Multi-level Cross-view Contrastive Learning for Knowledge-aware Recommender System

Ding Zou[1,6], Wei Wei[1,6 4], Xian-Ling Mao[2], Ziyang Wang[1,6], Minghui Qiu[3], Feida Zhu[4], Xin Cao[5]
[1] Cognitive Computing and Intelligent Information Processing (CCIIP) Laboratory, School of Computer Science and Technology, Huazhong University of Science and Technology, China
[2] School of Computer Science and Technology, Beijing Institute of Technology, China
[3] Alibaba Group, China
[4] Singapore Management University, Singapore
[5] School of Computer Science and Engineering, The University of New South Wales, Australia
[6]Joint Laboratory of HUST and Pingan Property & Casualty Research (HPL), China
[1] {m202173662, weiw, ziyang1997 }@hust.edu.cn [2] maoxl@bit.edu.cn [3] minghuiqiu@gmail.com
4fdzhu@smu.edu.sg 5 xin.cao@unsw.edu.au

——SIGIR2022

https://github.com/CCIIPLab/MCCLK

Leibniz-Institut für Sozialwissenschaften

Reported by Yabo Yin

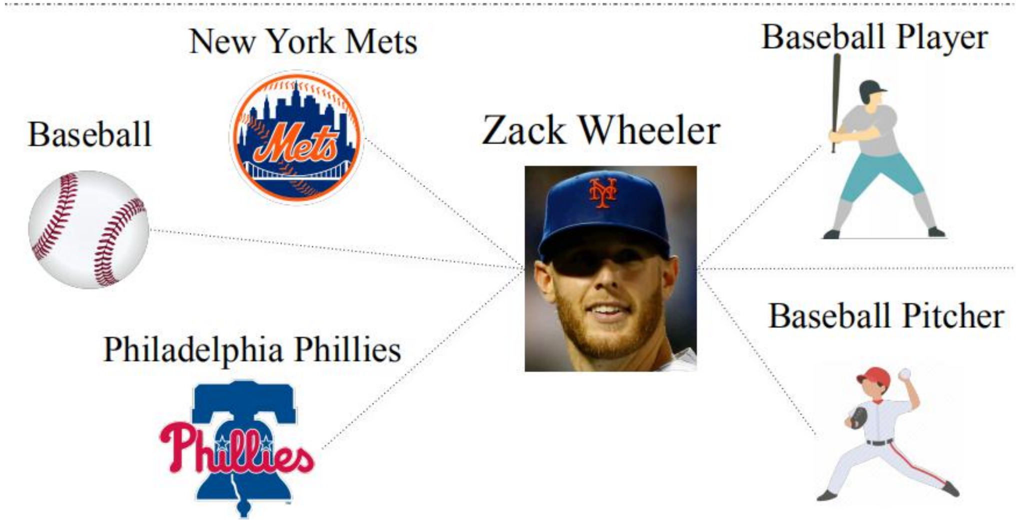
Artificial

1.Introduction

2.Method

3.Experiments

Introduction



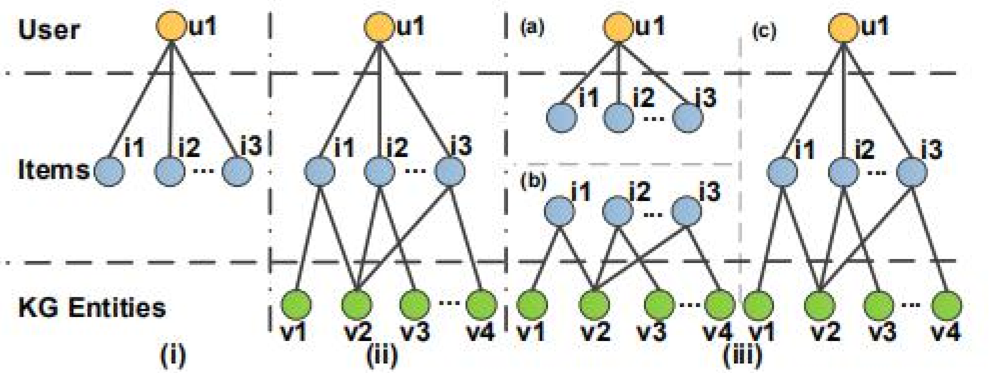


Figure 1: A toy example of our selected multi views. (i) Traditional CF-based recommendation learns from collaborative view. (ii) Previous KGR methods learn from the structural view. (iii) MCCLK learns from three selected views, including local-level collaborative view (a) and semantic view (b), global-level structural view (c).

1. Current GNN-based models greatly suffer from sparse supervision signal problem. leading to the indiscrimination of generated node representations.
2. Since previously models treat each item entity relation independently for learning. the learning process is incapable of distilling

Method

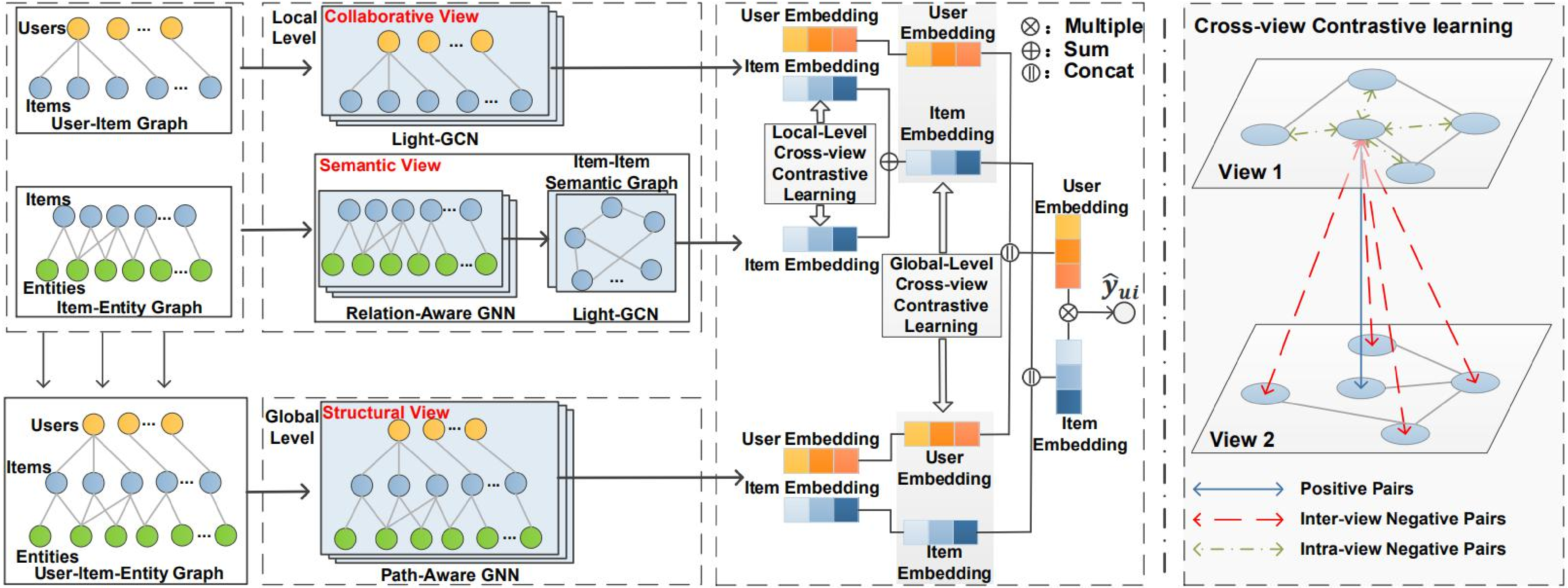


Figure 2: Illustration of the proposed MCCLK model. The left subfigure shows model framework of MCCLK; and the right subfigure presents the details of cross-view contrastive learning mechanism. Best viewed in color.

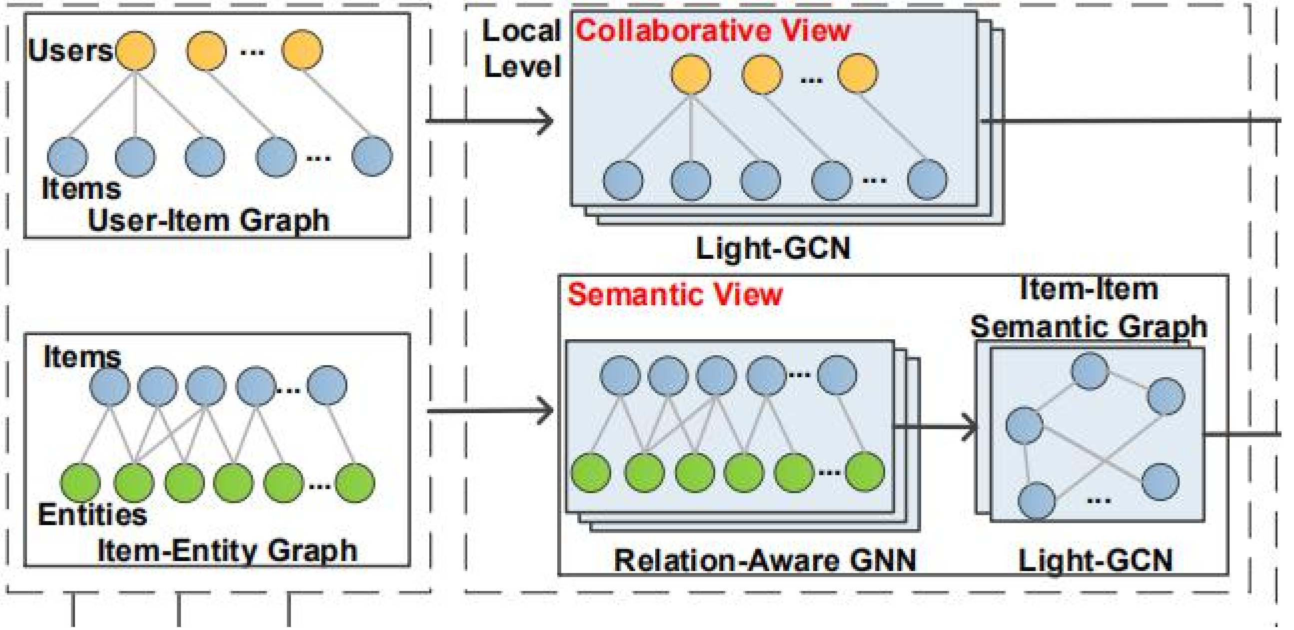
Method

Problem Definition

$\mathcal{U} = \{u_1, u_2, \ldots, u_M\}$ and $\mathcal{V} = \{v_1, v_2, \ldots, v_N\}$

user-item interaction matrix $\mathrm{Y} \in \mathrm{R}^{M \times N}$

$\mathcal{G} = \{(h, r, t) \mid h, t \in \mathcal{E}, r \in \mathcal{R}\}$



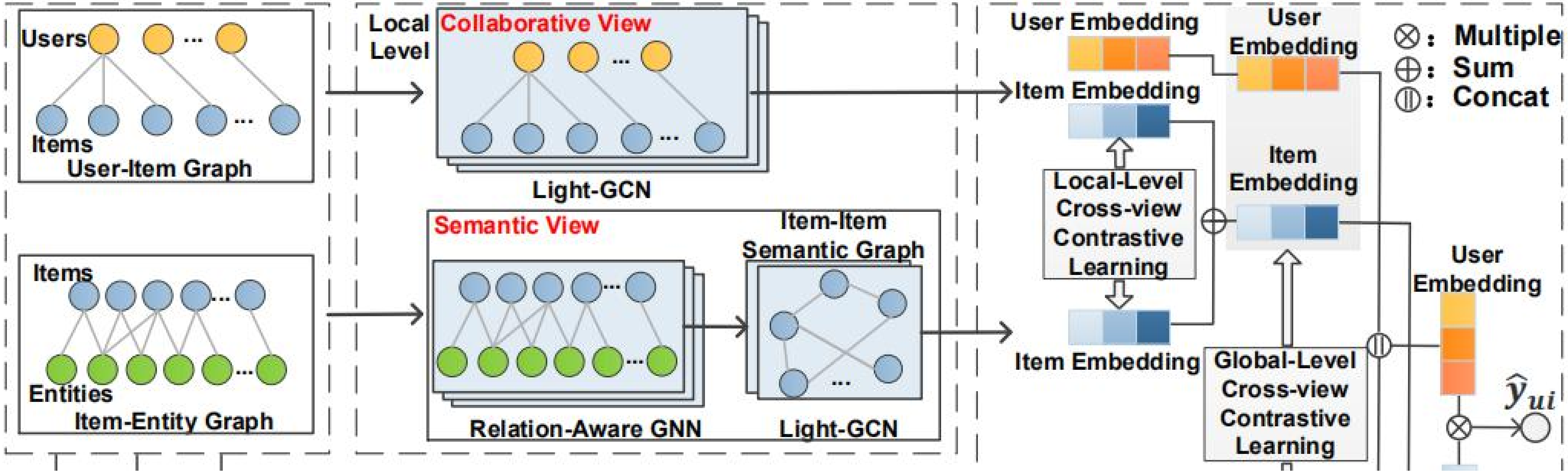
$$\mathbf{e}_i^{(k+1)} = \frac{1}{|\mathcal{N}_i|} \sum_{(r,v) \in \mathcal{N}_i} \mathbf{e}_r \odot \mathbf{e}_v^{(k)},$$
$$\mathbf{e}_v^{(k+1)} = \frac{1}{|\mathcal{N}_v|} \left(\sum_{(r,v) \in \mathcal{N}_v} \mathbf{e}_r \odot \mathbf{e}_v^{(k)} + \sum_{(r,i) \in \mathcal{N}_v} \mathbf{e}_r \odot \mathbf{e}_i^{(k)} \right), \tag{1}$$

$$S_{ij} = \frac{\left(\mathbf{e}_i^{(K')}\right)^{\top} \mathbf{e}_j^{(K')}}{\left\|\mathbf{e}_i^{(K')}\right\| \left\|\mathbf{e}_j^{(K')}\right\|}. \tag{2}$$

$$\widehat{S}_{ij} = \begin{cases} S_{ij}, & S_{ij} \in \text{top-k}(S_i), \\ 0, & \text{otherwise}, \end{cases} \tag{3}$$

$$\widetilde{S} = (D)^{-\frac{1}{2}} \widehat{S} (D)^{-\frac{1}{2}}, \tag{4}$$

Method



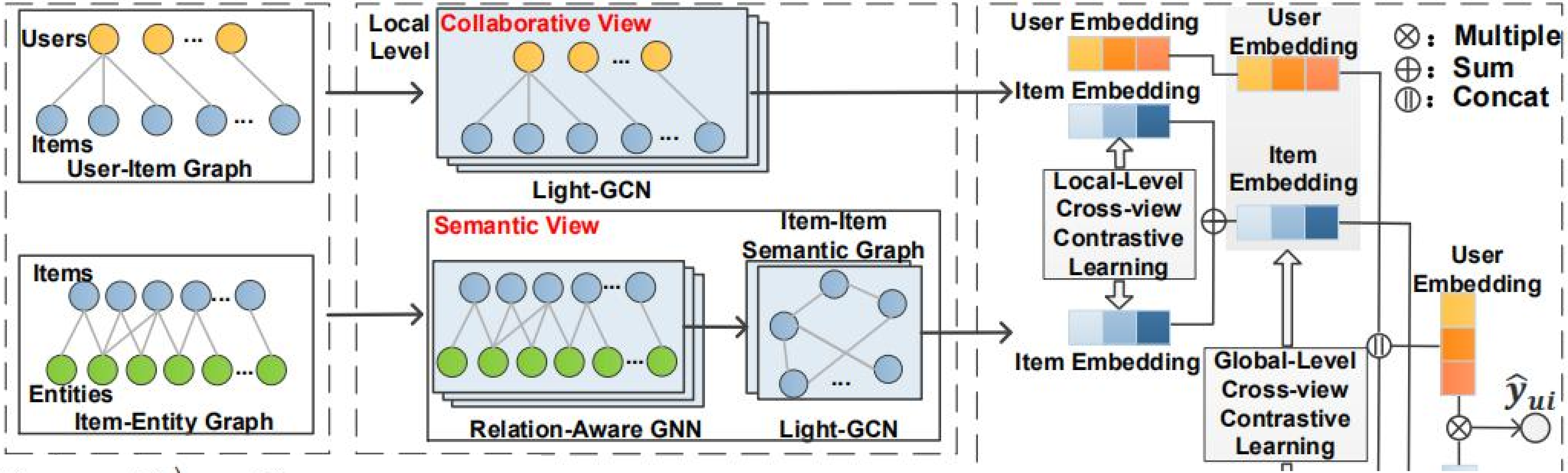
$$\mathbf{e}_u^{(k+1)} = \sum_{i \in \mathcal{N}_u} \frac{1}{\sqrt{|\mathcal{N}_u||\mathcal{N}i|}} \mathbf{e}_i^{(k)},$$
$$\mathbf{e}_i^{(k+1)} = \sum_{u \in \mathcal{N}_i} \frac{1}{\sqrt{|\mathcal{N}_u||\mathcal{N}i|}} \mathbf{e}_u^{(k)}, \quad (5)$$

$$\mathbf{e}_i^{(l+1)} = \sum_{j \in \mathcal{N}(i)} \widetilde{S} \mathbf{e}_j^{(l)}, \quad (7)$$

$$\mathbf{z}_u^c = \mathbf{e}_u^{(0)} + \cdots + \mathbf{e}_u^{(K)}, \quad \mathbf{z}_i^c = \mathbf{e}_i^{(0)} + \cdots + \mathbf{e}_i^{(K)}. \quad (6)$$

$$\mathbf{z}_i^s = \mathbf{e}_i^{(0)} + \cdots + \mathbf{e}_i^{(L)}. \quad (8)$$

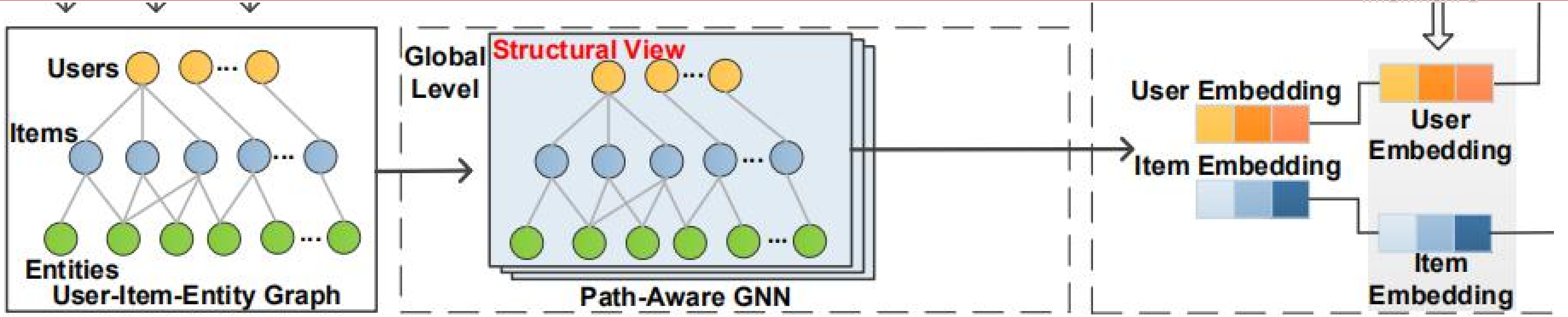
Method



$$\mathbf{z}_i^c\_\mathrm{p} = W^{(2)}\sigma\left(W^{(1)}\mathbf{z}_i^c + b^{(1)}\right) + b^{(2)},$$
$$\mathbf{z}_i^s\_\mathrm{p} = W^{(2)}\sigma\left(W^{(1)}\mathbf{z}_i^s + b^{(1)}\right) + b^{(2)}, \quad (9)$$

$$\mathcal{L}^{local} = -\log \frac{e^{s\left(\mathbf{z}_i^s\_\mathrm{p},\mathbf{z}_i^c\_\mathrm{p}\right)/\tau}}{\underbrace{e^{s\left(\mathbf{z}_i^s\_\mathrm{p},\mathbf{z}_i^c\_\mathrm{p}\right)/\tau}}_{\text{positive pair}} + \underbrace{\sum_{k\neq i} e^{s\left(\mathbf{z}_i^s\_\mathrm{p},\mathbf{z}_k^s\_\mathrm{p}\right)/\tau}}_{\text{intra-view negative pairs}} + \underbrace{\sum_{k\neq i} e^{s\left(\mathbf{z}_i^s\_\mathrm{p},\mathbf{z}_k^c\_\mathrm{p}\right)/\tau}}_{\text{inter-view negative pairs}}}, \quad (10)$$

Method



$$\mathbf{e}_u^{(l+1)} = \frac{1}{|\mathcal{N}_u|} \sum_{i \in \mathcal{N}_u} \mathbf{e}_i^{(l)},$$
$$\mathbf{e}_i^{(l+1)} = \frac{1}{|\mathcal{N}_i|} \sum_{(r,v) \in \mathcal{N}_i} \beta(i,r,v)\mathbf{e}_r \odot \mathbf{e}_v^{(l)}, \tag{11}$$

$$\begin{aligned}\beta(i,r,v) &= \mathrm{softmax}\left((\mathbf{e}_i||\mathbf{e}_r)^T \cdot (\mathbf{e}_v||\mathbf{e}_r)\right) \\ &= \frac{\exp\left((\mathbf{e}_i||\mathbf{e}_r)^T \cdot (\mathbf{e}_v||\mathbf{e}_r)\right)}{\sum_{(v',r) \in \hat{\mathrm{N}}(i)} \exp\left((\mathbf{e}_i||\mathbf{e}_r)^T \cdot (\mathbf{e}_{v'}||\mathbf{e}_r)\right)},\end{aligned} \tag{12}$$

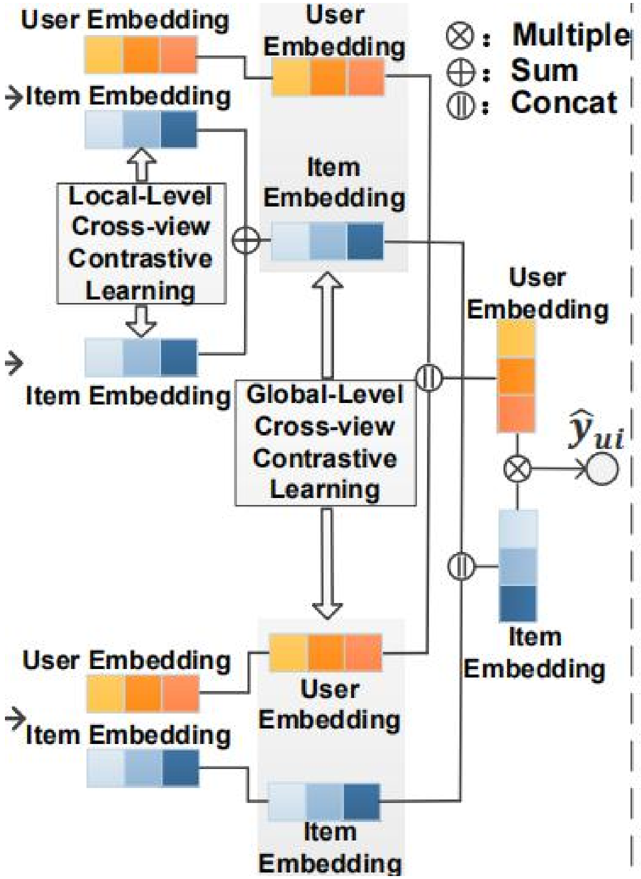
$$\mathbf{z}_u^g = \mathbf{e}_u^{(0)} + \cdots + \mathbf{e}_u^{(L')}, \quad \mathbf{z}_i^g = \mathbf{e}_i^{(0)} + \cdots + \mathbf{e}_i^{(L')}. \tag{13}$$

Method

$$\mathcal{L}_i^g = -\log \frac{e^{s\left(\mathbf{z}_i^g\_\mathrm{p},\mathbf{z}_i^l\_\mathrm{p}\right)/\tau}}{\underbrace{e^{s\left(\mathbf{z}_i^g\_\mathrm{p},\mathbf{z}_i^l\_\mathrm{p}\right)/\tau}}_{\text{positive pair}} + \underbrace{\sum_{k\neq i} e^{s\left(\mathbf{z}_i^g\_\mathrm{p},\mathbf{z}_k^g\_\mathrm{p}\right)/\tau}}_{\text{intra-view negative pairs}} + \underbrace{\sum_{k\neq i} e^{s\left(\mathbf{z}_i^g\_\mathrm{p},\mathbf{z}_k^l\_\mathrm{p}\right)/\tau}}_{\text{inter-view negative pairs}}},$$

$$\mathcal{L}_i^l = -\log \frac{e^{s\left(\mathbf{z}_i^l\_\mathrm{p},\mathbf{z}_i^g\_\mathrm{p}\right)/\tau}}{\underbrace{e^{s\left(\mathbf{z}_i^l\_\mathrm{p},\mathbf{z}_i^g\_\mathrm{p}\right)/\tau}}_{\text{positive pair}} + \underbrace{\sum_{k\neq i} e^{s\left(\mathbf{z}_i^l\_\mathrm{p},\mathbf{z}_k^l\_\mathrm{p}\right)/\tau}}_{\text{intra-view negative pairs}} + \underbrace{\sum_{k\neq i} e^{s\left(\mathbf{z}_i^l\_\mathrm{p},\mathbf{z}_k^g\_\mathrm{p}\right)/\tau}}_{\text{inter-view negative pairs}}}, \quad (15)$$

$$\mathcal{L}^{global} = \frac{1}{2N}\sum_{i=1}^{N}(\mathcal{L}_i^g + \mathcal{L}_i^l) + \frac{1}{2M}\sum_{i=1}^{M}(\mathcal{L}_u^g + \mathcal{L}_u^l). \quad (16)$$

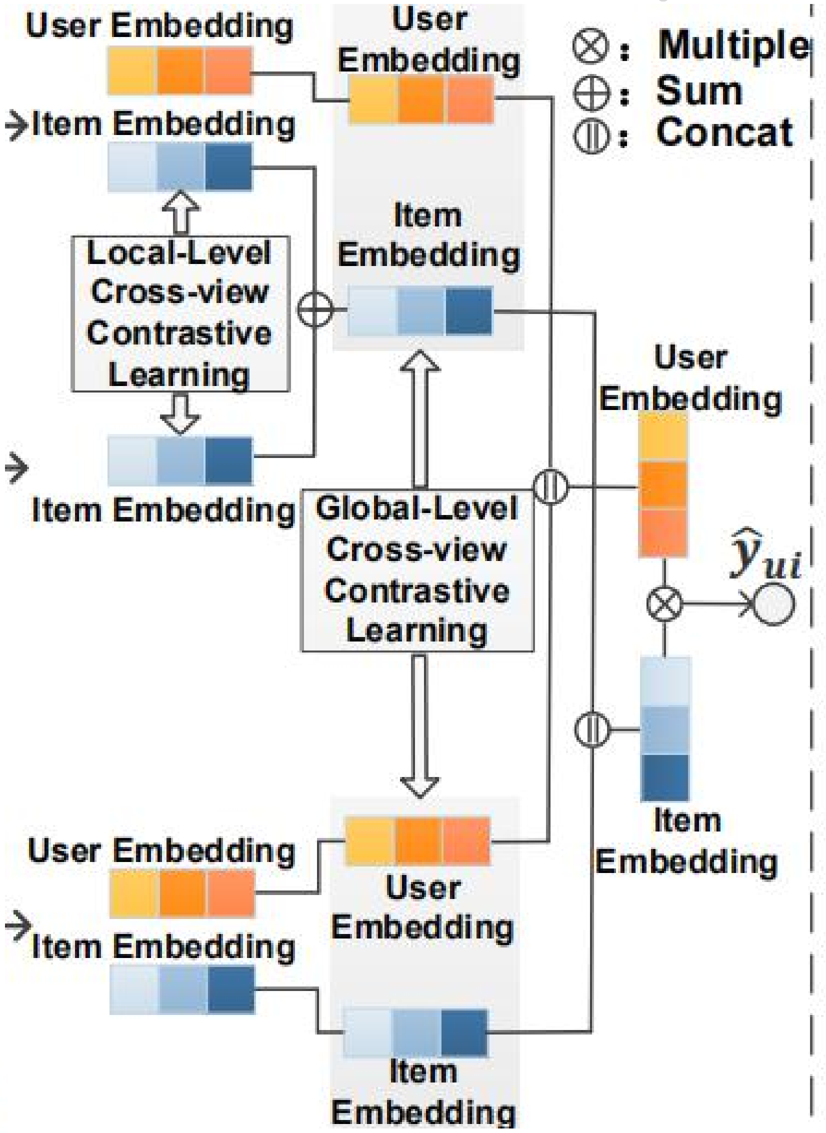


Method

$$
\begin{aligned}
\mathbf{z}_u^* &= \mathbf{z}_u^g || \mathbf{z}_u^c, \\
\mathbf{z}_i^* &= \mathbf{z}_i^g || (\mathbf{z}_i^c + \mathbf{z}_i^s), \\
\hat{\mathbf{y}}(u, i) &= \mathbf{z}_u^{*\top} \mathbf{z}_i^* .
\end{aligned}
\tag{17}
$$

$$
\mathcal{L}_{\mathrm{BPR}} = \sum_{(u,i,j) \in O} -\ln \sigma \left(\hat{\mathbf{y}}_{ui} - \hat{\mathbf{y}}_{uj} \right), \tag{18}
$$

$$
\mathcal{L}_{MCCLK} = \mathcal{L}_{\mathrm{BPR}} + \beta(\alpha \mathcal{L}^{local} + (1-\alpha)\mathcal{L}^{global}) + \lambda ||\Theta||_2^2, \tag{19}
$$



Experiments

		Book-Crossing	MovieLens-1M	Last.FM
User-item Interaction	# users	17,860	6,036	1,872
	# items	14,967	2,445	3,846
	# interactions	139,746	753,772	42,346
Knowledge Graph	# entities	77,903	182,011	9,366
	# relations	25	12	60
	# triplets	151,500	1,241,996	15,518
Hyper-parameter Settings	# α	0.2	0.2	0.2
	# β	0.1	0.1	0.1
	# K	2	2	3
	# K'	2	2	2
	# L	1	1	2
	# L'	2	2	2

Table 1: Statistics and hyper-parameter settings for the three datasets.(α: local-level contrastive loss weight, β: contrastive loss weight, K: local collaborative aggregation depth, K': aggregation depth of item-item semantic graph construction, L: local semantic aggregation depth, L': global structural aggregation depth.)

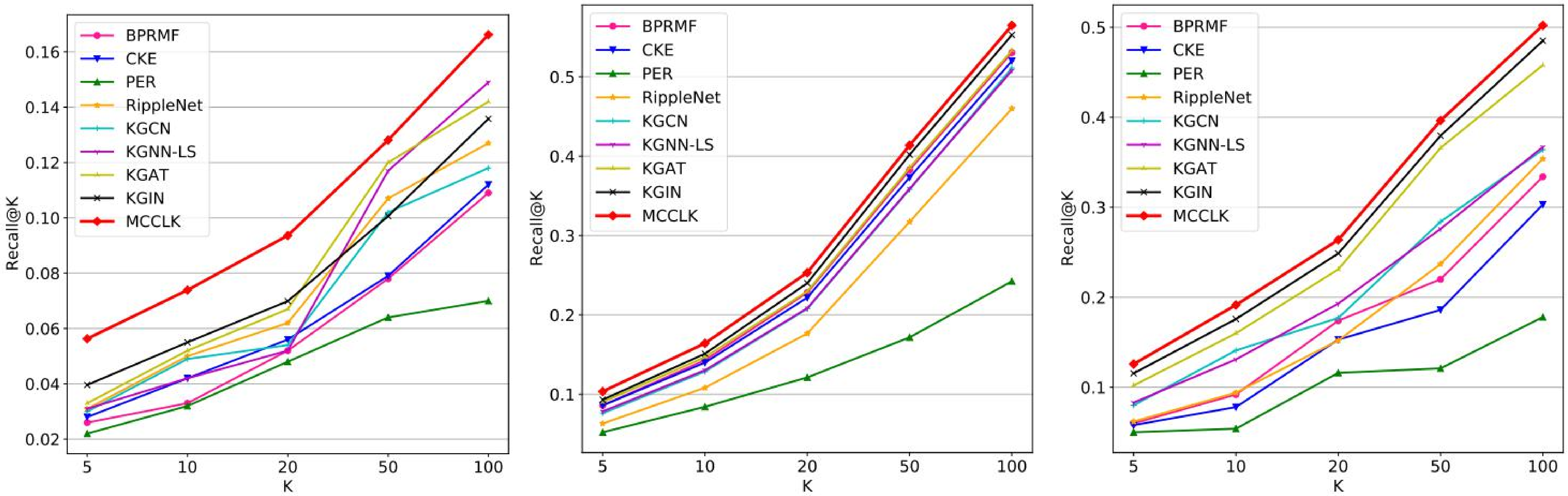
Experiments

Model	Book-Crossing *AUC*	Book-Crossing *F1*	MovieLens-1M *AUC*	MovieLens-1M *F1*	Last.FM *AUC*	Last.FM *F1*
BPRMF	0.6583(−10.42%)	0.6117(−6.60%)	0.8920(−4.31%)	0.7921(−7.10%)	0.7563(−12.00%)	0.7010(−9.98%)
CKE	0.6759(−8.66%)	0.6235(−5.42%)	0.9065(−2.86%)	0.8024(−6.07%)	0.7471(−12.92%)	0.6740(−12.68%)
RippleNet	0.7211(−4.14%)	0.6472(−3.05%)	0.9190(−1.61%)	0.8422(−2.09%)	0.7762(−10.01%)	0.7025(−9.83%)
PER	0.6048(−15.77%)	0.5726(−10.51%)	0.7124(−22.27%)	0.6670(−19.61%)	0.6414(−23.49%)	0.6033(−19.75%)
KGCN	0.6841(−7.84%)	0.6313(−4.64%)	0.9090(−2.61%)	0.8366(−2.65%)	0.8027(−7.36%)	0.7086(−9.22%)
KGNN-LS	0.6762(−8.63%)	0.6314(−4.63%)	0.9140(−2.11%)	0.8410(−2.21%)	0.8052(−7.11%)	0.7224(−7.84%)
KGAT	0.7314(−3.11%)	0.6544(−2.33%)	0.9140(−2.11%)	0.8440(−1.91%)	0.8293(−4.70%)	0.7424(−5.84%)
KGIN	0.7273(−3.52%)	0.6614(−1.63%)	0.9190(−1.61%)	0.8441(−1.90%)	0.8486(−2.77%)	0.7602(−4.06%)
MCCLK	**0.7625***	**0.6777***	**0.9351***	**0.8631***	**0.8763***	**0.8008***

Table 2: The result of AUC and $F1$ in CTR prediction. The best results are in boldface and the second best results are underlined.
*** denotes statistically significant improvement by unpaired two-sample t-test with $p < 0.001$.**

OOM	OOM	0.9360	0.8629	0.8804	0.8010

Experiments



(a) Book-Crossing (b) MovieLens-1M (c) Last.FM

Figure 3: The result of Recall@K in top-K recommendation.

Experiments

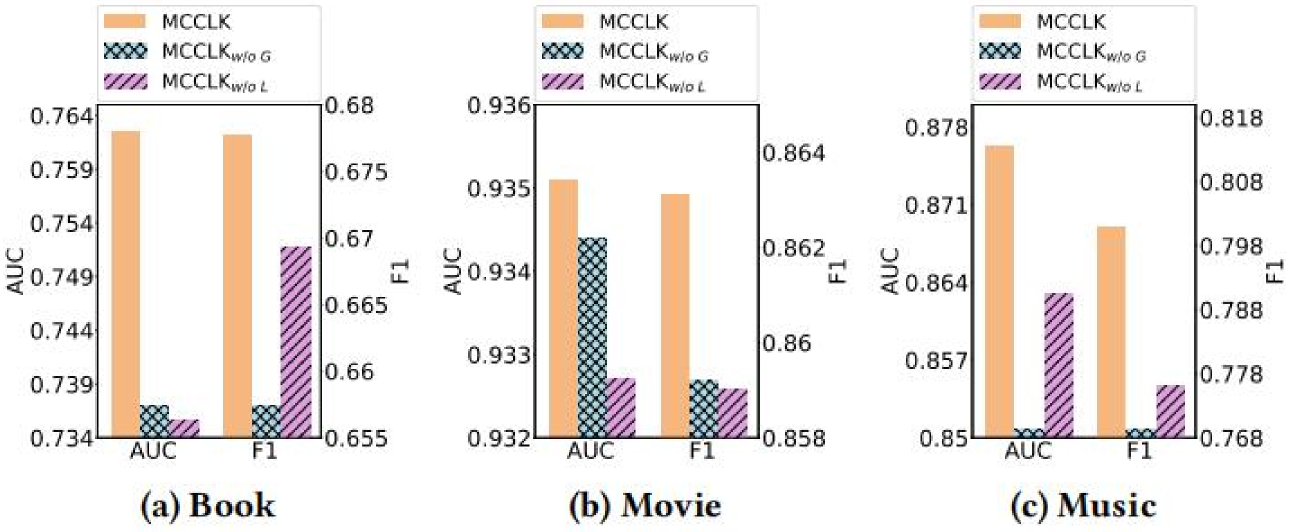


Figure 4: Effect of ablation study.

	Book		Movie		Music	
	Auc	F1	Auc	F1	Auc	F1
L=1	**0.7602**	**0.6777**	**0.9350**	**0.8631**	0.8711	0.7858
L=2	0.7601	0.6768	0.9347	0.8628	**0.8742**	**0.7945**
L=3	0.7591	0.6733	0.9345	0.8627	0.8726	0.7891
L=4	0.7583	0.6749	0.9343	0.8627	0.8720	0.7846

Table 3: Impact of aggregation depth in semantic view.

Experiments

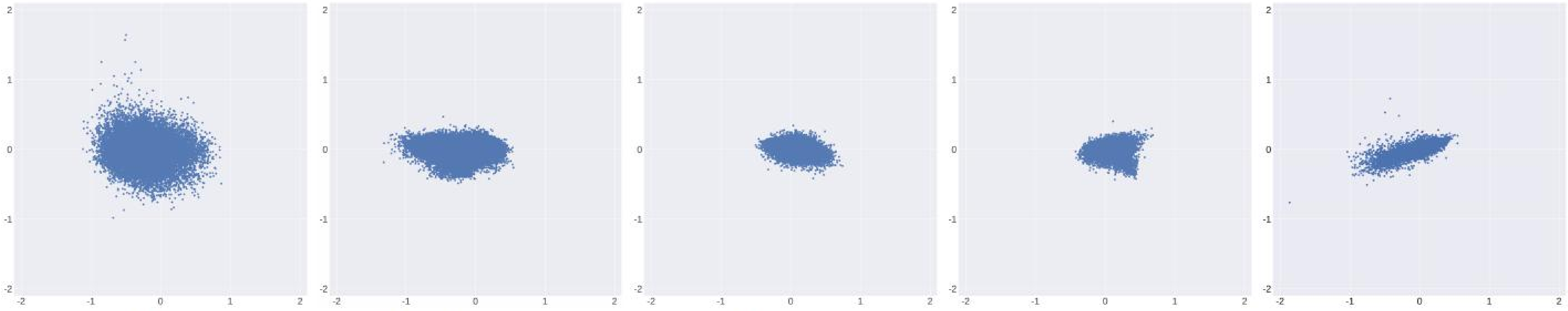
	Book		Movie		Music	
	Auc	F1	Auc	F1	Auc	F1
L'=1	0.7602	0.6776	0.9350	0.8628	0.8711	0.7858
L'=2	**0.7625**	**0.6777**	**0.9351**	**0.8631**	**0.8763**	**0.8008**
L'=3	0.7550	0.6719	0.9334	0.8589	0.8713	0.7899
L'=4	0.7569	0.6680	0.9320	0.8574	0.8706	0.7841

Table 4: Impact of aggregation depth in structural view.

	Book		Movie		Music	
	Auc	F1	Auc	F1	Auc	F1
β=1	0.7520	0.6649	0.9337	0.8593	0.8735	0.7938
β=0.1	**0.7625**	**0.6713**	**0.9351**	**0.8622**	**0.8758**	**0.7972**
β=0.01	0.7608	0.6689	0.9346	0.8610	0.8721	0.7913
β=0.001	0.7607	0.6675	0.9343	0.8604	0.8714	0.7856

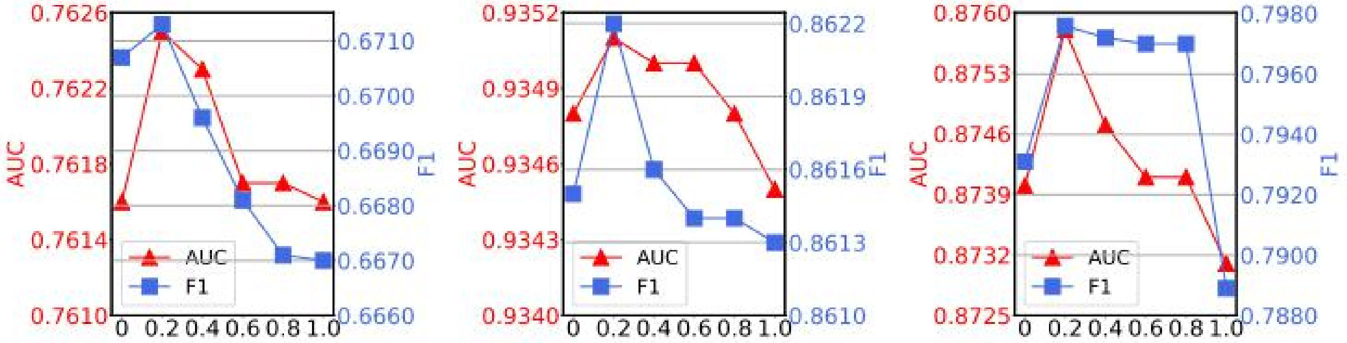
Table 5: Impact of contrastive loss weight β.

Experiments



(a) MCCLK (b) MCCLK$_{w/o\ G}$ (c) MCCLK$_{w/o\ L}$ (d) KGIN (e) RippleNet

Figure 5: Visualization of model representation learning ability on Book-Crossing.



(a) Book (b) Movie (c) Music

Figure 6: Impact of local-level contrastive loss weight α.

Thank you!