A Topology-aware Graph Coarsening Framework for Continual Graph Learning

Xiaoxue Han, Zhuo Feng, Yue Ning

Stevens Institute of Technology

Catastrophic forgetting on GNN

- Graph Neural Networks (GNNs) are oblivious
- Fails to remember pre-existing knowledge in a **Continual Learning** Setting:
  - The model learns a sequence of tasks incrementally
  - No access to the data from previous task
- **Challenge:** Preservation of knowledge from old tasks when learning the new ones

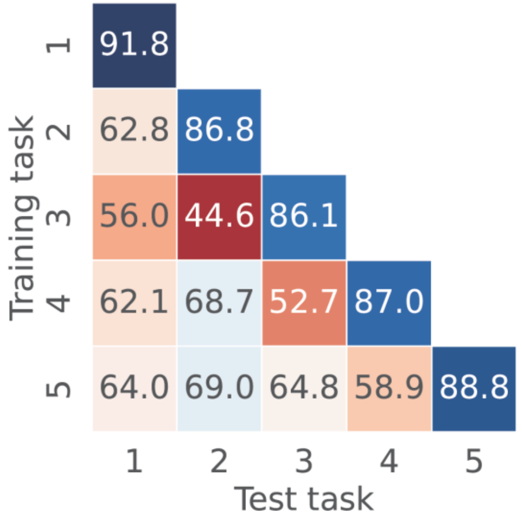


Fig. 1. The F1 score on the test task (x-axis) after GCN learns the training tasks (y-axis) on the Amazon-kindle dataset

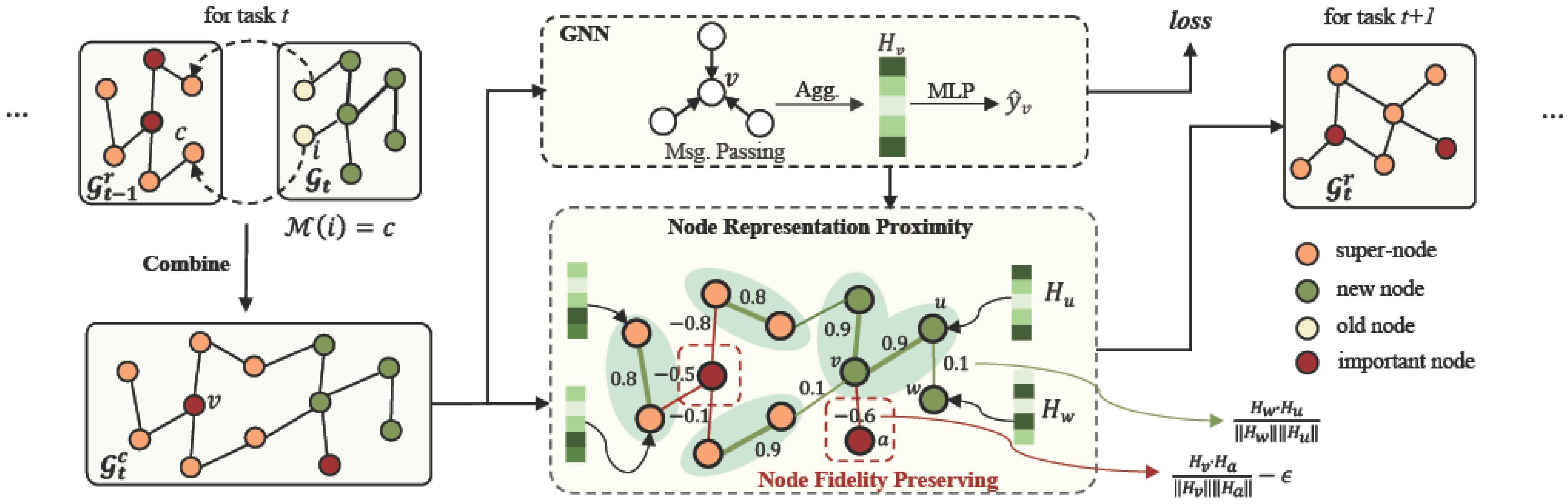
Experience Replay

- Use a small memory buffer to save sampled data from previous task
- Relay the saved samples to the model when training new tasks
- **Straight-forward and efficient**

Challenges:

- Preservation of graph topological properties from previous tasks with small memory
- Capturing the inter-dependencies among tasks

TACO:
A Topology-**a**ware Graph **Co**arsening Framework



Evaluation

	Kindle				DBLP				ACM			
Method	F1-AP(%) ↑		F1-AF (%) ↓		F1-AP (%) ↑		F1-AF (%) ↓		F1-AP (%) ↑		F1-AF (%) ↓	
joint train	87.21	± 0.55	0.45	± 0.25	86.33	± 1.38	0.77	± 0.13	75.35	± 1.49	1.87	± 0.60
finetune	69.10	± 10.85	18.99	± 11.19	67.85	± 8.05	20.43	± 7.07	60.53	± 9.35	19.09	± 9.23
simple-reg	68.80	± 10.02	18.21	± 10.49	69.70	± 9.16	18.69	± 8.48	61.63	± 10.09	17.83	± 9.99
EWC	77.08	± 8.37	10.87	± 8.62	79.38	± 4.86	8.85	± 4.11	66.48	± 6.43	12.73	± 6.26
TWP	78.90	± 4.71	8.99	± 4.93	80.05	± 3.71	8.23	± 3.28	65.98	± 7.26	13.33	± 6.94
OTG	69.01	± 10.55	18.94	± 10.79	68.24	± 10.12	20.12	± 9.34	61.45	± 9.94	18.33	± 9.86
GEM	76.08	± 6.70	11.01	± 7.27	80.04	± 3.24	7.90	± 2.68	67.17	± 4.24	11.69	± 3.94
ERGNN-rs	77.63	± 3.61	9.64	± 4.19	78.02	± 5.79	10.08	± 5.16	64.82	± 7.89	14.43	± 7.68
ERGNN-rb	75.87	± 6.41	11.46	± 6.98	75.16	± 7.24	12.85	± 6.54	63.58	± 8.82	15.66	± 8.71
ERGNN-mf	77.28	± 5.91	10.15	± 6.31	77.42	± 5.25	10.64	± 4.38	64.80	± 8.49	14.59	± 8.41
DyGrain	69.14	± 10.47	18.88	± 10.72	67.52	± 10.88	20.83	± 10.16	61.40	± 9.57	18.47	± 9.50
IncreGNN	69.45	± 10.34	18.48	± 10.66	69.40	± 9.60	18.92	± 8.75	61.32	± 9.70	18.42	± 9.64
SSM	78.99	± 3.13	8.19	± 3.63	82.71	± 1.76	4.20	± 1.26	68.77	± 2.93	9.50	± 2.47
SSRM	77.37	± 4.06	9.99	± 4.55	77.43	± 5.34	10.66	± 4.47	64.39	± 7.43	14.72	± 7.48
CaT	75.12	±4.01	11.83	±4.22	76.24	±3.78	9.06	±3.14	63.72	±2.21	11.86	±2.32
DeLoMe	76.93	±3.83	10.16	±4.68	77.27	±2.85	8.01	±2.16	64.54	±2.42	10.75	±2.04
TACO	**82.97**	**± 2.05**	**4.91**	**± 1.90**	**84.60**	**± 2.01**	**2.51**	**± 1.03**	**70.96**	**± 2.68**	**8.02**	**± 2.33**
p-value	<0.0001		<0.0001		0.002		<0.0001		0.005		0.02	

Average performance:

$$AP = \frac{1}{T}\sum_{j=1}^{T} a_{T,j}$$

Average forgetting:

$$AF = \frac{1}{T}\sum_{j=1}^{T} \max_{l=\{1,\ldots\ldots,T\}} a_{l,j} - a_{T,j}$$

Contributions

- A dynamic graph coarsening framework to effectively perverse graph topology information.
- An efficient graph coarsening method.
- Extensive evaluations on real-world datasets.

More information

Please check out the paper for more details of our methods, results, and ablation studies. Also, feel free to play with the code.