Decoupling the Depth and Scope of Graph Neural Networks

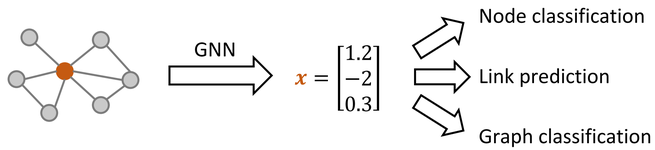
https://github.com/facebookresearch/shadow_GNN

Hanqing Zeng, Muhan Zhang, Yinglong Xia, Ajitesh Srivastava, Andrey Malevich, Rajgopal Kannan, Viktor Prasanna, Long Jin, Ren Chen

NEURAL INFORMATION PROCESSING SYSTEMS

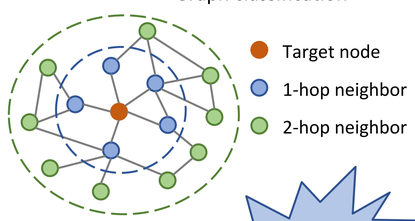
Background & Challenges

Graph representation learning & GNNs



Message passing in GNNs

- Scope: from what neighbors?
- Depth: by how many layers?



GNN designs by default

Increasing the **depth** (L-layer GNN) ⟺ Expanding the **scope** (L-hop neighborhood)

Neighborhood explosion $\sim d^L$

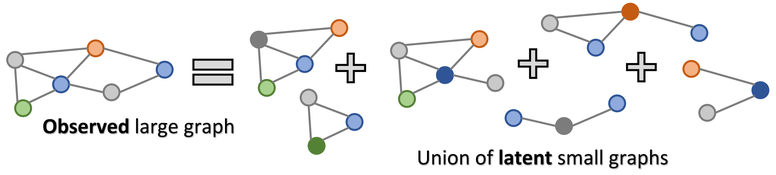
Decoupling Design Principle

Intuitions

- Some neighbors are irrelevant → No need to pass their messages
- Some neighbors are important → Worth passing their messages many times

Scope, Depth } is a property of the { Data, Model

Alternative view on the input graph



Algorithm: Generate embedding for a target node v of the full graph $\mathcal{G}$

```
1. Extract a subgraph G_[v] around v
2. for round i = 1 to L':
       Perform message passing along all edges in G_[v]
3.Take v's embedding from all node embeddings of G_[v]
```

Theoretical Justifications

Decoupling improves GNN scalability

- Deep model + Large graph ≠ Exploding scope
- With fixed-size scope, complexity is linear with the model depth

Decoupling improves GNN expressivity

Graph signal processing perspective

Decoupled-GCN addressed oversmoothing of original GCN

- ∞ layers: different scope of different target nodes → distinctive embeddings

Function approximation perspective

Decoupled-SAGE is more expressive than GraphSAGE

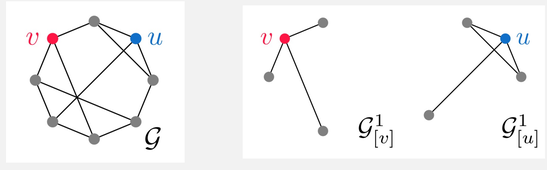
- Consider neighborhood $\mathcal{G}'$ & function τ for linear combination of $\mathcal{G}'$ features
- Decoupled-SAGE can approximate τ where

Scope $\mathcal{G}_{[v]} = \mathcal{G}'$ — **Depth** reduces the error exponentially

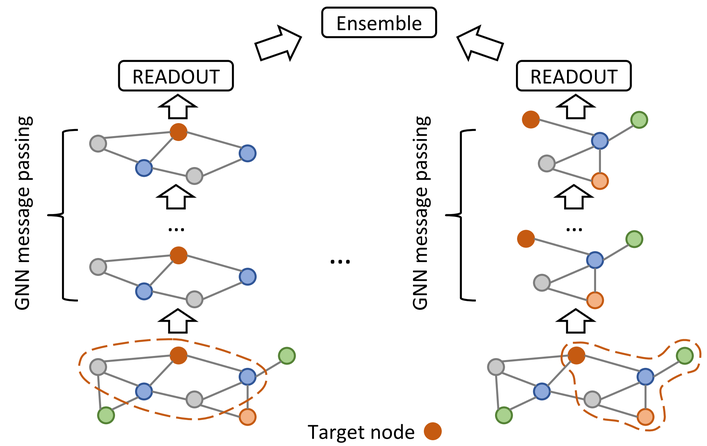
Topology learning perspective

Decoupled-GIN is more expressive than original GIN / 1-WL

- Non-isomorphic regular graphs: *subgraphs of a regular graph is not regular*



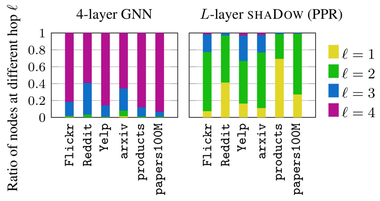
Architecture: shaDow-GNN



Evaluation

- Decoupling improves accuracy & reduces cost by orders of magnitude
- Subgraph EXTRACT functions are critical
- Scalable to production-scale graphs with low-end server configurations

Neighborhood composition



Scaling to 100M-node graph

Table 2: Leaderboard comparison on `papers100M`

Method	Test accuracy	Val accuracy	Neigh size
GraphSAGE+incep	0.6707±0.0017	0.7032±0.0011	4E5
SIGN-XL	0.6606±0.0019	0.6984±0.0006	> 4E5
SGC	0.6329±0.0019	0.6648±0.0020	> 4E5
SHADOW-GAT$_{200}$	0.6681±0.0016	0.7019±0.0011	2E2
SHADOW-GAT$_{400}$	**0.6710**±0.0015	**0.7067**±0.0012	3E2

Baseline comparisons

Method	Layers	Flickr		Reddit		Yelp		ogbn-arxiv		ogbn-products	
		Accuracy	Cost	Accuracy	Cost	F1-micro	Cost	Accuracy	Cost	Accuracy	Cost
GCN	3	0.5159±0.0017	2E0	0.9532±0.0003	6E1	0.4028±0.0019	2E1	0.7170±0.0026	1E1	0.7567±0.0018	5E0
	5	0.5217±0.0016	2E2	0.9495±0.0012	1E3	OOM	1E3	0.7186±0.0017	1E3	OOM	9E2
GCN	3	0.5155±0.0027	2E0	0.9523±0.0003	6E1	0.5110±0.0012	2E1	0.7093±0.0003	1E1	**0.8003**±0.0024	5E0
+ GraphSAINT-RW	5	0.5165±0.0026	2E2	0.9523±0.0012	1E3	0.5012±0.0021	1E3	0.7039±0.0020	1E3	0.7992±0.0021	9E2
SHADOW-GCN	3	0.5262±0.0018	**(1)**	0.9581±0.0004	**(1)**	0.5255±0.0012	**(1)**	0.7192±0.0025	**(1)**	0.7778±0.0030	**(1)**
+PPR	5	**0.5270**±0.0024	1E0	**0.9583**±0.0002	1E0	**0.5272**±0.0018	2E0	**0.7207**±0.0030	2E0	0.7844±0.0029	2E0
GraphSAGE	3	0.5140±0.0014	3E0	0.9653±0.0002	5E1	0.6178±0.0033	2E1	0.7192±0.0027	1E1	0.7858±0.0013	4E0
	5	0.5154±0.0052	2E2	0.9626±0.0004	1E3	OOM	2E3	0.7193±0.0037	1E3	OOM	1E3
GraphSAGE	3	0.5176±0.0032	3E0	0.9671±0.0003	5E1	0.6453±0.0011	2E1	0.7107±0.0003	1E1	0.7923±0.0023	4E0
+ GraphSAINT-RW	5	0.5201±0.0032	2E2	0.9670±0.0010	1E3	0.6394±0.0003	2E3	0.7013±0.0021	1E3	0.7964±0.0022	1E3
SHADOW-SAGE	3	0.5288±0.0014	1E0	0.9660±0.0003	1E0	0.6493±0.0001	1E0	0.7163±0.0012	1E0	0.7993±0.0012	1E0
+ 2-hop	5	0.5338±0.0038	2E0	0.9661±0.0002	2E0	0.6503±0.0001	2E0	0.7183±0.0012	2E0	0.8014±0.0020	2E0
SHADOW-SAGE	3	0.5344±0.0028	**(1)**	**0.9693**±0.0002	**(1)**	**0.6512**±0.0002	**(1)**	0.7234±0.0032	**(1)**	0.7945±0.0021	**(1)**
+ PPR	5	**0.5424**±0.0025	2E0	0.9691±0.0003	2E0	0.6502±0.0001	2E0	**0.7255**±0.0013	2E0	**0.8043**±0.0026	2E0
GAT	3	0.5070±0.0032	2E1	OOM	3E2	OOM	2E2	0.7201±0.0011	1E2	OOM	3E1
	5	0.5164±0.0033	2E2	OOM	3E3	OOM	2E3	OOM	3E3	OOM	2E3
GAT	3	0.5225±0.0053	2E1	0.9671±0.0003	3E2	0.6459±0.0002	2E2	0.6977±0.0003	1E2	0.8027±0.0028	3E1
+ GraphSAINT-RW	5	0.5153±0.0034	2E2	0.9651±0.0024	2E3	0.6478±0.0012	2E3	0.6954±0.0013	3E3	0.7990±0.0072	2E3
SHADOW-GAT	3	**0.5383**±0.0032	**(1)**	0.9703±0.0010	**(1)**	**0.6549**±0.0002	**(1)**	0.7243±0.0011	**(1)**	0.8014±0.0012	**(1)**
+ PPR	5	0.5342±0.0023	2E0	**0.9710**±0.0008	2E0	0.6537±0.0004	2E0	**0.7283**±0.0027	2E0	**0.8094**±0.0034	2E0