LIGHTGCL: SIMPLE YET EFFECTIVE GRAPH CONTRASTIVE LEARNING FOR RECOMMENDATION

Xuheng Cai Chao Huang* Lianghao Xia Xubin Ren
Department of Computer Science, University of Hong Kong
{rickcai, lhaoxia}@hku.hk chaohuang75gmail.com xubinrencs@gmail.com

https://github.com/HKUDS/LightGCL

(ICLR-2022)

gesis
Leibniz-Institut für Sozialwissenschaften

Reported by Ke Gan

1. Introduction
2. Approach
3. Experiments

Introduction

i) Graph augmentation with random perturbation may lose useful structural information, which misleads the representation learning

ii) The success of heuristic-guided representation contrasting schemes is largely built upon the view generator, which limits the model generality and is vulnerable to the noisy user behaviors.

iii) Most of current GNN-based contrastive recommenders are limited by the over-smoothing issue which leads to indistinguishable representations.

Approach

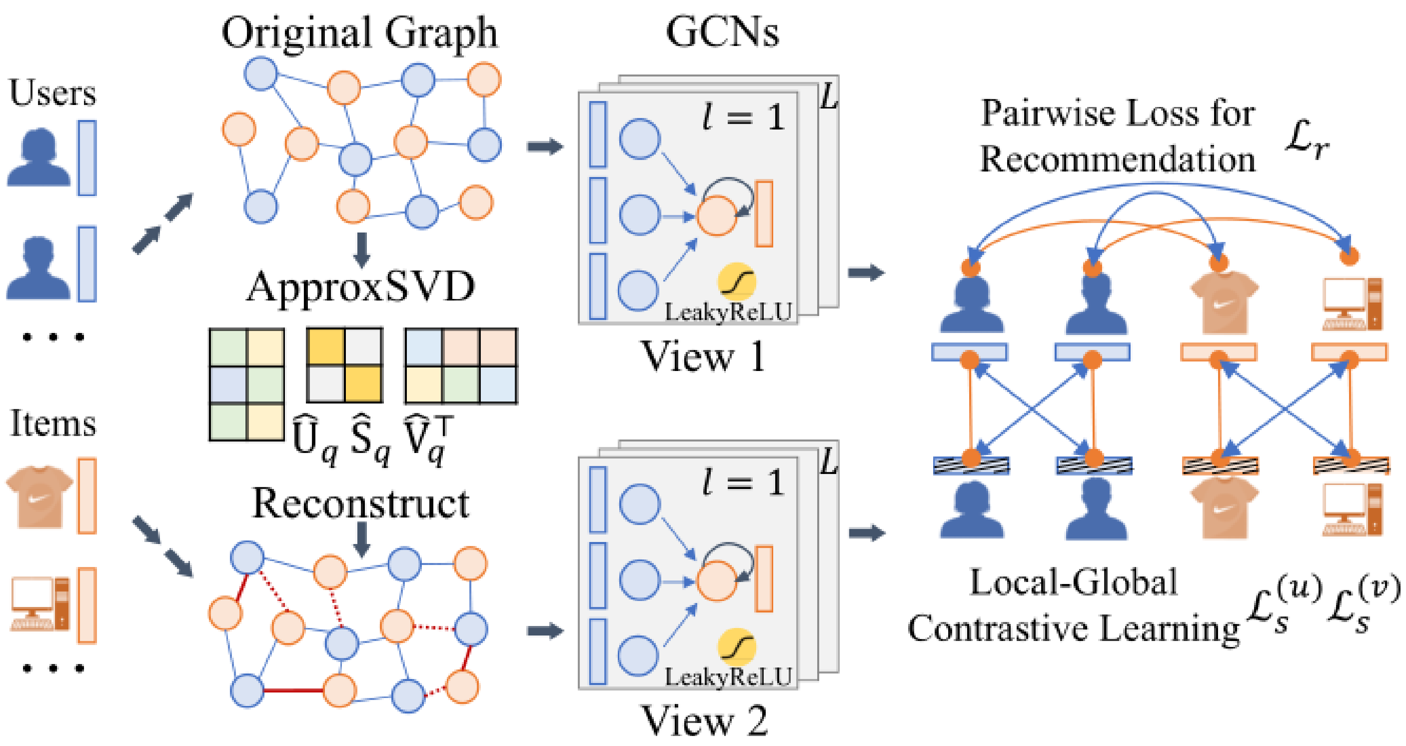

Figure 1: Overall structure of LightGCL.

Approach

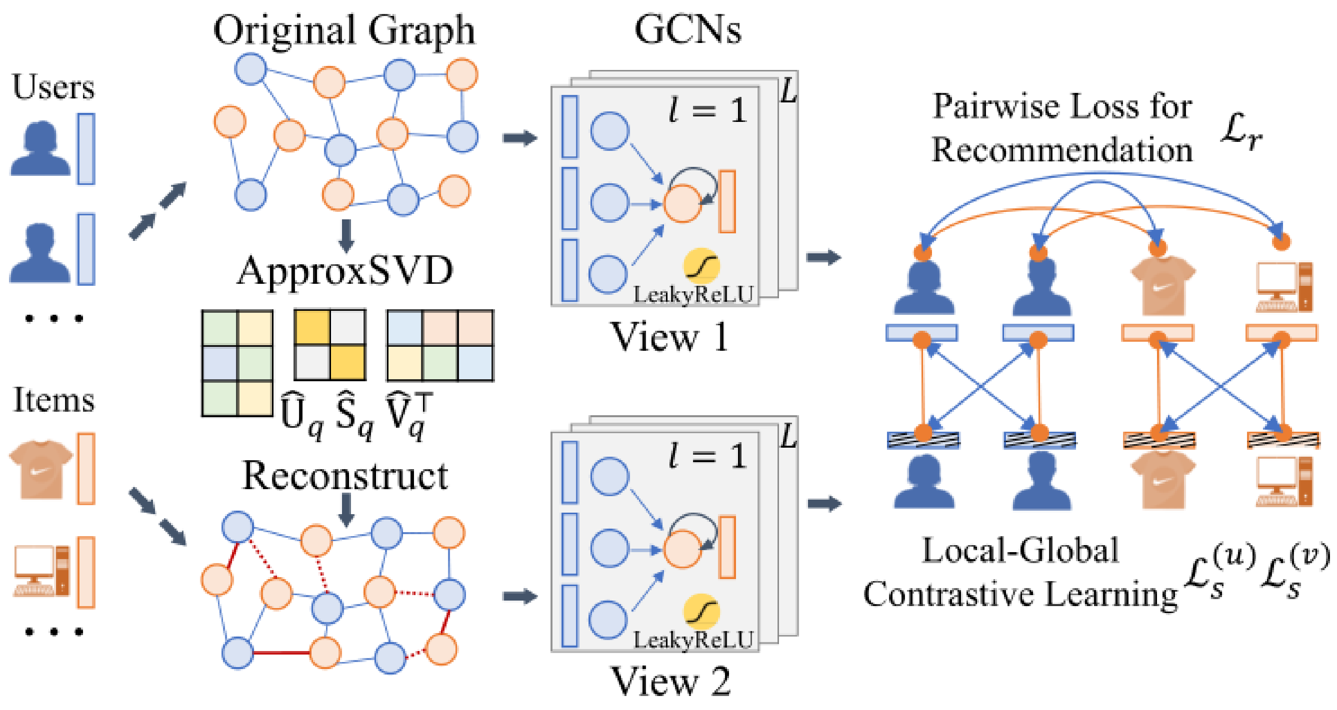

Figure 1: Overall structure of LightGCL.

$$z_{i,l}^{(u)} = \sigma(p(\tilde{\mathcal{A}}_{i,:}) \cdot \boldsymbol{E}_{l-1}^{(v)}), \quad z_{j,l}^{(v)} = \sigma(p(\tilde{\mathcal{A}}_{:,j}) \cdot \boldsymbol{E}_{l-1}^{(u)}) \tag{1}$$

$$\boldsymbol{e}_{i,l}^{(u)} = \boldsymbol{z}_{i,l}^{(u)} + \boldsymbol{e}_{i,l-1}^{(u)}, \quad \boldsymbol{e}_{j,l}^{(v)} = \boldsymbol{z}_{j,l}^{(v)} + \boldsymbol{e}_{j,l-1}^{(v)} \tag{2}$$

$$\boldsymbol{e}_i^{(u)} = \sum_{l=0}^{L} \boldsymbol{e}_{i,l}^{(u)}, \ \boldsymbol{e}_j^{(v)} = \sum_{l=0}^{L} \boldsymbol{e}_{j,l}^{(v)}, \quad \hat{y}_{i,j} = \boldsymbol{e}_i^{(u)\top} \boldsymbol{e}_j^{(v)} \tag{3}$$

Approach

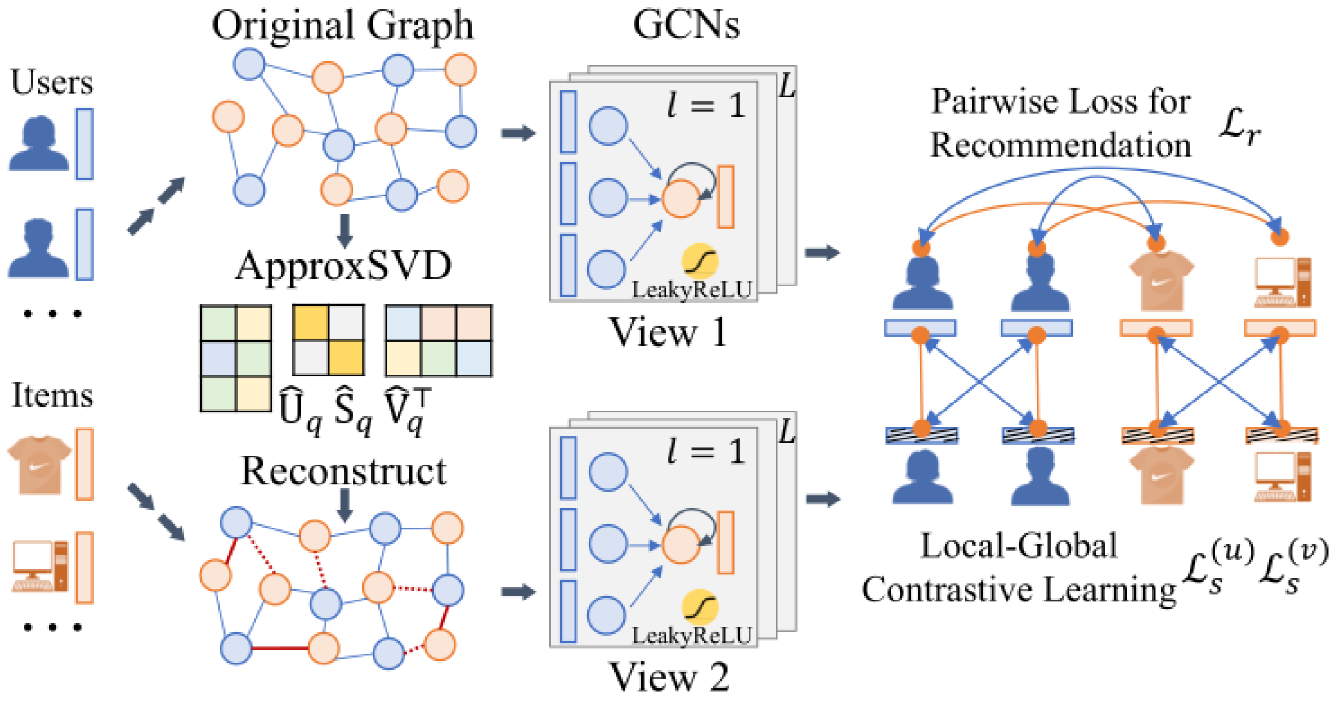

Figure 1: Overall structure of LightGCL.

$$\hat{\mathcal{A}} = \boldsymbol{U}_q \boldsymbol{S}_q \boldsymbol{V}_q^\top$$

$$\boldsymbol{g}_{i,l}^{(u)} = \sigma(\hat{\mathcal{A}}_{i,:} \cdot \boldsymbol{E}_{l-1}^{(v)}), \quad \boldsymbol{g}_{j,l}^{(v)} = \sigma(\hat{\mathcal{A}}_{:,j} \cdot \boldsymbol{E}_{l-1}^{(u)}) \tag{4}$$

$$\hat{\boldsymbol{U}}_q, \hat{\boldsymbol{S}}_q, \hat{\boldsymbol{V}}_q^\top = \mathrm{ApproxSVD}(\mathcal{A}, q), \quad \hat{\mathcal{A}}_{SVD} = \hat{\boldsymbol{U}}_q \hat{\boldsymbol{S}}_q \hat{\boldsymbol{V}}_q^\top \tag{5}$$

$$\boldsymbol{G}_l^{(u)} = \sigma(\hat{\mathcal{A}}_{SVD} \boldsymbol{E}_{l-1}^{(v)}) = \sigma(\hat{\boldsymbol{U}}_q \hat{\boldsymbol{S}}_q \hat{\boldsymbol{V}}_q^\top \boldsymbol{E}_{l-1}^{(v)}); \quad \boldsymbol{G}_l^{(v)} = \sigma(\hat{\mathcal{A}}_{SVD}^\top \boldsymbol{E}_{l-1}^{(u)}) = \sigma(\hat{\boldsymbol{V}}_q \hat{\boldsymbol{S}}_q \hat{\boldsymbol{U}}_q^\top \boldsymbol{E}_{l-1}^{(u)}) \tag{6}$$

Approach

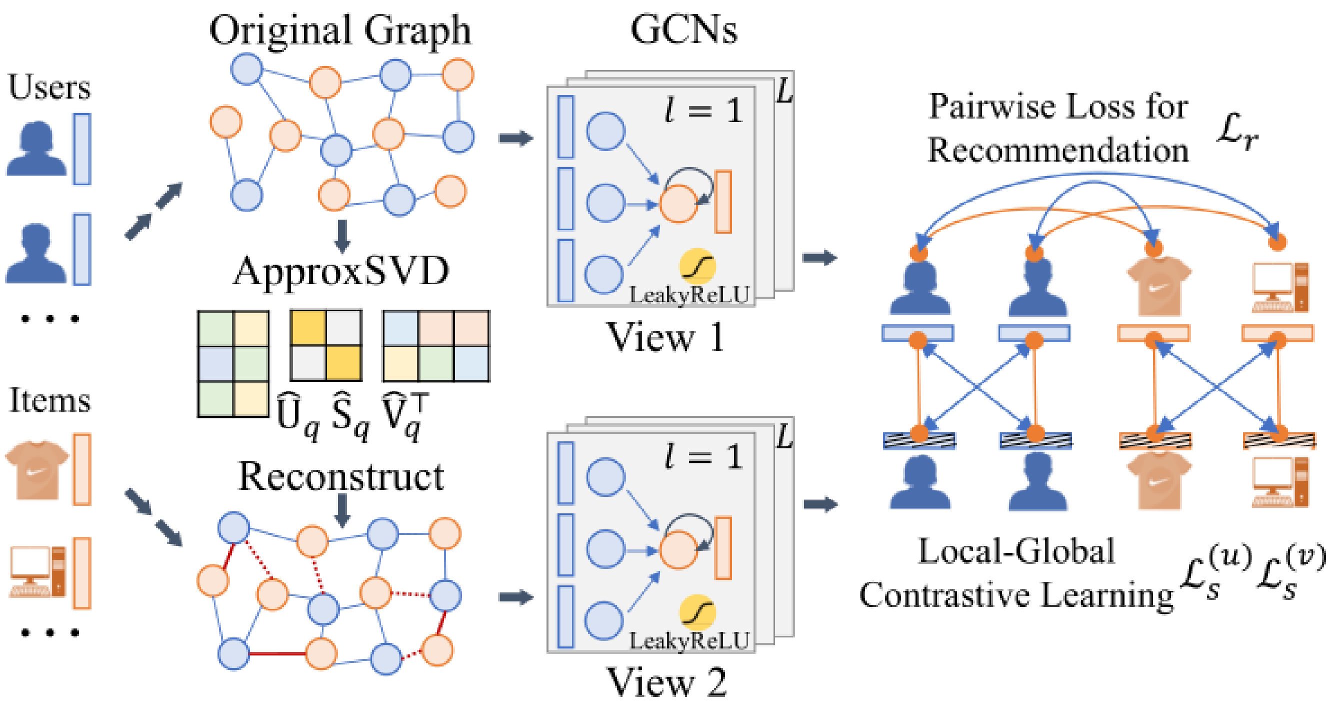

Figure 1: Overall structure of LightGCL.

$$\mathcal{L}_s^{(u)} = \sum_{i=0}^{I} \sum_{l=0}^{L} -\log \frac{\exp(s(\boldsymbol{z}_{i,l}^{(u)}, \boldsymbol{g}_{i,l}^{(u)}/\tau))}{\sum_{i'=0}^{I} \exp(s(\boldsymbol{z}_{i,l}^{(u)}, \boldsymbol{g}_{i',l}^{(u)})/\tau)} \quad (7)$$

$$\mathcal{L} = \mathcal{L}_r + \lambda_1 \cdot (\mathcal{L}_s^{(u)} + \mathcal{L}_s^{(v)}) + \lambda_2 \cdot \|\Theta\|_2^2;$$

$$\mathcal{L}_r = \sum_{i=0}^{I} \sum_{s=1}^{S} \max(0, 1 - \hat{y}_{i,p_s} + \hat{y}_{i,n_s}) \quad (8)$$

Experiment

Table 1: Performance comparison with baselines on five datasets.

Data	Metric	DGCF	HyRec	LightGCN	MHCN	SGL	SimGRACE	GCA	HCCF	SHT	SimGCL	LightGCL	*p-val.*	*impr.*
Yelp	R@20	0.0466	0.0472	0.0482	0.0503	0.0526	0.0603	0.0621	0.0626	0.0651	0.0718	**0.0793**	7e-9	10%
	N@20	0.0395	0.0395	0.0409	0.0424	0.0444	0.0435	0.0530	0.0527	0.0546	0.0615	**0.0668**	8e-9	8%
	R@40	0.0774	0.0791	0.0803	0.0826	0.0869	0.0989	0.1021	0.1040	0.1091	0.1166	**0.1292**	2e-9	10%
	N@40	0.0511	0.0522	0.0527	0.0544	0.0571	0.0656	0.0677	0.0681	0.0709	0.0778	**0.0852**	2e-9	9%
Gowalla	R@20	0.0944	0.0901	0.0985	0.0955	0.1030	0.0869	0.0896	0.1070	0.1232	0.1357	**0.1578**	1e-6	16%
	N@20	0.0522	0.0498	0.0593	0.0574	0.0623	0.0528	0.0537	0.0644	0.0731	0.0818	**0.0935**	2e-6	14%
	R@40	0.1401	0.1356	0.1431	0.1393	0.1500	0.1276	0.1322	0.1535	0.1804	0.1956	**0.2245**	3e-6	14%
	N@40	0.0671	0.0660	0.0710	0.0689	0.0746	0.0637	0.0651	0.0767	0.0881	0.0975	**0.1108**	3e-6	13%
ML-10M	R@20	0.1763	0.1801	0.1789	0.1497	0.1833	0.2254	0.2145	0.2219	0.2173	0.2265	**0.2613**	1e-9	15%
	N@20	0.2101	0.2178	0.2128	0.1814	0.2205	0.2686	0.2613	0.2629	0.2573	0.2613	**0.3106**	3e-9	18%
	R@40	0.2681	0.2685	0.2650	0.2250	0.2768	0.3295	0.3231	0.3265	0.3211	0.3345	**0.3799**	7e-10	13%
	N@40	0.2340	0.2340	0.2322	0.1962	0.2426	0.2939	0.2871	0.2880	0.3318	0.2880	**0.3387**	1e-9	17%
Amazon	R@20	0.0211	0.0302	0.0319	0.0296	0.0327	0.0381	0.0309	0.0322	0.0441	0.0474	**0.0585**	2e-7	23%
	N@20	0.0154	0.0225	0.0236	0.0219	0.0249	0.0291	0.0238	0.0247	0.0328	0.0360	**0.0436**	2e-6	21%
	R@40	0.0351	0.0432	0.0499	0.0489	0.0531	0.0621	0.0498	0.0525	0.0719	0.0750	**0.0933**	1e-7	24%
	N@40	0.0201	0.0246	0.0290	0.0284	0.0312	0.0371	0.0301	0.0314	0.0420	0.0451	**0.0551**	9e-7	22%
Tmall	R@20	0.0235	0.0233	0.0225	0.0203	0.0268	0.0222	0.0373	0.0314	0.0387	0.0473	**0.0528**	3e-5	11%
	N@20	0.0163	0.0160	0.0154	0.0139	0.0183	0.0152	0.0252	0.0213	0.0262	0.0328	**0.0361**	1e-4	10%
	R@40	0.0394	0.0350	0.0378	0.0340	0.0446	0.0367	0.0616	0.0519	0.0645	0.0766	**0.0852**	1e-5	11%
	N@40	0.0218	0.0199	0.0208	0.0188	0.0246	0.0203	0.0337	0.0284	0.0352	0.0429	**0.0473**	7e-5	10%

Experiment

Table 2: Comparisons of computational complexity against baselines.

Stage	Computation	LightGCN	SGL	SimGCL	LightGCL
Pre-processing	Normalization	$O(E)$	$O(E)$	$O(E)$	$O(E)$
	SVD	–	–	–	$O(qE)$
Training	Augmentation	–	$O(2\rho E)$	–	–
	Graph Convolution	$O(2ELd)$	$O(2ELd + 4\rho ELd)$	$O(6ELd)$	$O[2ELd + 2q(I+J)Ld]$
	BPR Loss	$O(2Bd)$	$O(2Bd)$	$O(2Bd)$	$O(2Bd)$
	InfoNCE Loss	–	$O(Bd + BMd)$	$O(Bd + BMd)$	$O[(Bd + BMd)L]$

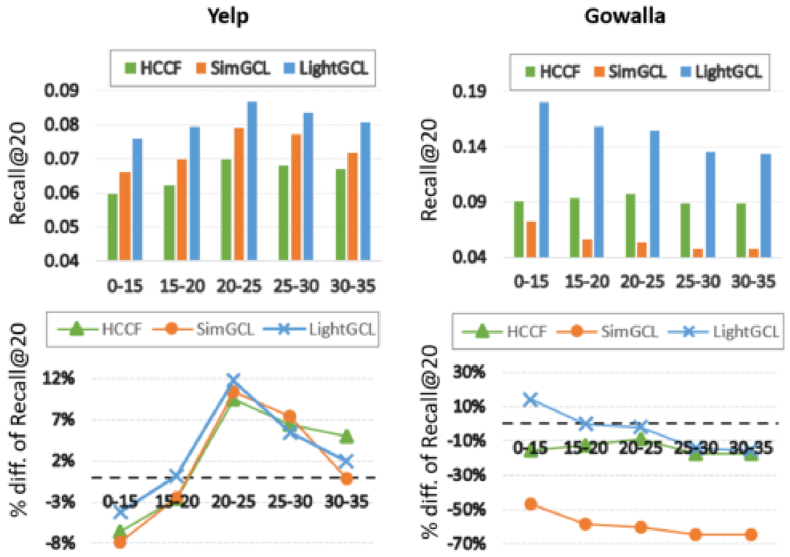

Figure 2: Performance on users of different sparsity degrees, in terms of *Recall* (histograms) and relative *Recall w.r.t* overall performances (charts).

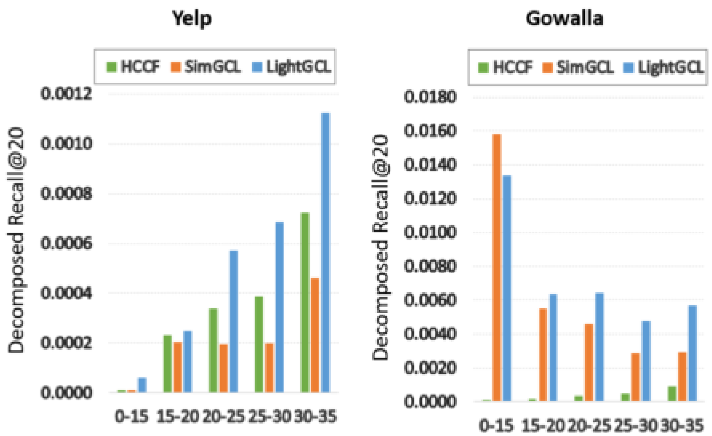

Figure 3: LightGCL's ability to alleviate popularity bias in comparison to SOTA CL-based methods HCCF and SimGCL.

Experiment

Table 3: Mean Average Distance (MAD) of the embeddings learned by different methods.

Dataset	MHCN	LightGCN	LightGCL	SGL	SimGCL
Yelp	0.8806	0.9469	0.9657	0.9962	0.9956
Gowalla	0.9247	0.9568	0.9721	0.9859	0.9897

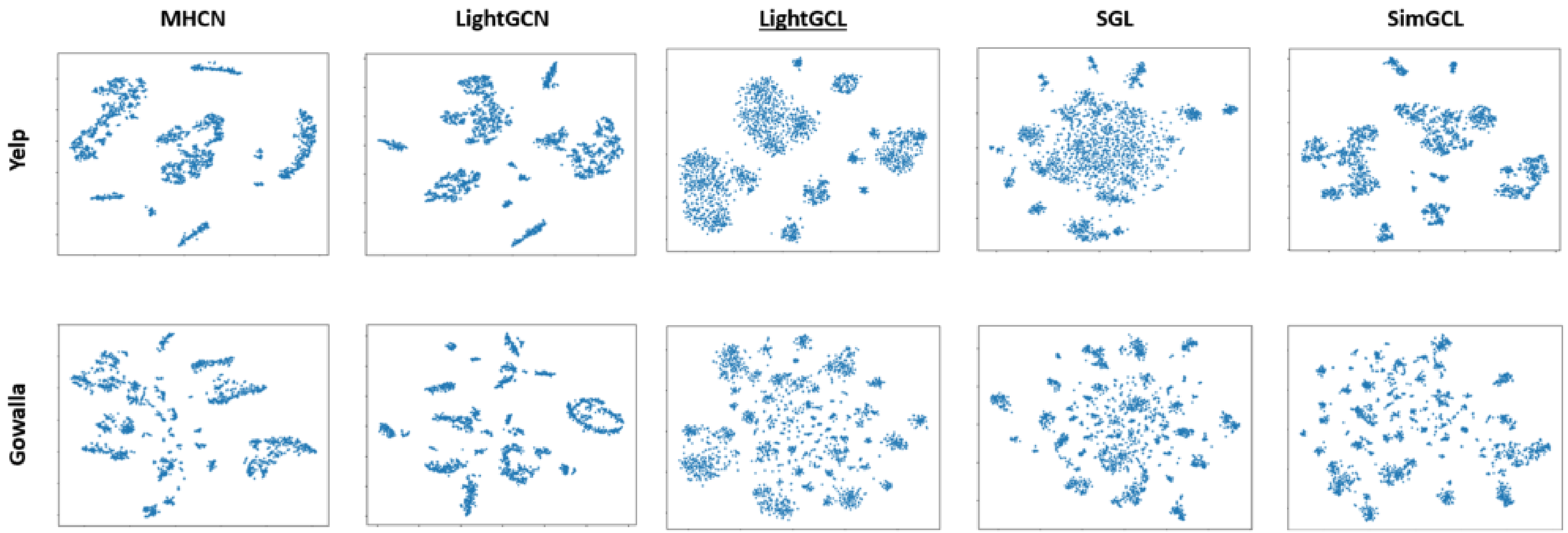

Figure 4: Embedding distributions on *Yelp* and *Gowalla* visualized with t-SNE.

Experiment

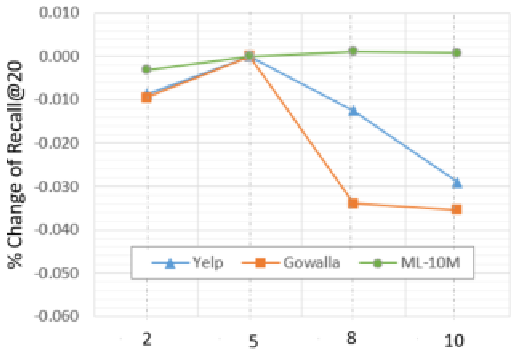

Figure 5: *Recall* change *w.r.t.* q.

Table 4: Ablation study on LightGCL.

Variant	Yelp		Gowalla	
	Recall@20	NDCG@20	Recall@20	NDCG@20
CL-MF	0.0781	0.0659	0.1561	0.0929
CL-SVD++	0.0788	0.0666	0.1568	0.0932
LightGCL	**0.0793**	**0.0668**	**0.1578**	**0.0935**

Experiment

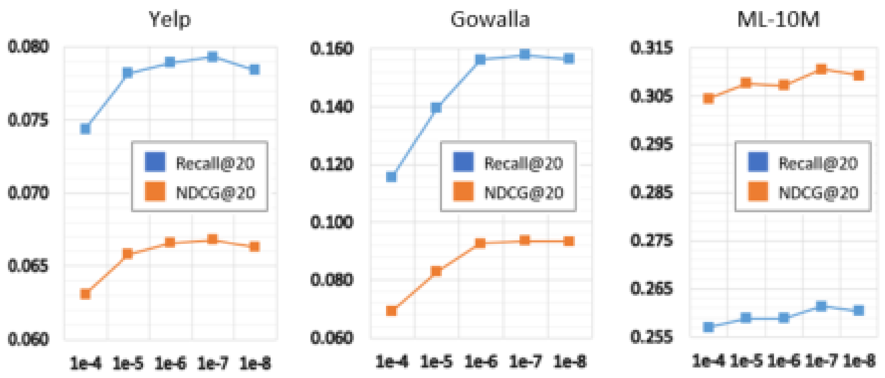

Figure 6: Impact of λ_1.

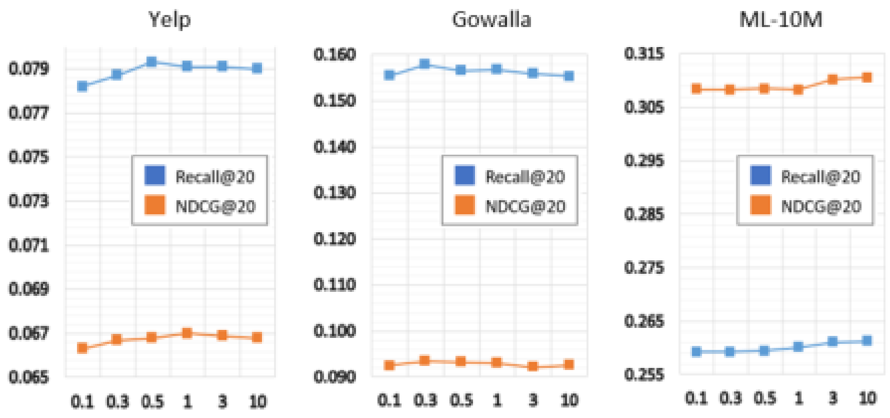

Figure 7: Impact of τ

Thank you!