

| | Alibaba-iFashion | | Yelp2018 | | Last-FM | |
|---|---|---|---|---|---|---|
| | recall | ndcg | recall | ndcg | recall | ndcg |
| KRDN-1 | 0.1356 | 0.0866 | 0.0840 | 0.0544 | 0.1017 | 0.0939 |
| KRDN-2 | 0.1365 | 0.0873 | 0.0842 | 0.0545 | 0.1023 | 0.0946 |
| KRDN-3 | 0.1372 | 0.0879 | 0.0841 | 0.0545 | 0.1021 | 0.0941 |

**Table 5: Impact of the iteration times $n$.**

| | Alibaba-iFashion | | Yelp2018 | | Last-FM | |
|---|---|---|---|---|---|---|
| | recall | ndcg | recall | ndcg | recall | ndcg |
| n-1 | 0.1344 | 0.0852 | 0.0826 | 0.0527 | 0.1009 | 0.0929 |
| n-2 | 0.1369 | 0.0875 | 0.0842 | 0.0545 | 0.1023 | 0.0946 |
| n-3 | 0.1372 | 0.0879 | 0.0841 | 0.0543 | 0.1023 | 0.0945 |



# Experiment

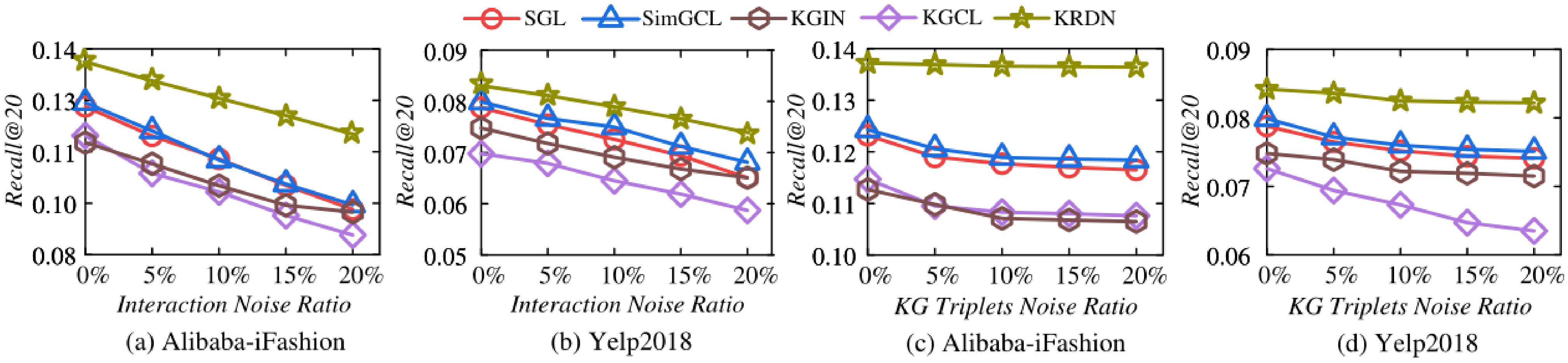



**Figure 3: Impact of different ratio of noise in user-item graph and knowledge graph.**

# Thank you!