Knowledge Enhancement for Contrastive Multi-Behavior Recommendation

Hongrui Xuan
Nanjing University of Aeronautics and Astronautics
Nanjing, China
1692595335@nuaa.edu.cn

Yi Liu
Nanjing University of Aeronautics and Astronautics
Nanjing, China
liuyi-sx21@nuaa.edu.cn

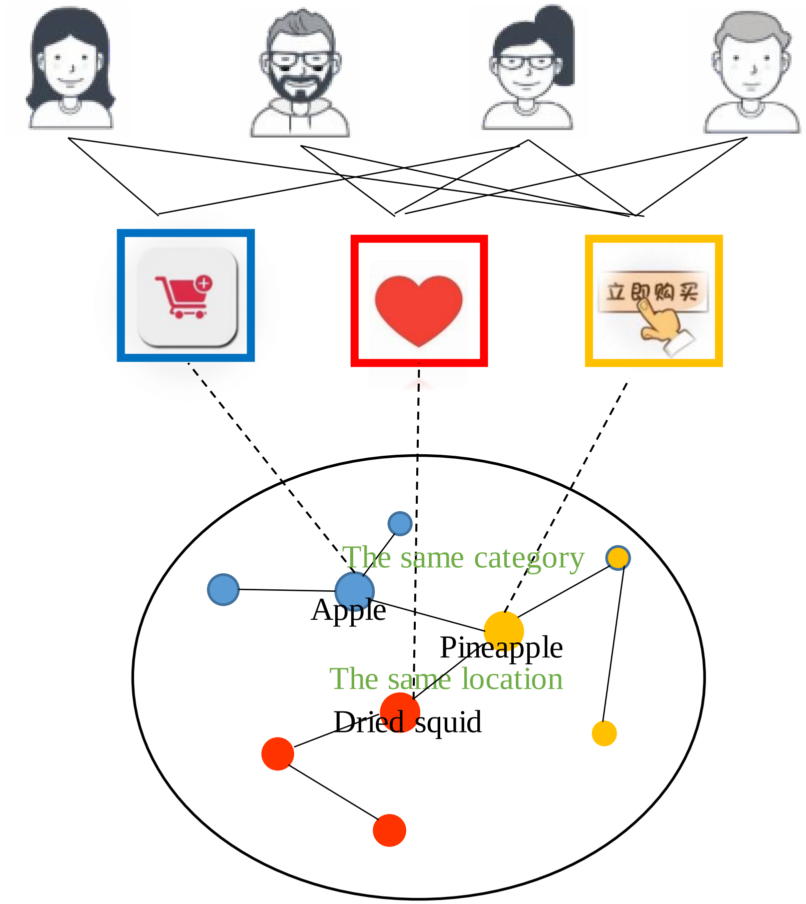
Bohan Li*
Nanjing University of Aeronautics and Astronautics
Nanjing, China
bhli@nuaa.edu.cn

Hongzhi Yin
The University of Queensland
Brisbane, Australia
h.yin1@uq.edu.au

WSDM 2023

Reported by Nengqiang Xiang

Introduction



There are two limitations in the existing methods:

(1) Target behavior data sparsity: people's purchase data on e-commerce platforms is far lower than click data

(2) Personalized behavioral diversity: how to model the coarse-grained and fine-grained commonality between different behaviors

Inspired by work related to CL and KG, we propose a Knowledge Enhancement Multi-Behavior Contrastive Learning Recommendation (KMCLR) framework to tackle above limitations.

Method

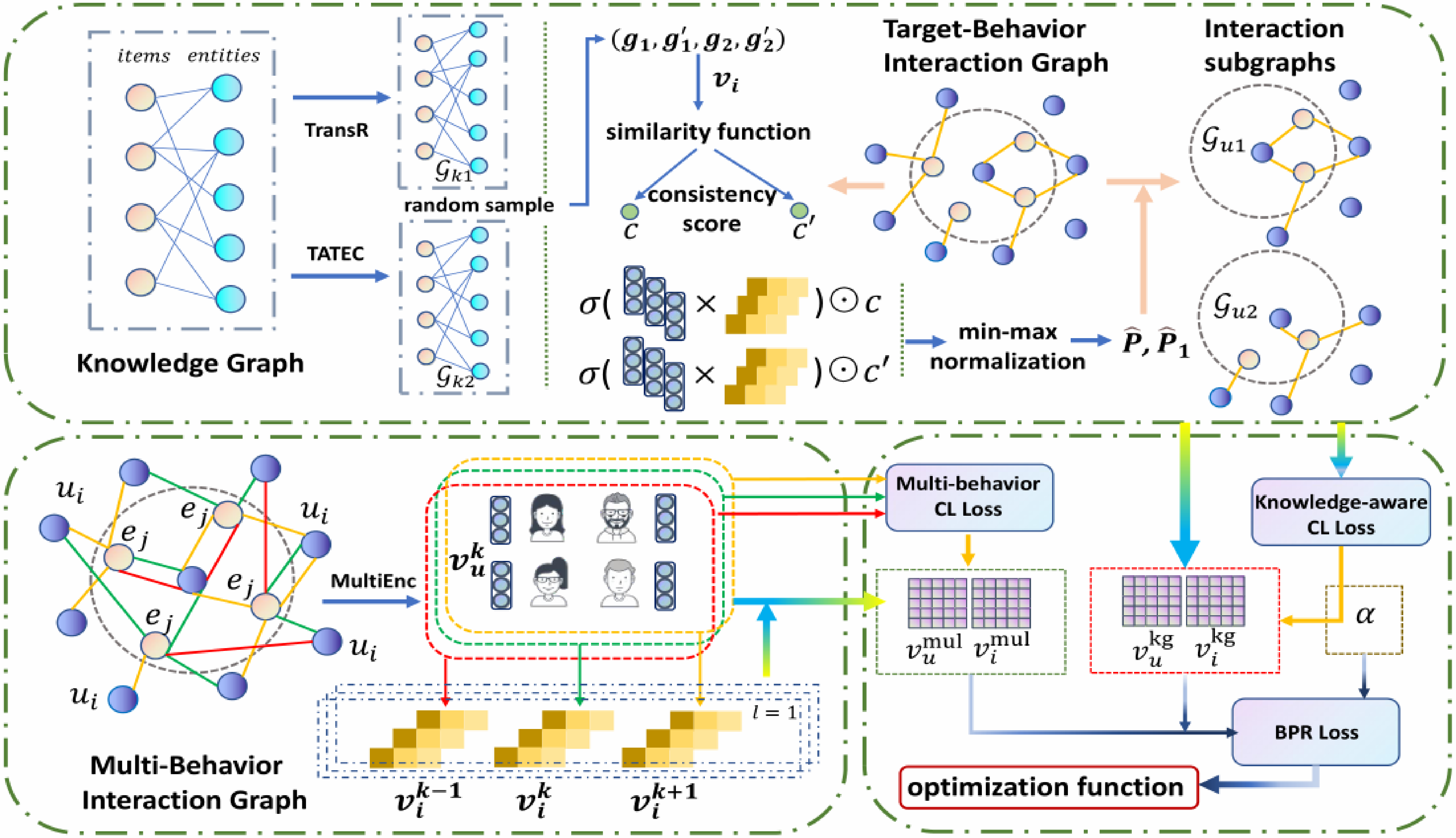
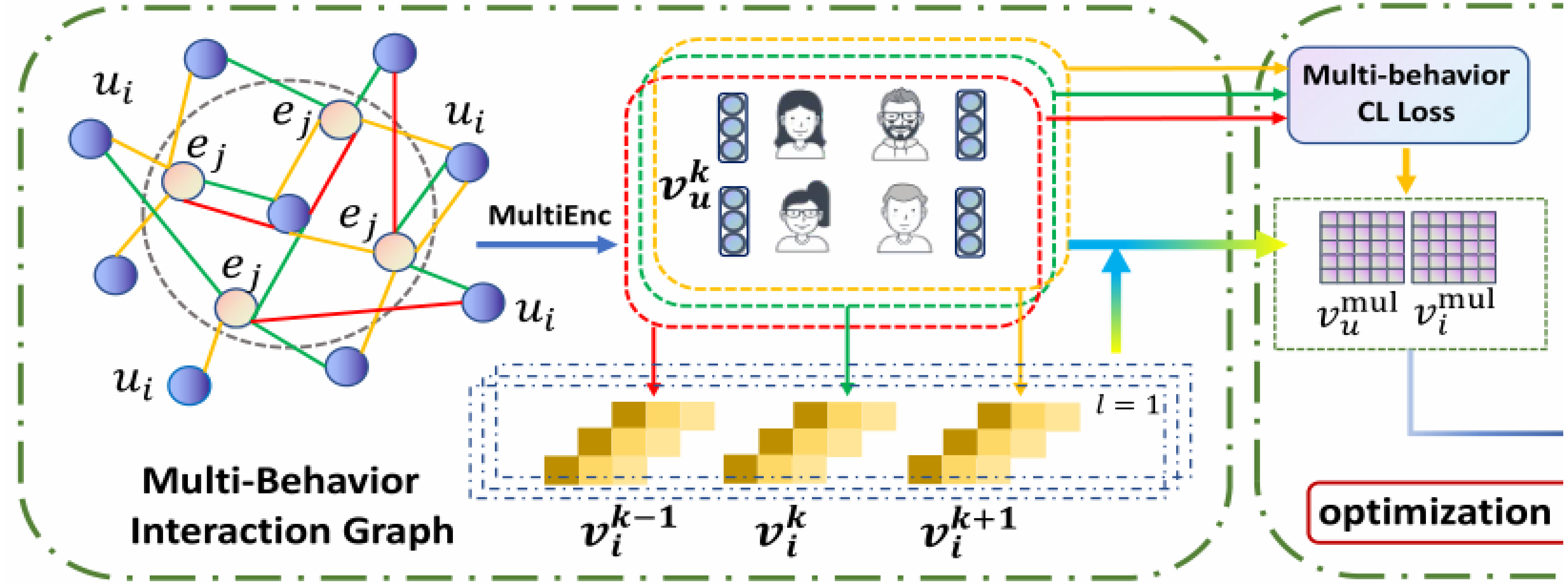


Figure 1: The model architecture of KMCLR framework.

Method



Multi-behavior Information Encoding and Aggregation:

$$v_u^k, v_i^k = MultiEnc^k(G_u, u, k), \quad (1)$$

$$v_u^{k,(l+1)} = \sum_{i \in N_u^k} v_i^{k,l}; \quad v_i^{k,(l+1)} = \sum_{u \in N_i^k} v_u^{k,l}, \quad (2)$$

$$v_u = PReLU((v_u^0||,...,||v_u^l||,...,||v_u^L) \times W^l), \quad (3)$$

$$v_u^l = \sigma(W^u \times mean(v_u^{1,l} \oplus,...,\oplus v_u^{k,l} \oplus,...,\oplus v_u^{K,l})). \quad (4)$$

PRELIMINARIES:

$$\mathcal{U} = \{u_1, ..., u_i, ..., u_I\} \quad \mathcal{V} = \{e_1, ..., e_j, ..., e_J\}$$

User-Item Multi-Behavior Interaction Graph:

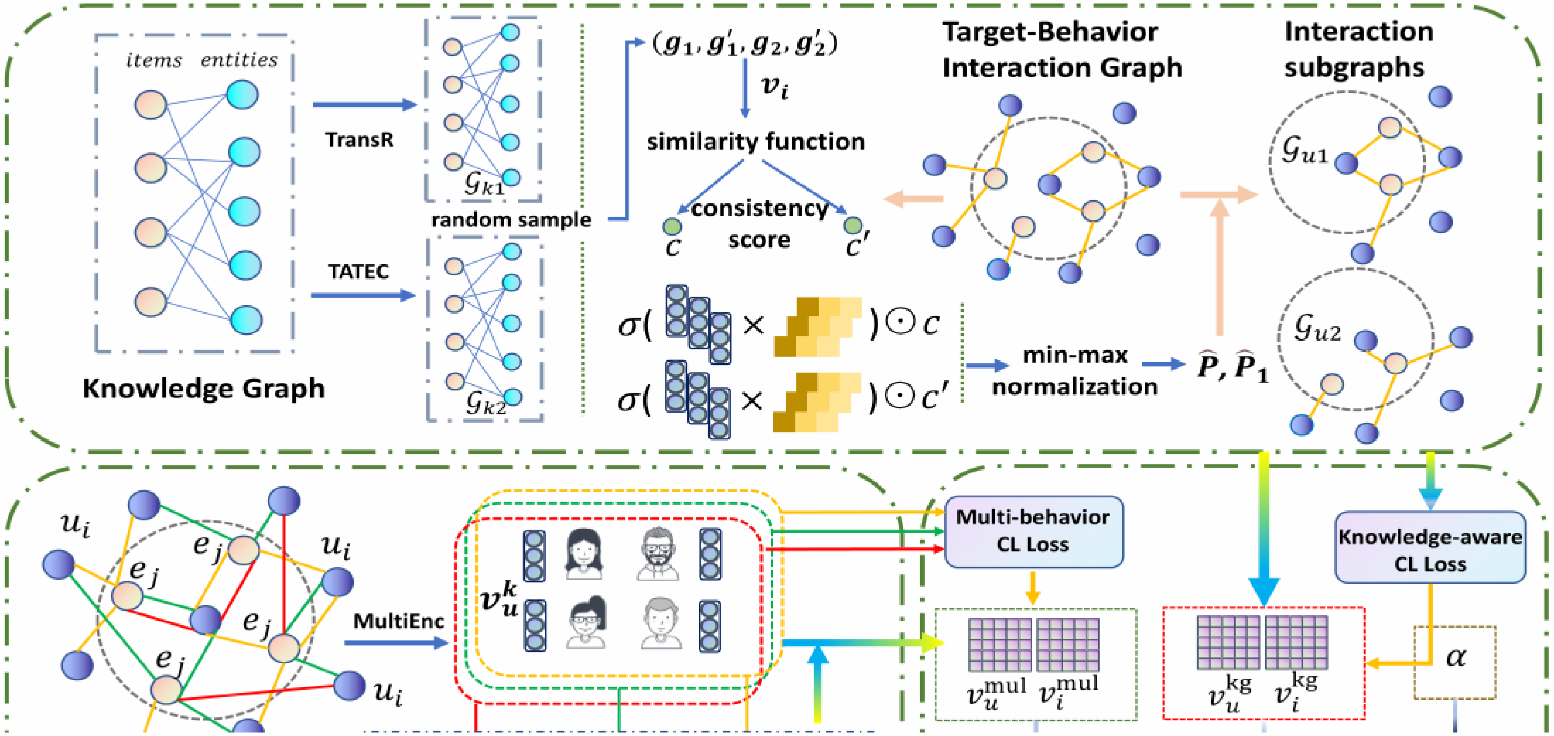
$$G_u = (\mathcal{U}, \Theta, \mathcal{V}) \quad \Theta = \{\theta^1, ..., \theta^k, ..., \theta^K\}$$

Item-Item Relation Knowledge Graph:

$$G_v = (H, R, T) \quad H, T \in \mathcal{V}$$

Multi-behavior Contrastive Learning:

$$\mathcal{L}_{MulCL} = \sum_{k'=1}^{K} \sum_{x \in \mathcal{U}} -log \frac{exp(s(v_x^k, v_x^{k'})/\tau)}{\sum_{y \in \mathcal{U}} exp(s(v_x^k, v_y^{k'})/\tau)}. \quad (5)$$



Knowledge-based Enhancement and Augmentation:

$$f_{td}(h,r,t) = -||M_r v_h + v_r - M_r v_t||_2^2, \tag{8}$$

$$f_{sm}(h,r,t) = v_h^T M_r v_t + v_h^T v_r + v_t^T v_r + v_h^T d D v_t \tag{9}$$

$$\mathcal{L}_{td} = \sum_{(h,r,t,t')\in\mathcal{G}_{k1}} -ln\sigma(f_{td}(h,r,t') - f_{td}(h,r,t)), \tag{10}$$

$$\mathcal{L}_{sm} = \sum_{(h,r,t,t')\in\mathcal{G}_{k2}} -ln\sigma(f_{sm}(h,r,t') - f_{sm}(h,r,t)), \tag{11}$$

$$c = s(g_1(v_i), g_1'(v_i)); (g_1, g_1' \in \mathcal{G}_{k1}^{sub}), \tag{12}$$

$$c' = s(g_2(v_i), g_2'(v_i)); (g_2, g_2' \in \mathcal{G}_{k2}^{sub}). \tag{13}$$

$$P_{u,i} = \sigma(v_u^T v_i) \odot c, \tag{14}$$

$$\hat{P_{u,i}} = (1 - Min\_Max(P_{u,i}))a + Min\_Max(P_{u,i})b. \tag{15}$$

Item-side Information Encoding and Aggregation:

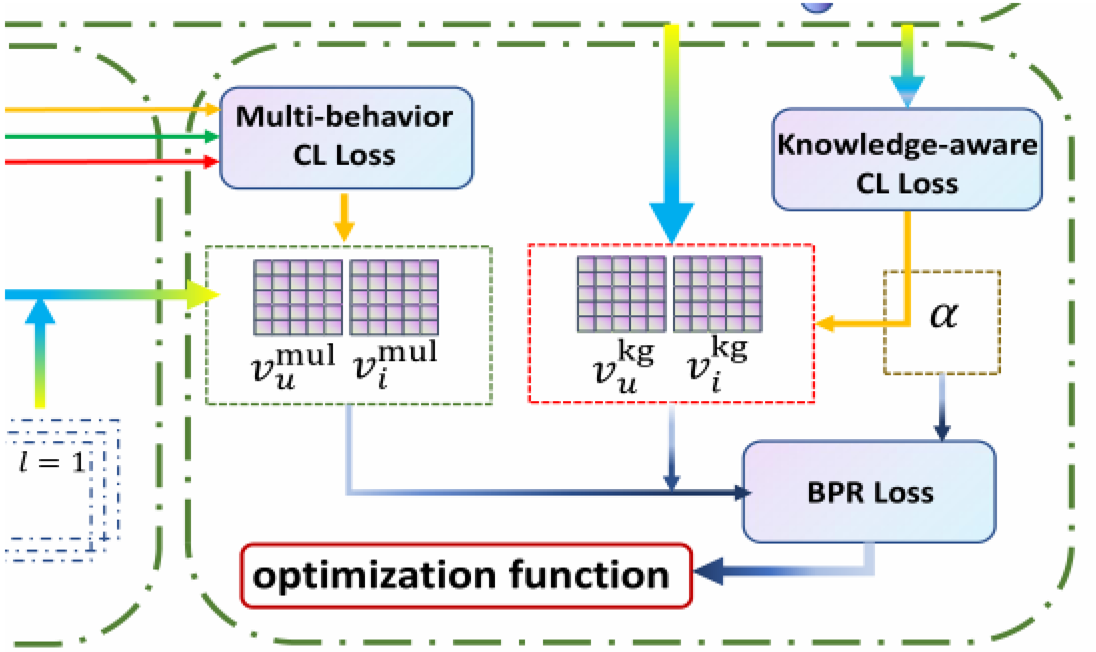
$$\mathcal{W}_{i,r,e} = \frac{f_{att}(v_i, v_r, v_e)}{\sum_{e'\in\mathcal{N}_i} f_{att}(v_i, v_r, v_{e'})}, \tag{6}$$

$$f_{att}(v_i, v_r, v_e) = exp[\sigma(W_1(v_i||v_r||v_e)) + b_1],$$

$$v_i^{l+1} = \sigma(W_2(v_i^l + \sum_{e\in\mathcal{N}_i} \mathcal{W}_{i,r,e} v_e) + b_2), \tag{7}$$

Knowledge-aware Contrastive Learning:

$$\mathcal{L}_{KCL} = \sum_{x\in\mathcal{U}\cup\mathcal{V}} -log\frac{exp(s(v_x^1, v_x^2)/\tau)}{\sum_{y\in\mathcal{U}\cup\mathcal{V}} exp(s(v_x^1, v_y^2)/\tau)}. \tag{16}$$



training paradigm:

$$argmin \triangleq main\_opt\{mul\_opt(L^{mul}, V^{mul}), kg\_opt(L^{kg}, V^{kg}), V^{main}\}. \tag{19}$$

$$L = \mathcal{L}_{BPR} + \lambda 1 \mathcal{L}_{cl} + \lambda 2 ||\Omega||_2^2, \tag{20}$$

Joint Learning Module:

$$v_u = (1-\alpha)v_u^{mul} + \alpha v_u^{kg}, \qquad v_i = (1-\alpha)v_i^{mul} + \alpha v_i^{kg}. \tag{17}$$

$$\mathcal{L}_{BPR} = -\sum_{u \in \mathcal{U}} \sum_{i \in \mathcal{N}_u} \sum_{j \notin \mathcal{N}_u} In\sigma(v_u^T v_i - v_u^T v_j), \tag{18}$$

Experiments

Table 1: Statistics of datasets

Dataset	Users	Items	Interactions	Interactive Behavior Type
Yelp	19,800	22,734	1,400,002	{Dislike, Neutral, Tip, Like}
Tmall	31,882	31,232	1.451,219	{Page View, Favorite, Cart, Purchase}
Retail	147,894	99,037	1,584,238	{Favorite, Cart, Purchase}

Table 2: Performance comparison on different datasets in terms of HR@10 and NDCG@10

Data	Metric	BPR	LightGCN	AutoRec	NGCF	KGAT	KGCL	NMTR	MATN	MBGCN	KHGT	CML	KMCLR
Yelp	HR	0.7435	0.7884	0.7592	0.7903	0.8211	0.8331	0.7880	0.8210	0.7962	0.8767	<u>0.8774</u>	**0.8897**
	NDCG	0.4497	0.4993	0.4705	0.5063	0.5362	0.5516	0.4673	0.5331	0.5033	<u>0.5970</u>	0.5965	**0.6038**
Tmall	HR	0.2443	0.3418	0.3274	0.3288	0.3853	0.4027	0.3623	0.4302	0.3910	0.4032	<u>0.5185</u>	**0.5671**
	NDCG	0.1501	0.2048	0.1903	0.1964	0.2219	0.2293	0.2097	0.2437	0.2261	0.2351	<u>0.3055</u>	**0.3540**
Retail	HR	0.2608	0.3065	0.2779	0.3027	0.3485	0.3502	0.3112	0.3490	0.3137	0.4199	<u>0.4256</u>	**0.4557**
	NDCG	0.1653	0.1847	0.1683	0.1844	0.2191	0.2197	0.1869	0.2183	0.1904	0.2471	<u>0.2503</u>	**0.2735**

Experiments

Table 3: Results of ablation experiments

model \ dataset	Yelp		Tmall		Retail	
	HR	NDCG	HR	NDCG	HR	NDCG
w/o-Mcl	0.8352	0.5749	0.4583	0.2645	0.3612	0.2157
w/o-Kcl	0.8687	0.5916	0.5231	0.3064	0.4247	0.2513
w/o NorT	0.8701	0.5943	0.5339	0.3106	0.4268	0.2502
KMCLR	**0.8897**	**0.6038**	**0.5671**	**0.3540**	**0.4557**	**0.2735**

Experiments

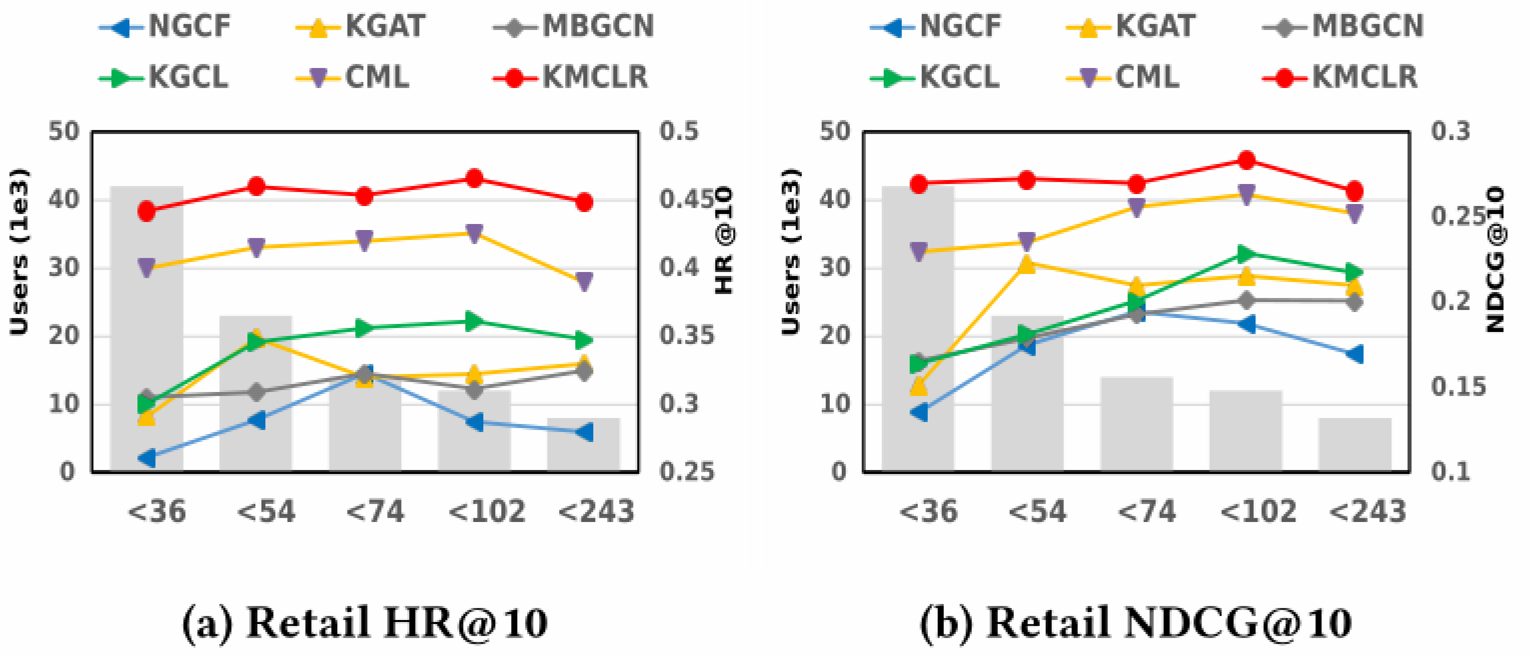


Figure 2: Performance of KMCLR and baseline methods w.r.t different data sparsity degrees on Retail data.

Experiments

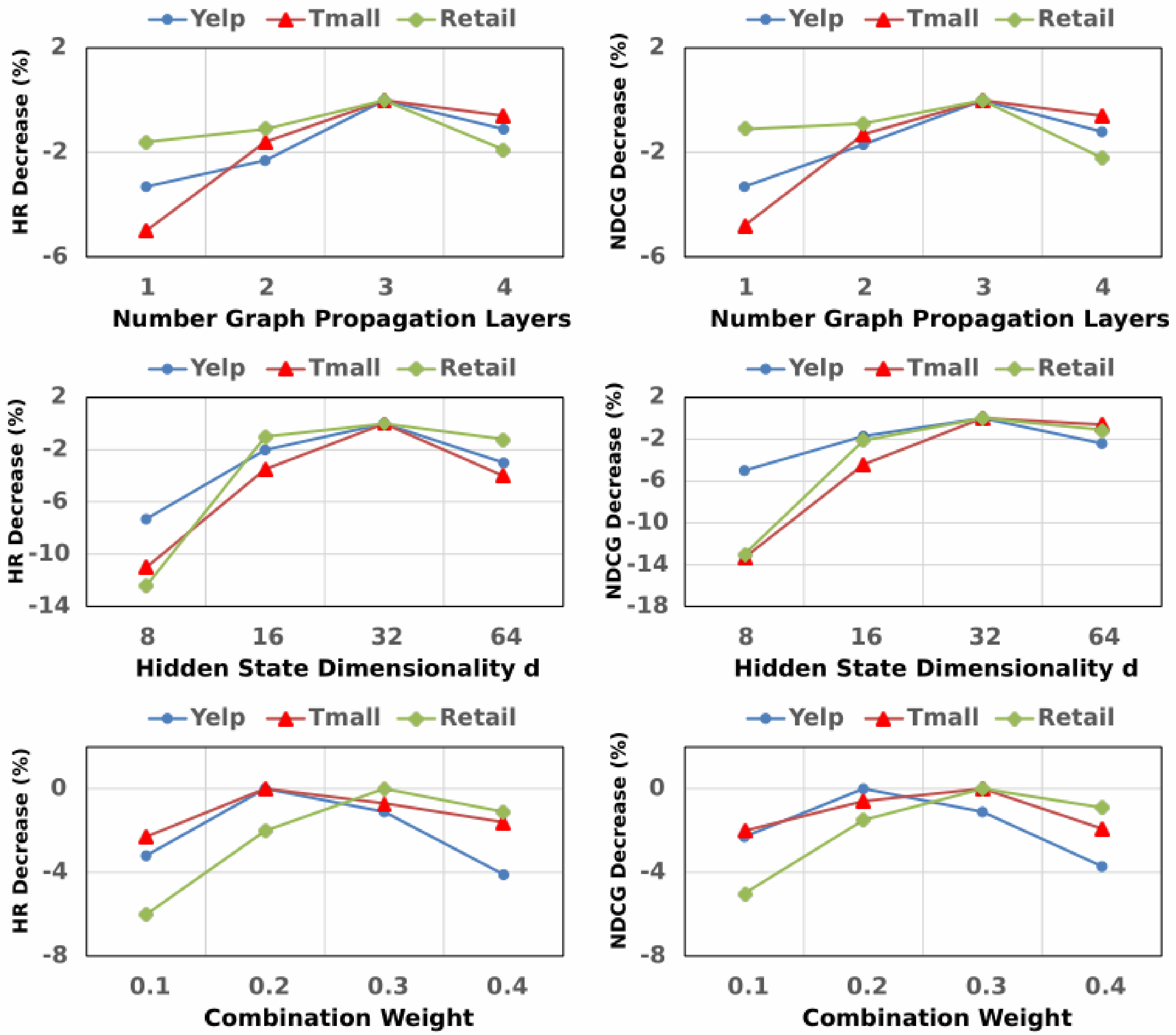
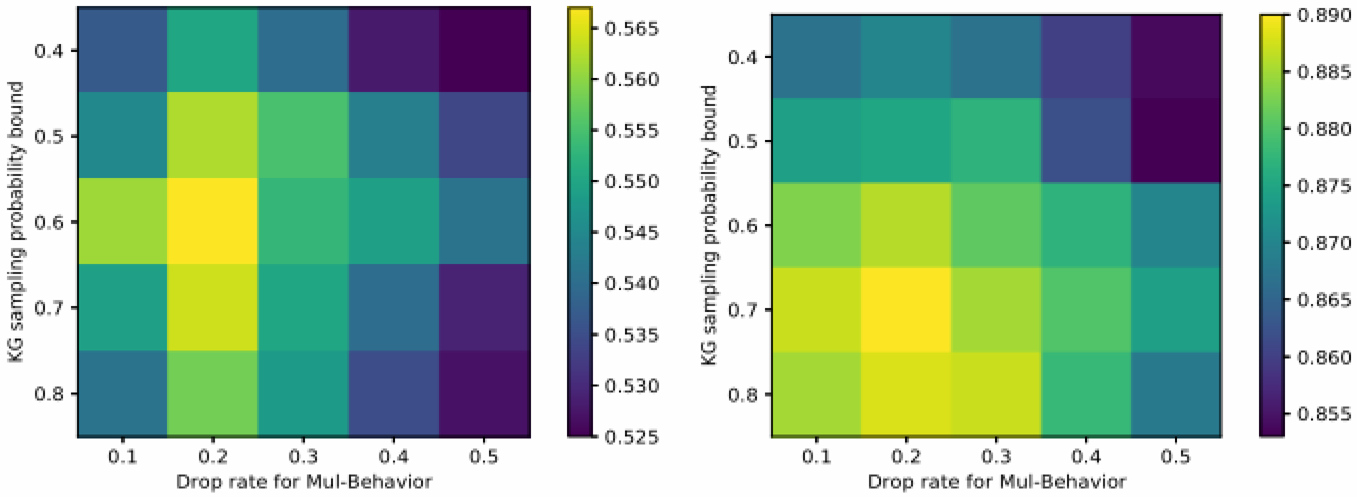


Figure 3: Hyperparameter analyses of KMCLR.

Experiments



(a) Tmall HR@10　　(b) Yelp HR@10

Figure 4: Parameter settings to alleviate noise.