Tackling Over-squashing in Graph Neural Networks via Higher-order Neighborhood Disentanglement

Yu Wang  Tyler Derr 

Network and Data Science Lab, Vanderbilt University

VANDERBILT UNIVERSITY

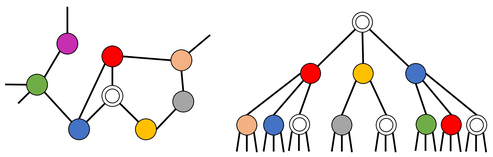
NDS Lab Network and Data Science

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant ISS-2011666.

Motivation

Over-smoothing

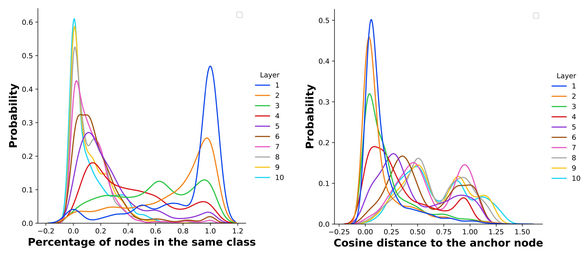


Node features are fused together to become indistinguishable when stacking multiple graph convolution layers.

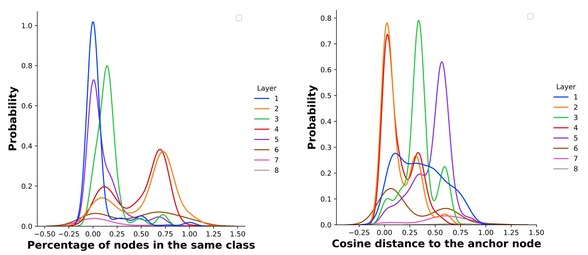
Over-squashing

Information from the exponentially-growing receptive field is compressed into fixed-length node vectors.

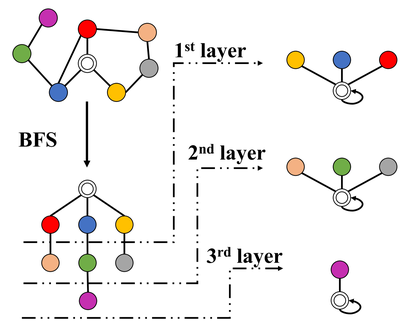
Cora



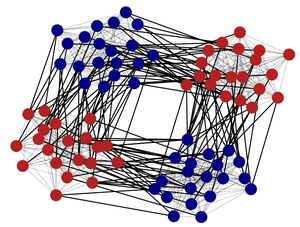
Texas

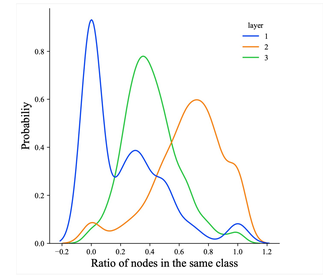


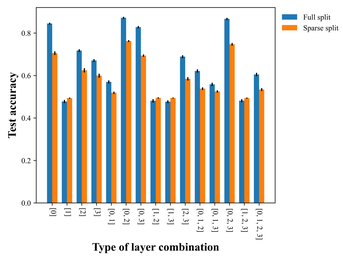
Tree Decomposition and Synthetic Evaluation



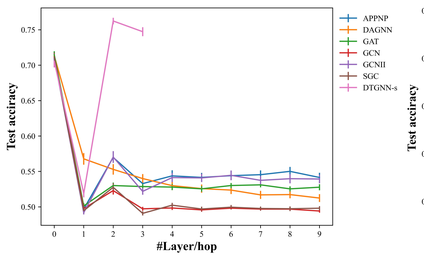
Contextual Stochastic Block Model



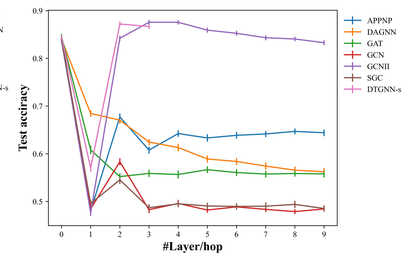




Sparse split (5%/5%/90%)



Dense split (60%/20%/20%)



Decomposition Tree of Graph Neural Network

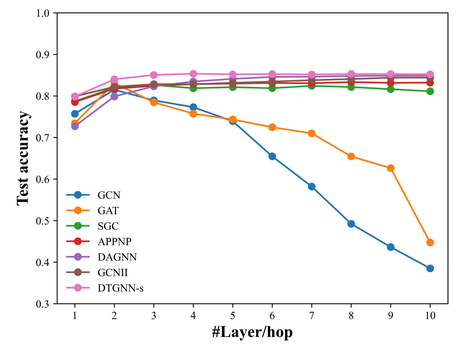
Real World Graph Result

Model	Cora	Cite.	Pub.	Corn.	Tex.	Wisc.	Avg. Rank
MLP	74.75	72.41	86.65	80.97	81.32	85.38	5.83
GCN	86.97	76.37	88.19	58.57	58.68	53.14	7.33
GAT	87.30	76.55	86.33	61.89	52.16	49.41	7.00
SGC	87.07	76.01	85.11	58.68	60.43	53.49	7.33
Geom-GCN	85.35	78.02	89.95	60.54	66.76	64.51	5.00
APPNP	86.76	77.08	88.45	74.59	74.30	81.10	4.66
DAGNN	87.26	74.14	84.57	58.05	58.46	52.61	8..50
GCNII	**88.27**	77.06	**90.26**	76.70	77.08	80.94	2.83
DTGNN-s	88.26	76.94	89.13	80.97	82.95	85.47	2.67
DTGNN-w	88.01	76.42	89.22	**82.92**	**83.00**	**85.57**	**2.50**

Model	Cora		Citeseer		Pubmed		Avg. Rank
	Fixed	Random	Fixed	Random	Fixed	Random	
GCN	81.50	79.91	71.42	68.78	79.12	77.84	6.17
GAT	83.10	80.80	70.80	68.90	79.10	77.80	6.00
SGC	82.63	80.18	72.10	69.33	79.12	76.74	5.83
APPNP	83.34	82.26	72.22	70.53	80.14	79.54	3.17
DAGNN	84.88 (10)	83.47 (10)	73.39 (9)	70.87 (10)	**80.51 (20)**	79.52 (20)	2.17
GCNII	**85.57 (64)**	82.58 (64)	73.24 (32)	70.04 (32)	80.00 (16)	79.03 (16)	3.00
DTGNN-s	85.35 (4)	**83.84 (6)**	**73.78 (8)**	**71.27 (8)**	80.07 (4)	**79.89 (4)**	**1.67**

Shallow layers with long range dependencies achieves the comparable or even better performance

Cora



Citeseer