

A Training procedures for GNN models

The experiment is conducted for 200 epochs on the baseline. No batch processing is applied on Cora, Citeseer and Pubmed. The Degree-Sorted Residual Mamba pathway consists of three Degree-Sorted Residual Mamba blocks.

- Early stopping: no early stopping. Full-batch training on Cora, Citeseer and Pubmed. Batch size is set 512 on the other datasets.
- Maximum number of epochs: 200.
- Coauthor and Amazon split:
Train set ratio: 0.6; validation set ratio: 0.2; test set ratio: 0.2.
Cora, Citeseer and Pubmed split: default (as in the Planetoid split)

B Parameters about Replacing Different GCN Variants in ResGCN Branches

The experiment is conducted for 50 epochs on the Pubmed dataset with a batch size of 512.

C Parameters about Ablation for Degree-Sorted Residual Mamba pathway

The experiment is conducted for 50 epochs on the Pubmed dataset with a batch size of 512.

D Parameters about Ablation experiment for ResGCN pathway

The experiment is conducted for 50 epochs on the Pubmed dataset with a batch size of 512.

E Parameters about Ablation experiment for reordering mechanism

The experiment is conducted for 50 epochs on the Pubmed dataset with a batch size of 512.

F Datasets Description and Statistics

Amazon Computers and Amazon Photo are two subsets of the Amazon co-purchase graph originally introduced by McAuley et al. (2015), and further processed and split by Shchur et al. (2018)[19]. In these graphs, nodes represent items and edges indicate that two items are often purchased together. The node features come from the bag-of-words representation of product reviews, while the category label is the category to which the product belongs.

Coauthor CS and Coauthor Physics graphs are from the Microsoft Academic Graph for the 2016 KDD Cup. The nodes in the graph are authors, and if two authors have published a paper together, there is an edge between them. Node features are based on the keywords of each author’s published papers, while the tags represent the author’s most active research areas. These four datasets are downloaded from the GitHub website¹.

Dataset	Nodes	Edges	Node Features	number of Classees	Type
Cora	2708	5429	1433	7	Citation
Citeseer	3327	4732	3703	6	Citation
PubMed	19717	44338	500	3	Citation
Reddit	232965	11606919	602	41	Social
Coauthor CS	18333	81894	6805	15	Coauthor
Coauthor Physics	34493	247962	8415	5	Coauthor
Amazon Computers	13381	245778	767	10	Co-purchase
Amazon Photo	7487	119043	745	8	Co-purchase

Table 6: Features of Test Graph Datasets

¹ <https://github.com/shchur/gnn-benchmark?tab=readme-ov-filedatasets>

G experiment results with 3-7 layers

In addition to the experimental results Table 2, there are also results of testing with other layers. The figure below Figure 5 includes the experimental results of GAT and MambaPlusResGCN with layers 3 to 7 on three datasets. The best accuracy of GAT is 80.37%, and the best result of MambaPlusResGCN is 77.83% on Cora.

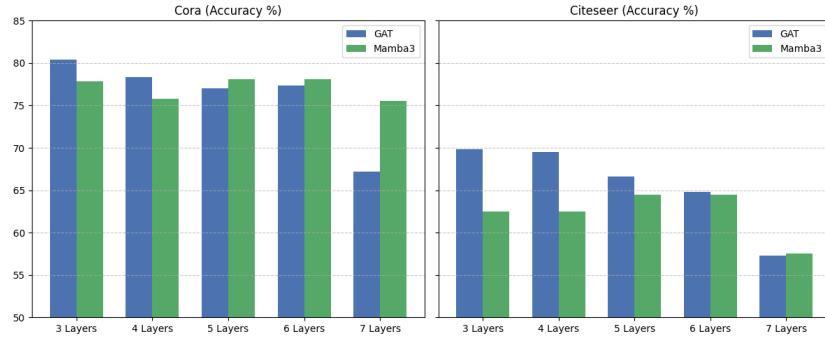


Fig. 5: experimental results of GAT and MambaPlusResGCN with layers 3 to 7

H t-SNE Graph

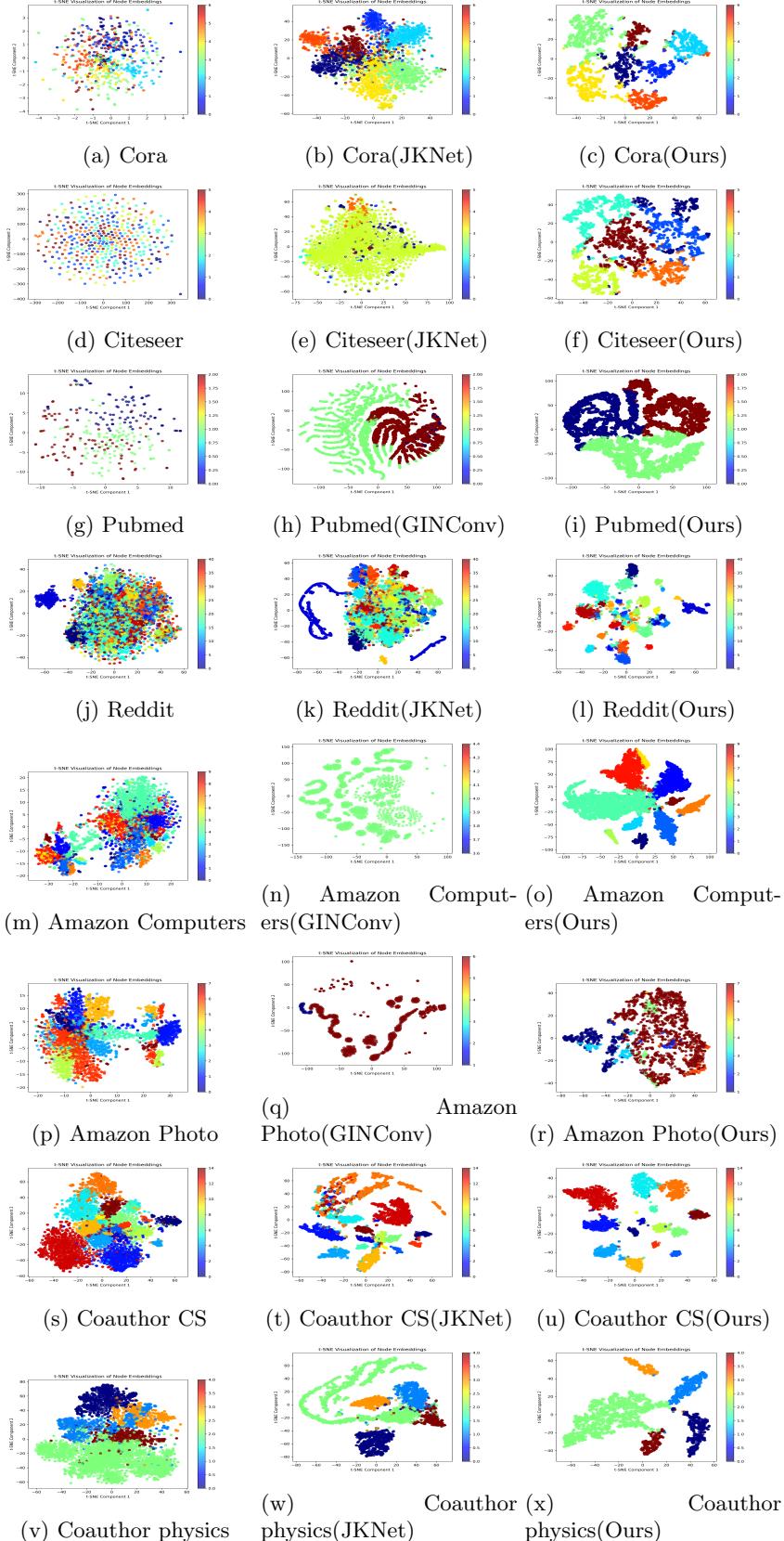


Fig. 6: t-SNE graph of embeddings produced by 32-layer GCN models and MambaPlusResGCN on GNN benchmarks