How Does Message Passing Improve Collaborative Filtering?

Clark Mingxuan Ju[1,2], William Shiao[3], Zhichun Guo[2], Yanfang Ye[2],
Yozen Liu[1], Neil Shah[1], Tong Zhao[1]

[1]Snap Inc. [2]Universtiy of Notre Dame, [3]UCR

NEURAL INFORMATION PROCESSING SYSTEMS

Code: Paper:

1. Neighbor information is more helpful than accompanying gradients from message passing

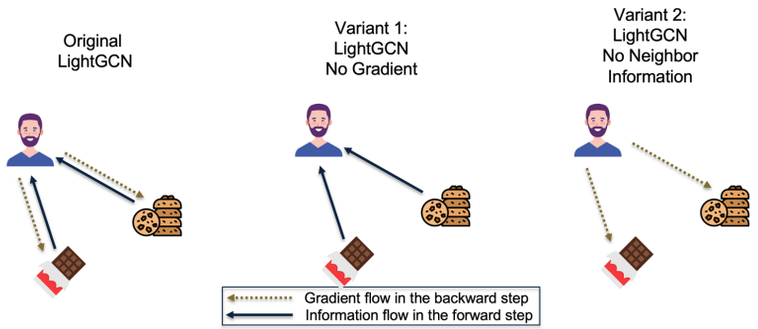
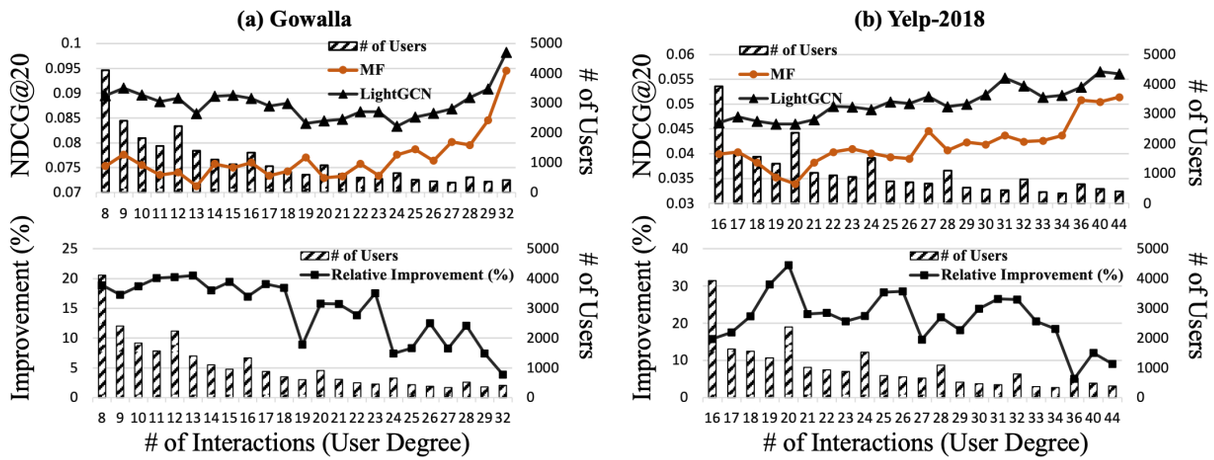


Table 1: Performance of LightGCN variants.

Method	Yelp-2018	Gowalla	Amazon-book
NDCG@20			
LightGCN	6.36	9.88	8.13
w/o grad.	6.16 (3.1%↓)	9.87 (0.1%↓)	7.80 (4.1%↓)
w/o neigh. info	4.71 (25.9%↓)	6.95 (29.7%↓)	6.95 (14.5%↓)
w/o both	6.09 (4.2%↓)	9.83 (0.5%↓)	7.75 (4.7%↓)
Recall@20			
LightGCN	11.21	18.53	12.97
w/o grad.	10.87 (3.0%↓)	18.51 (0.1%↓)	12.81 (1.2%↓)
w/o neigh. info	8.44 (24.7%↓)	13.06 (29.5%↓)	11.25 (13.3%↓)
w/o both	10.71 (4.5%↓)	18.42 (0.6%↓)	12.57 (3.1%↓)

We can skip message passing during training, yet achieving good enough performance.

2. Message Passing Helps Low-degree Users More Compared with High-degrees



We can skip message passing on high-degree users (very expensive)!

Seeking Internship?
We are hiring interns for 2025 to work on RecSys!

Apply here:

mju@snap.com

3. Proposed Method: TAG-CF -- Test-time Aggregation for CF

Mix-up user and item embeddings during the testing only.

$$\mathbf{u}_i^* = \mathbf{u}_i + \sum_{i_j \in N(u_i)} |N(u_i)|^m |N(i_j)|^n \cdot \mathbf{i}_j, \mathbf{i}_i^* = \mathbf{i}_i + \sum_{u_j \in N(i_i)} |N(i_i)|^m |N(u_j)|^n \cdot \mathbf{u}_j$$