Global Heterogeneous Graph and Target Interest Denoising for Multi-behavior Sequential Recommendation

Xuewei Li
College of Intelligence and
Computing, Tianjin University
Tianjin, China
lixuewei@tju.edu.cn

Hongwei Chen
College of Intelligence and
Computing, Tianjin University
Tianjin, China
chw@tju.edu.cn

Jian Yu
College of Intelligence and
Computing, Tianjin University
Tianjin, China
yujian@tju.edu.cn

Mankun Zhao
College of Intelligence and
Computing, Tianjin University
Tianjin, China
zmk@tju.edu.cn

Tianyi Xu
College of Intelligence and
Computing, Tianjin University
Tianjin, China
tianyi.xu@tju.edu.cn

Wenbin Zhang
Information and Network Center,
Tianjin University
Tianjin, China
zhangwenbin@tju.edu.cn

Mei Yu*
College of Intelligence and
Computing, Tianjin University
Tianjin, China
yumei@tju.edu.cn

code:none

WSDM 2024

Reported by Minqin Li

Introduction

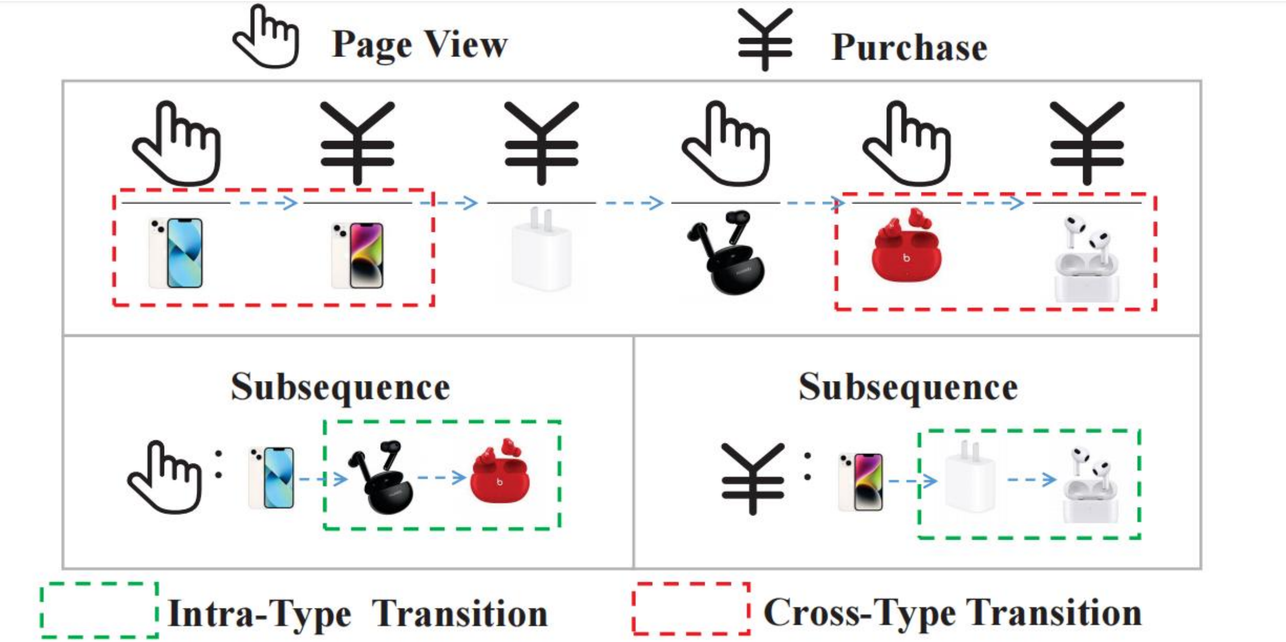


Figure 1: Illustration of heterogeneous item transitions, including intra-type transition and cross-type transition.

1.Most existing methods either do not explicitly model heterogeneous item transitions or ignore information from global users, and thus cannot learn heterogeneous item transitions.

2.Extracting specific interests for the target behavior from the auxiliary behaviors and filtering out the noise information in the auxiliary behaviors is another critical challenge.

Method

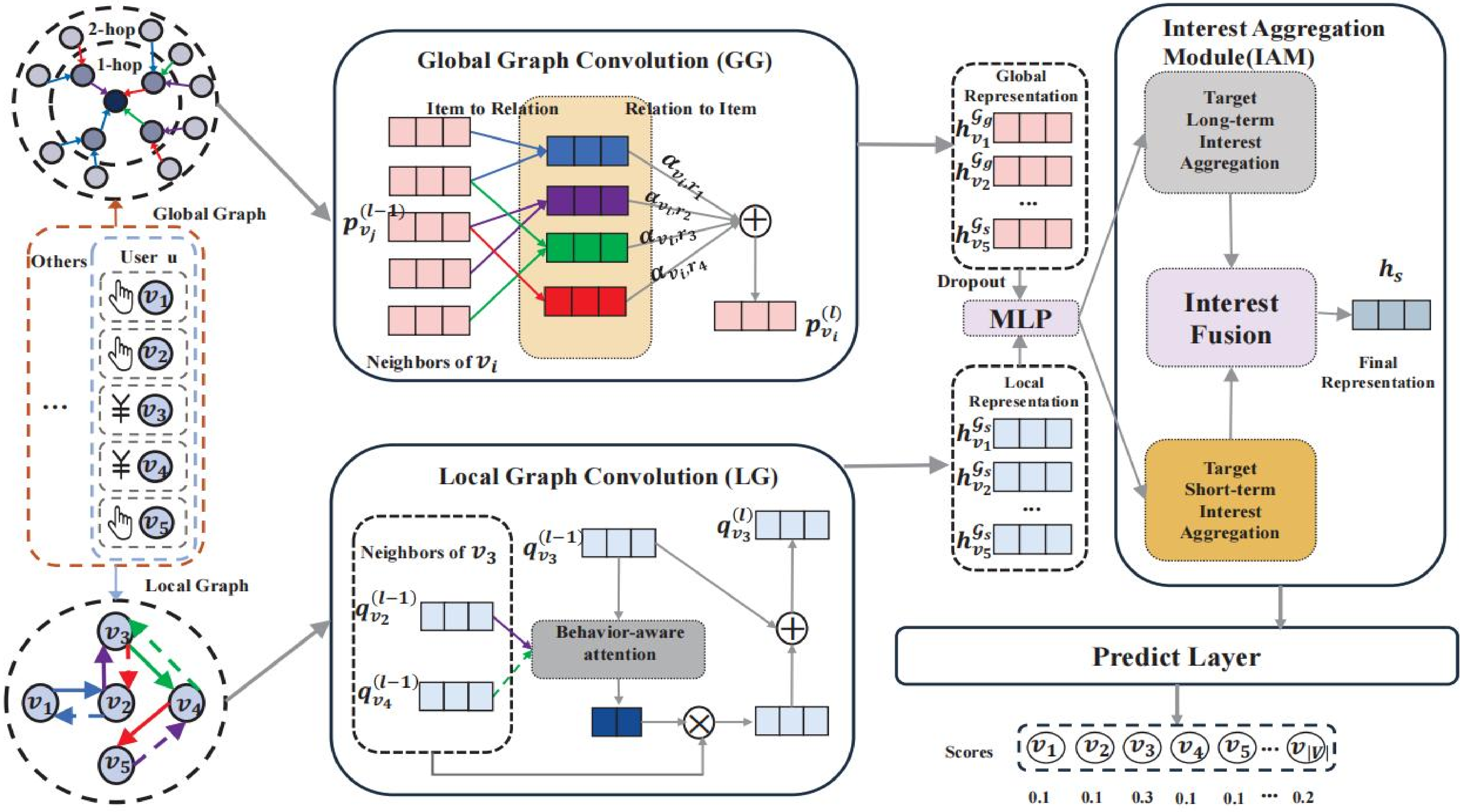
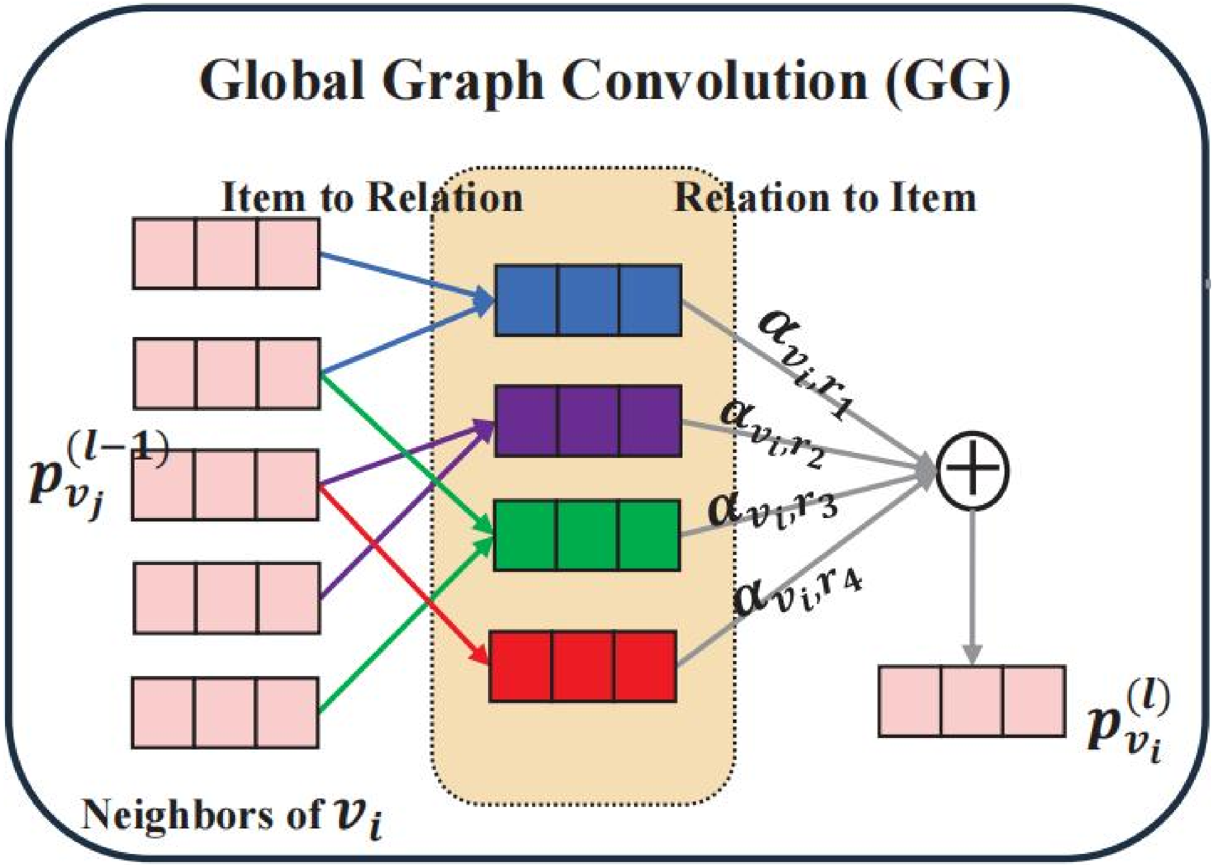


Figure 2: The overall framework of our proposed GHTID.

Method



$$\mathrm{PMI}(v_i, v_j, x \rightarrow y) = \log \frac{\mathrm{p}(v_i, v_j, x \rightarrow y)}{\mathrm{p}(v_i, x)\mathrm{p}(v_j, y)}$$

$$\mathrm{p}(v_i, v_j, x \rightarrow y) = \frac{|\mathcal{S}(v_i, v_j, x \rightarrow y)|}{|\mathcal{S}|} \quad (1)$$

$$\mathrm{p}(v_i, x) = \frac{|\mathcal{S}(v_i, x)|}{|\mathcal{S}|}$$

where $\mathcal{S}$ is the set of all item sequences, $\mathcal{S}(v_i, x)$ is the set of item sequences containing item v_i with behavior type x, and $\mathcal{S}(v_i, v_j, x \rightarrow y)$ is the set of item sequences containing both item v_i with behavior type x and item v_j with behavior type y. The co-occurrence coefficient of v_i and v_j under the relationship $x \rightarrow y$ is $w_{ij}^{x \rightarrow y} = \mathrm{PMI}(v_i, v_j, x \rightarrow y)$.

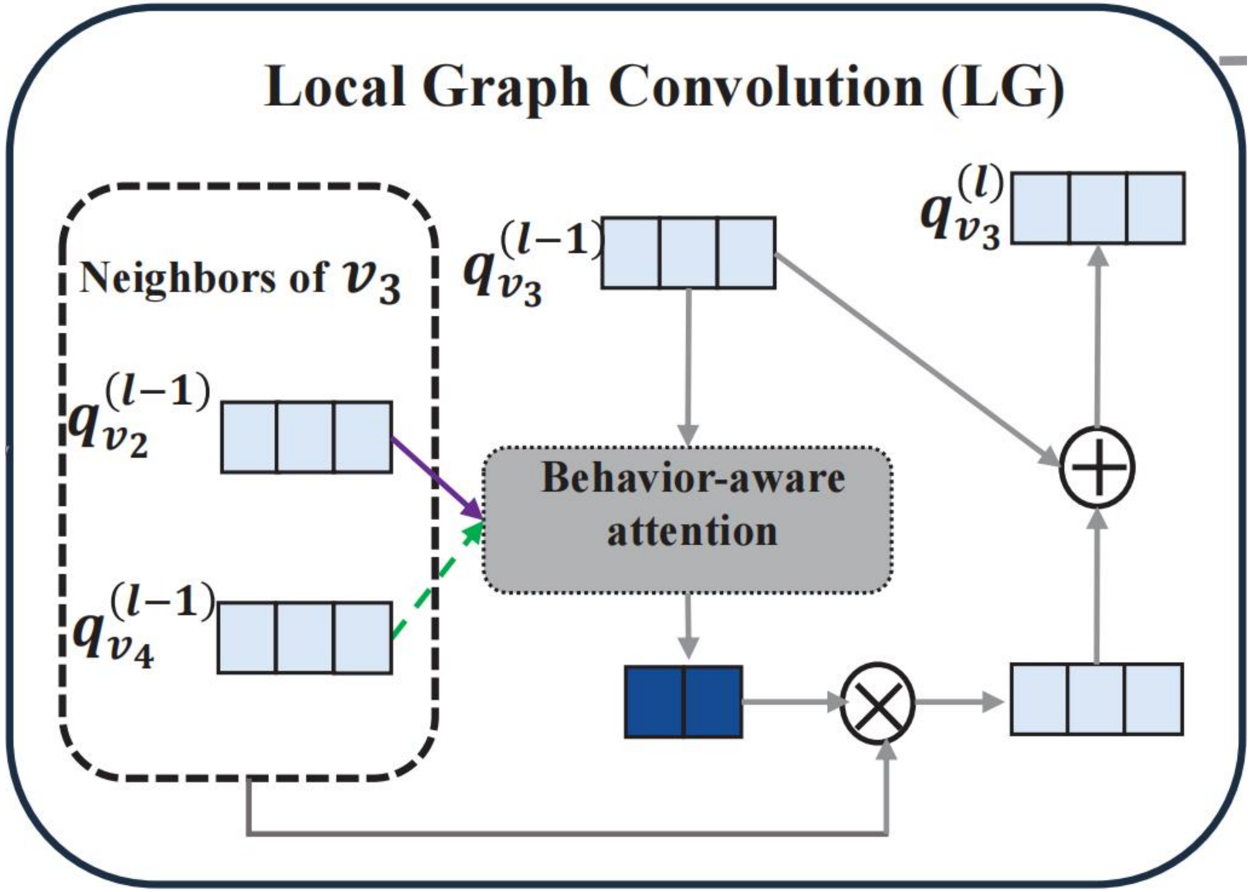
$$m_{v_i,r}^{(l)} = \frac{1}{|\sum_{v_j \in \mathcal{N}_r^{\mathcal{G}_g}(v_i)} w_{ij}^r|} \sum_{v_j \in \mathcal{N}_r^{\mathcal{G}_g}(v_i)} w_{ij}^r p_{v_j}^{(l-1)} \quad (2)$$

$$\pi(v_i, r) = a^T LeakyReLU(\mathbf{W}_r \ \$ \ [p_{v_i}^{(l-1)} || m_{v_i,r}^{(l)}])$$

$$\alpha_{v_i,r} = \frac{exp(\pi(v_i, r))}{\sum_{k \in \mathcal{R}_g} exp(\pi(v_i, k))} \quad (3)$$

$$p_{v_i}^{(l)} = \sum_{k \in \mathcal{R}_g} \alpha_{v_i,k} m_{v_i,k}^{(l)}$$

Method

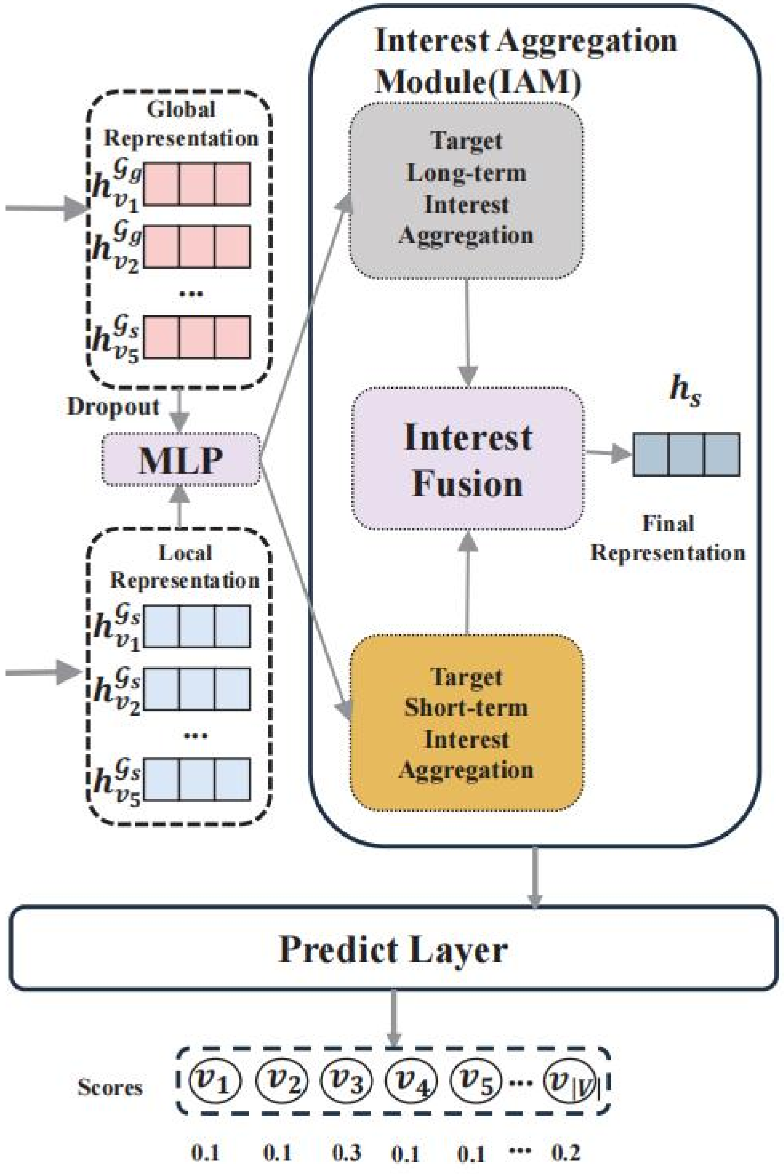


$$q_{v_i}^{(l)} = q_{v_i}^{(l-1)} + \sum_{v_j \in \mathcal{N}_{v_i}^{\mathcal{G}s}} \mathtt{attn}(v_i, v_j) q_{v_j}^{(l-1)} \tag{4}$$

$$\pi(v_i, v_j) = a_{r_{ij}}^T \text{LeakeyReLU}([\mathbf{W}_1 q_{v_i}^{l-1} || \mathbf{W}_2 q_{v_j}^{l-1}])$$

$$\mathtt{attn}(v_i, v_j) = \frac{exp(\pi(v_i, v_j))}{\sum_{v_k \in \mathcal{N}_{v_i}^{\mathcal{G}s}} exp(\pi(v_i, v_k))} \tag{5}$$

Method



$$h_{v_i} = \mathrm{Relu}(\mathbf{W}_3[\mathrm{Dropout}(h_{v_i}^{\mathcal{G}_g})||h_{v_i}^{\mathcal{G}_s}]) \quad (6)$$

$$q_{short} = h_{v_n}$$
$$\alpha_i = a_{short}^T \sigma(\mathbf{W}_4 q_{short} + \mathbf{W}_5 h_{v_i} + b_1) \quad (7)$$
$$h_s^{short} = \sum_{i=1}^{n} m_i \alpha_i h_{v_i}$$

$$q_{long} = \frac{1}{n}\sum_{i=1}^{n} m_i h_{v_i}$$
$$\beta_i = a_{long}^T \sigma(\mathbf{W}_6 q_{long} + \mathbf{W}_7 h_{v_i} + b_2) \quad (8)$$
$$h_s^{long} = \sum_{i=1}^{n} \beta_i h_{v_i}$$

$$h_s = \mathrm{LeakeyReLU}(\mathbf{W}_8[h_s^{short}||h_s^{long}]) \quad (9)$$

$$Loss = -\frac{1}{|O|}\sum_{(s,i,b)\in O} log \frac{exp(y_{s,i}^b)}{\sum_{j=1}^{N} exp(y_{s,j}^b)} \quad (10)$$

Experiments

Table 1: Statistics of the datasets after preprocessing.

Dataset	#Users	#items	Behavior Type	#Interaction
ML1M	5645	2357	Exam Like	628,892 223,305
UB	20858	30793	Page View Cart Favorite Purchase	470,731 85,910 28,242 136,250
Rec15	36917	9621	Click Purchase	446,442 233,263
Tmall	17209	16177	Page View Favorite Purchase	831,117 240,901 121,168

Experiments

Methods	ML1M		UB		Rec15		Tmall	
	HR@10	NDCG@10	HR@10	NDCG@10	HR@10	NDCG@10	HR@10	NDCG@10
FPMC	0.1086	0.0510	0.0467	0.0249	0.3829	0.2102	0.0352	0.0191
TransRec	0.0852	0.0409	0.0589	0.0342	0.3697	0.1928	0.0374	0.0325
SASRec	0.1350	0.0651	0.0744	0.0412	0.3615	0.1889	0.0862	0.0521
TiSASRec	0.1327	0.0638	0.0736	0.0415	0.4093	0.2056	0.0746	0.0445
GCEGNN	0.1319	0.0618	0.0611	0.0328	0.4026	0.2053	0.0823	0.0487
TransRec++	0.1088	0.0508	0.0661	0.0413	0.4064	0.2209	0.0593	0.0377
DMT	0.1158	0.0521	0.0613	0.0343	0.3942	0.2036	0.0642	0.0335
MGNN-Spred	0.1134	0.0548	0.0727	0.0386	0.4164	0.2108	0.0449	0.0236
M-SR	0.1349	0.0647	0.0784	0.0407	0.4315	0.2327	0.0811	0.0498
MBSTR	0.1431	0.0716	0.0904	0.0453	0.4239	0.2274	0.0905	0.0516
GPG4HSR	0.1460	0.0737	0.0830	0.0462	0.4198	0.2160	0.0944	0.0548
GHTID	**0.1663**	**0.0869**	**0.1124**	**0.0613**	**0.5081**	**0.2685**	**0.1115**	**0.0664**
Improv.	13.90%	17.90%	23.34%	32.68%	19.86%	18.07%	18.11%	21.17%

Table 2: Overall model performance on four datasets, with the metrics of HR@N and NDCG@N (N=10).

Experiments

Table 3: Ablation study of GHTID

Methods	UB		Tmall	
	HR@10	NDCG@10	HR@10	NDCG@10
GHTID w/o GG	0.1042	0.0541	0.1003	0.0528
GHTID w/o LG	0.0917	0.0498	0.0947	0.0512
GHTID w/o IAM	0.1057	0.0569	0.1038	0.0557
GHTID	**0.1124**	**0.0613**	**0.1115**	**0.0664**
SASRec	0.0744	0.0412	0.0862	0.0521
SASRec+IAM	**0.0847**	**0.0512**	**0.0922**	**0.0535**
MBSTR	0.0904	0.0453	0.0905	0.0516
MBSTR+IAM	**0.0961**	**0.0523**	**0.0967**	**0.0551**

Experiments

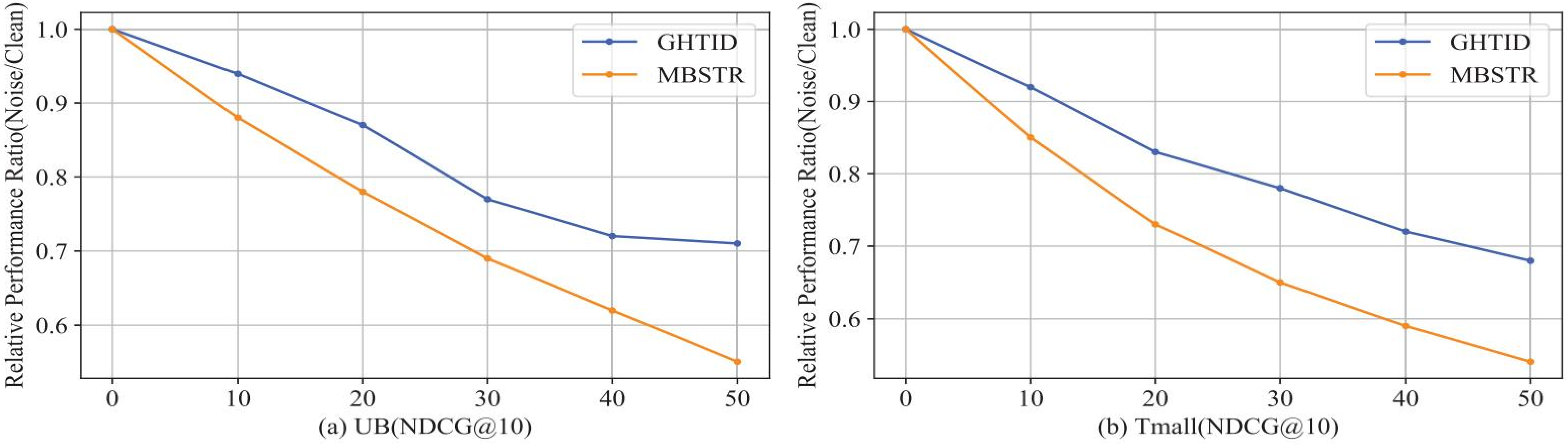


Figure 3: Relative performance drop on dataset UB and Tmall when the test data are corrupted by synthetic noises on auxiliary behavior. The x-axis is the percentage of corrupted data from auxiliary behavior. The y-axis is the ratio of the performance with noisy test data to the performance with clean training data.

Experiments

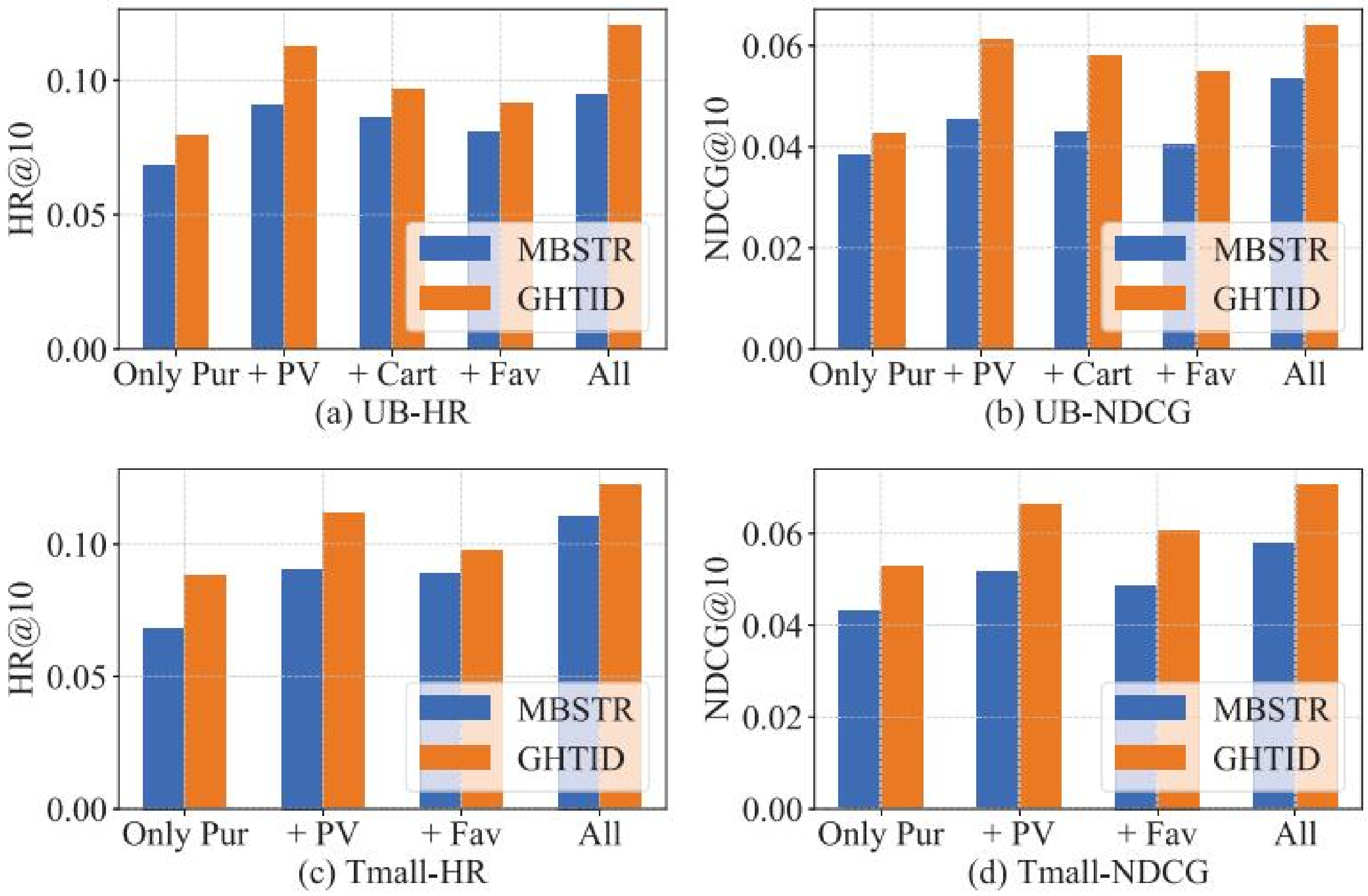


Figure 4: Effect of auxiliary behaviors.

Thanks