Tree Decomposed Graph Neural Network

Yu Wang  **Tyler Derr** 

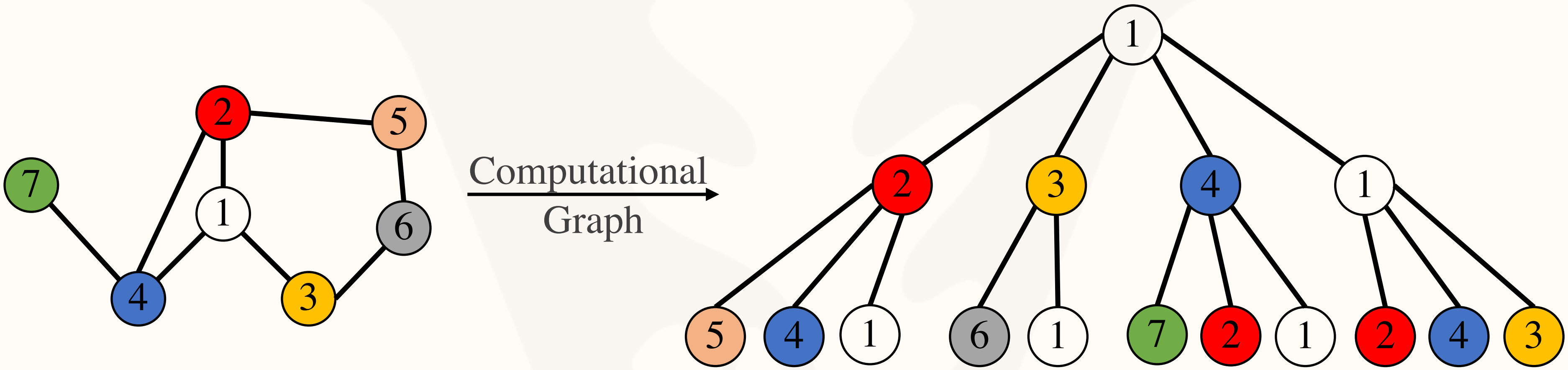
Network and Data Science Lab
Department of Computer Science

Vanderbilt University

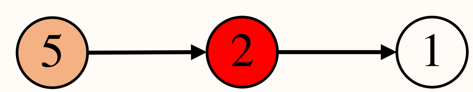
yu.wang.1@Vanderbilt.edu
https://yuwvandy.github.io/

Motivation – Tree Decomposition

Iterative propagation framework



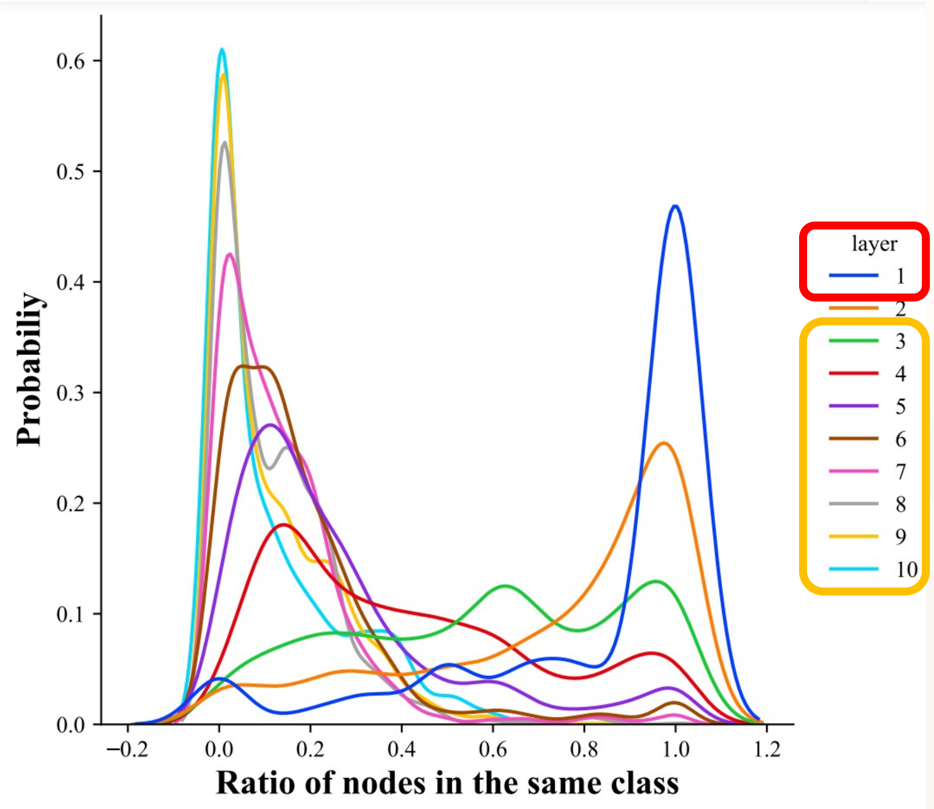
$$\widehat{\mathbf{h}}_i^l = \mathrm{AGGREGATION}^l(\mathbf{h}_i^{l-1}, \{\mathbf{h}_j^{l-1} | j \in \mathcal{N}_i\}),$$



Feature smoothing between different layers!

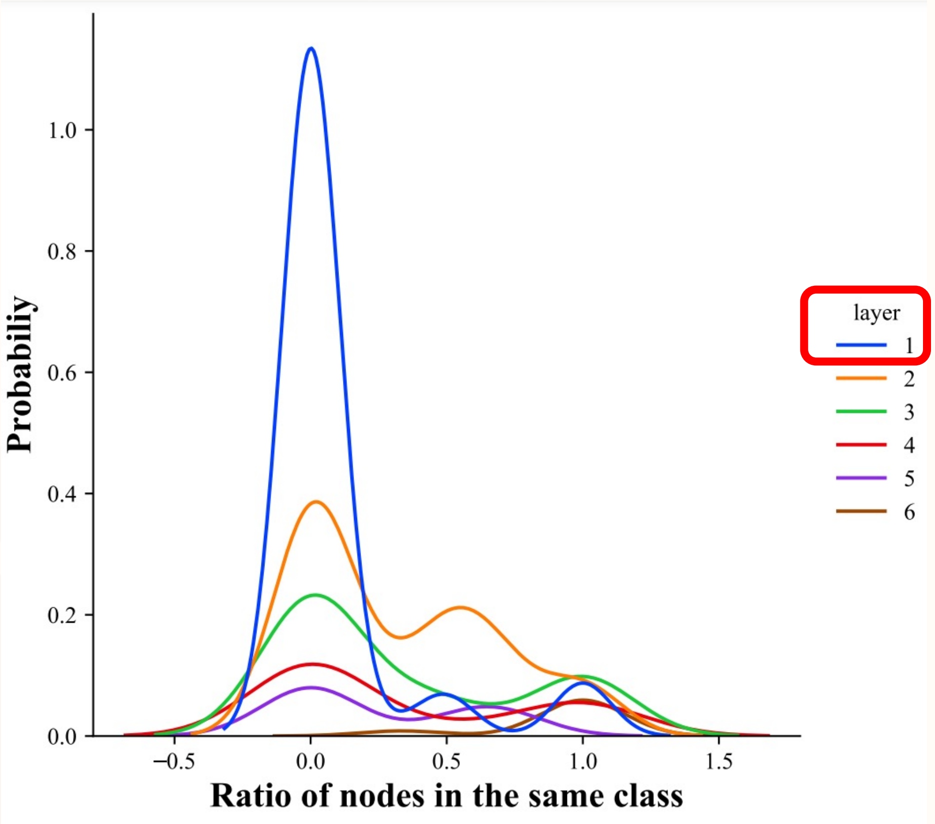
Motivation – Tree Decomposition

Cora



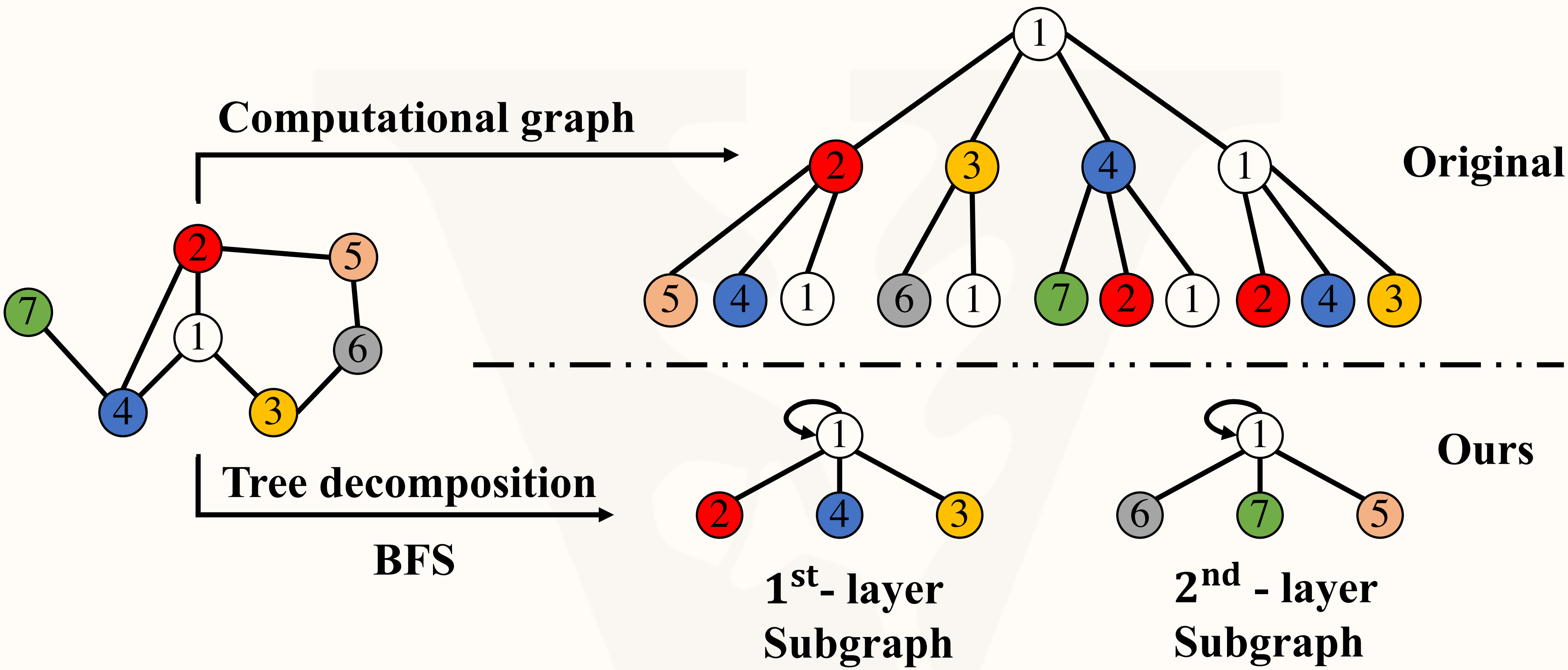
Homophily

Texas



Heterophily

Method – Tree Decomposition

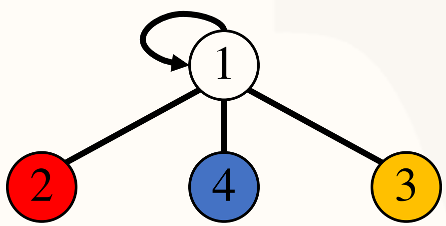


$$\mathbf{T}^k = \operatorname{sign}\left(\widehat{\mathbf{A}}^k\right) - \operatorname{sign}\left(\widehat{\mathbf{A}}^{k-1}\right) + \mathbf{I}$$

$$\operatorname{sign}\left(\widehat{\mathbf{A}}^k\right)_{ij} = \begin{cases} 1, \text{if } \widehat{\mathbf{A}}^k_{ij} > 0 \\ 0, \text{if } \widehat{\mathbf{A}}^k_{ij} = 0 \end{cases}$$

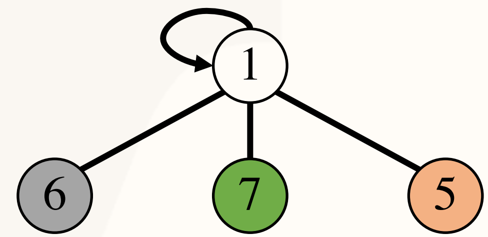
Motivation – Multi-hop dependency

What's the weight for these edges?

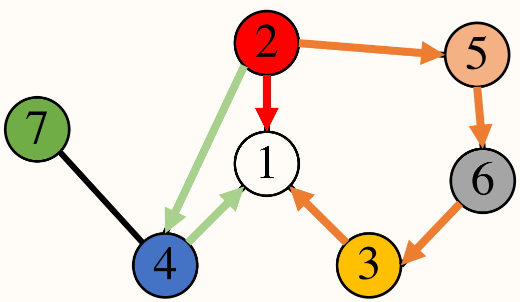


1st- layer Subgraph

Difference



2nd - layer Subgraph



$v_2 \rightarrow v_1$

$v_2 \rightarrow v_5 \rightarrow v_6 \rightarrow v_3 \rightarrow v_1$

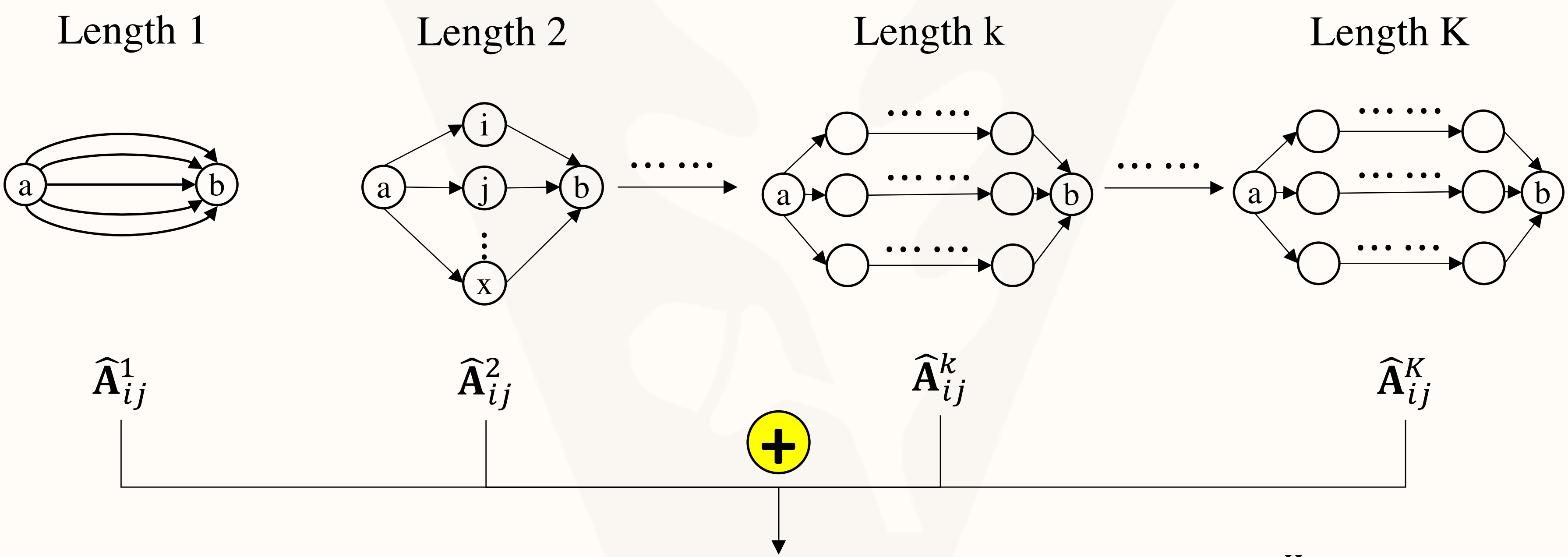
$v_2 \rightarrow v_4 \rightarrow v_1$

Multi-hop dependency

Width of GNNs

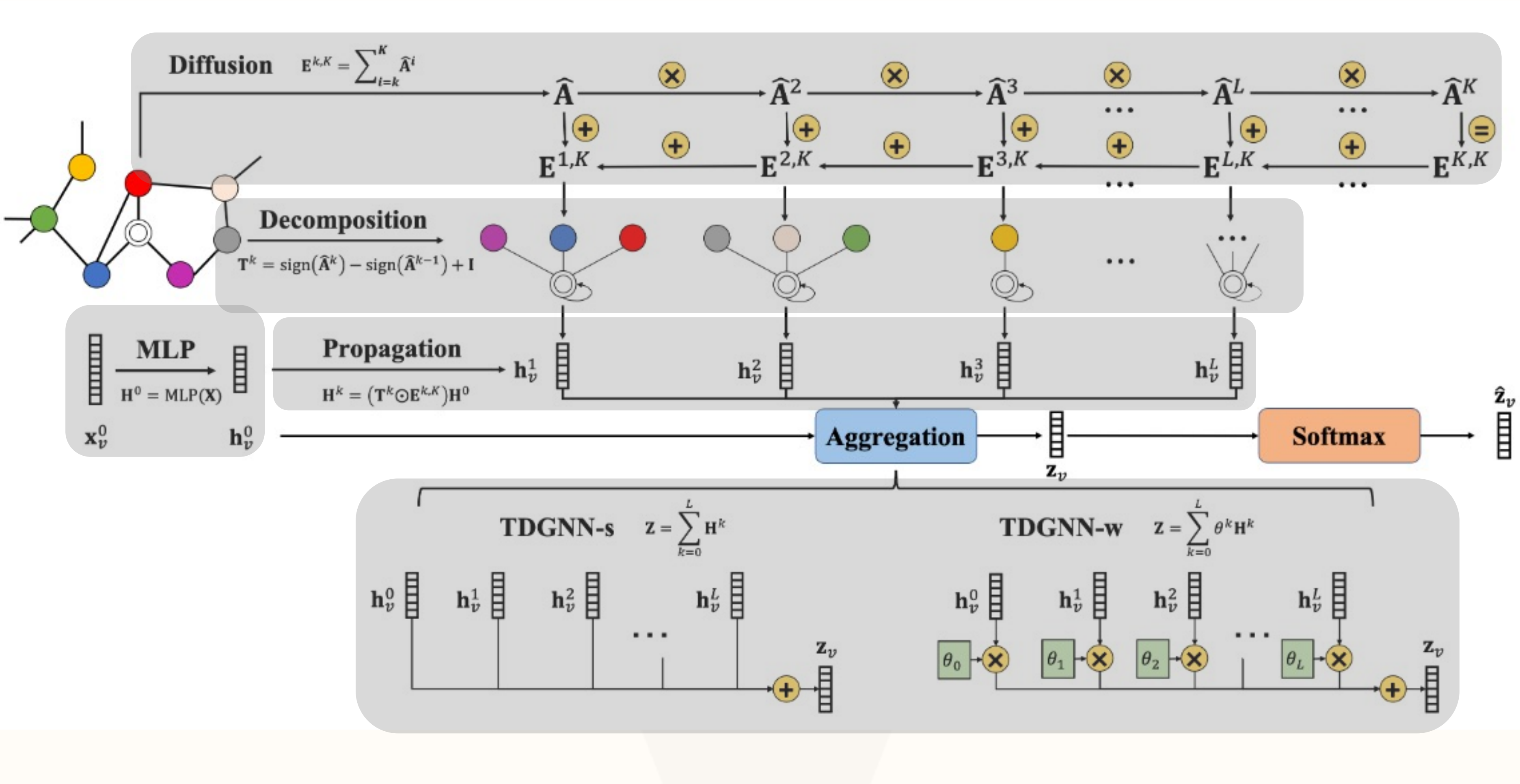
Method – Multi-hop dependency

Graph Diffusion



$$\mathbf{E}^{1,K} = \sum_{i=1}^{K} \widehat{\mathbf{A}}^{i} \qquad \mathbf{E}^{k,K} = \sum_{i=k}^{K} \widehat{\mathbf{A}}^{i}$$

Framework - TDGNN



Experiment

Table 1: Statistics of datasets.

Networks		Nodes	Edges	Features	Classes	Train/Val/Test	Type
Homophily	Cora	2708	5429	1433	7	140/500/1000	Citation network
	Citeseer	3327	4732	3703	6	120/500/1000	Citation network
	Pubmed	19717	44338	500	3	60/500/1000	Citation network
Non-homophily	Cornell	183	295	1703	5	48%/32%/20%	Webpage network
	Texas	183	309	1703	5	48%/32%/20%	Webpage network
	Wisconsin	251	499	1703	5	48%/32%/20%	Webpage network
	Actor	7600	33544	931	5	48%/32%/20%	Actor co-occurrence network

Semi-supervised setting

(1) 20 training nodes each class
(2) Fixed splitting and random splitting

Full-supervised setting

(1) 48% training nodes each class
(2) Fixed splitting

Experiment - Semi-supervised classification

Table 2: Summary of semi-supervised classification accuracy (%) ± stdev over Cora, Citeseer, and Pubmed datasets.

Method	Cora		Citeseer		Pubmed		Avg. Rank
	Fixed	Random	Fixed	Random	Fixed	Random	
GCN	81.50±0.79 (0-2)	79.91±1.64 (0-2)	71.42±0.48 (0-2)	68.78±2.01 (0-2)	79.12±0.46 (0-2)	77.84±2.36 (0-2)	7.17
GAT	83.10±0.40 (0-2)	80.80±1.60 (0-2)	70.80±0.50 (0-2)	68.90±1.70 (0-2)	79.10±0.40 (0-2)	77.80±2.10 (0-2)	7.00
SGC	82.63±0.01 (0-2)	80.18±1.57 (0-2)	72.10±0.14 (0-2)	69.33±1.90 (0-2)	79.12±0.10 (0-2)	76.74±2.84 (0-2)	6.83
APPNP	83.34±0.56 (0-10)	82.26±1.39 (0-10)	72.22±0.50 (0-10)	70.53±1.57 (0-10)	80.14±0.24 (0-10)	79.54±2.23 (0-10)	3.83
DAGNN	84.88±0.49 (0-10)	83.47±1.18 (0-10)	73.39±0.57 (0-9)	70.87±1.44 (0-10)	**80.51±0.42 (0-20)**	79.52±2.19 (0-20)	2.33
GCNII*	**85.57±0.45 (0-64)**	82.58±1.68 (0-64)	73.24±0.61 (0-32)	70.04±1.72 (0-10)	80.00±0.48 (0-16)	79.03±1.68 (0-16)	3.83
TDGNN-s	85.35±0.49 (0-4)	**83.84±1.45 (0-6)**	**73.78±0.60 (0-8)**	**71.27±1.71 (0-8)**	80.20±0.33 (0-5)	**80.01±1.96 (0-5)**	**1.33**
TDGNN-w	84.42±0.59 (0-4)	83.43±1.35 (0-6)	72.14±0.49 (0-6)	70.32±1.57 (0-6)	80.12±0.44 (0-5)	79.77±2.04 (0-5)	3.67

TDGNN-s, TDGNN-w rank 1st and 3rd

TDGNN-s, TDGNN-w utilize less layers of neighborhoods

Random setting – more robustness to data distribution

Experiment - Full-supervised classification

Table 3: Summary of full-supervised classification accuracy (%) ± stdev over 8 datasets

Method	Cora	Cite.	Pub.	Corn.	Tex.	Wisc.	Act.	Avg. Rank
MLP	75.78±1.84 (0)	73.81± 1.74 (0)	86.90±0.37 (0)	80.97±6.33 (0)	81.32± 4.19 (0)	85.38±3.95 (0)	36.60±1.25 (0)	5.57
GCN	86.97±1.32 (0-2)	76.37±1.47 (0-2)	88.19±0.48 (0-2)	58.57±3.57 (0-2)	58.68±4.64 (0-2)	53.14±6.25 (0-2)	28.65±1.38 (0-2)	8.14
GAT	87.30±1.01 (0-2)	75.55±1.32 (0-2)	85.33±0.48 (0-2)	61.89±5.05 (0-2)	58.38±6.63 (0-2)	55.29±4.09 (0-2)	28.45±0.89 (0-2)	8.00
SGC	87.07±1.20 (0-2)	76.01±1.78 (0-2)	85.11±0.52 (0-2)	58.68±3.75 (0-2)	60.43±5.11 (0-2)	53.49±5.13 (0-2)	27.46±1.46 (0-2)	8.57
Geom-GCN*	85.35±1.57 (0-2)	**78.02±1.15 (0-2)**	89.95±0.47 (N/A)	60.54±3.67 (0-2)	66.76±2.72 (N/A)	64.51±3.66 (N/A)	31.63±1.15 (N/A)	5.86
APPNP	86.76±1.74 (0-10)	77.08±1.56 (0-10)	88.45±0.42 (0-10)	74.59±5.11 (0-10)	74.30±4.74 (0-10)	81.10±2.93 (0-10)	34.36±1.09 (0-10)	5.43
DAGNN	87.26±1.42 (0-10)	76.47±1.54 (0-10)	87.49±0.63 (0-20)	80.97±6.33 (0)	81.32±4.19 (0)	85.38±3.95 (0)	36.60±1.25 (0)	4.71
GCNII*	**88.27±1.31 (0-64)**	77.06±1.67 (0-64)	**90.26±0.41 (0-64)**	76.70±5.40 (0-16)	77.08±5.84 (0-32)	80.94±4.94 (0-16)	35.18±1.30 (0-64)	3.71
TDGNN-s	88.26±1.32 (0-4)	76.64±1.54 (0-8)	89.13±0.39 (0-1)	80.97±6.33 (0)	82.95±4.59 (0, 4-5)	85.47±3.88 (0, 4-5)	36.70±1.28 (0, 3-4)	2.86
TDGNN-w	88.01±1.32 (0-5)	76.58±1.40 (0-2)	89.22±0.41 (0-1)	**82.92±6.61 (0, 2-6)**	**83.00±4.50 (0, 2)**	**85.57±3.78 (0, 3-5)**	**37.11±0.96 (0, 3-4)**	**2.14**

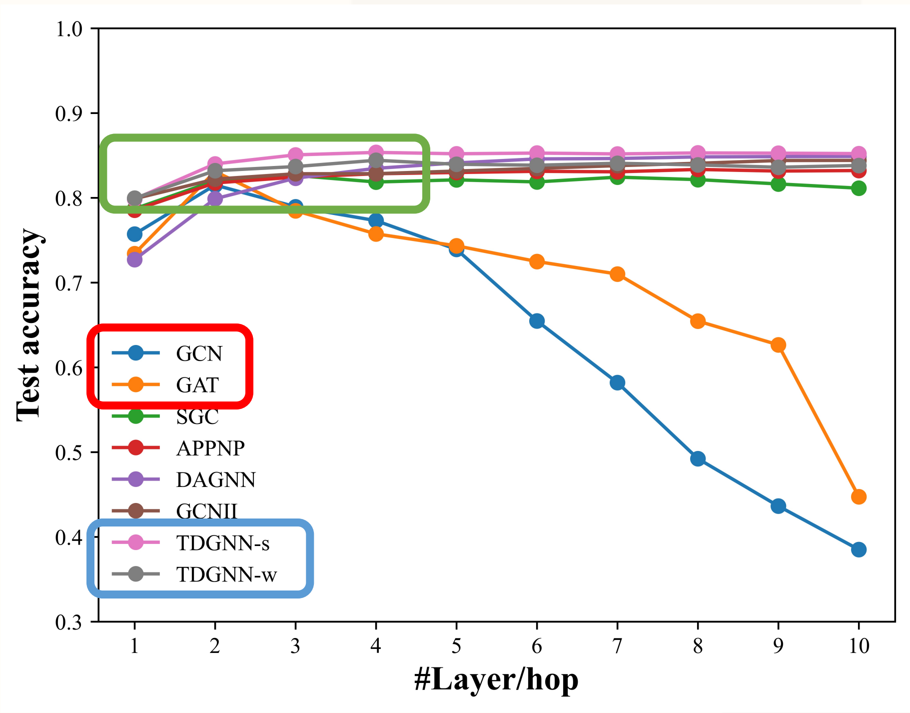
* We reuse the results reported in [33] for Geom-GCN. 'N/A' indicate the corresponding layers are not reported in the paper.

TDGNN-w, TDGNN-s rank 1st and 2nd

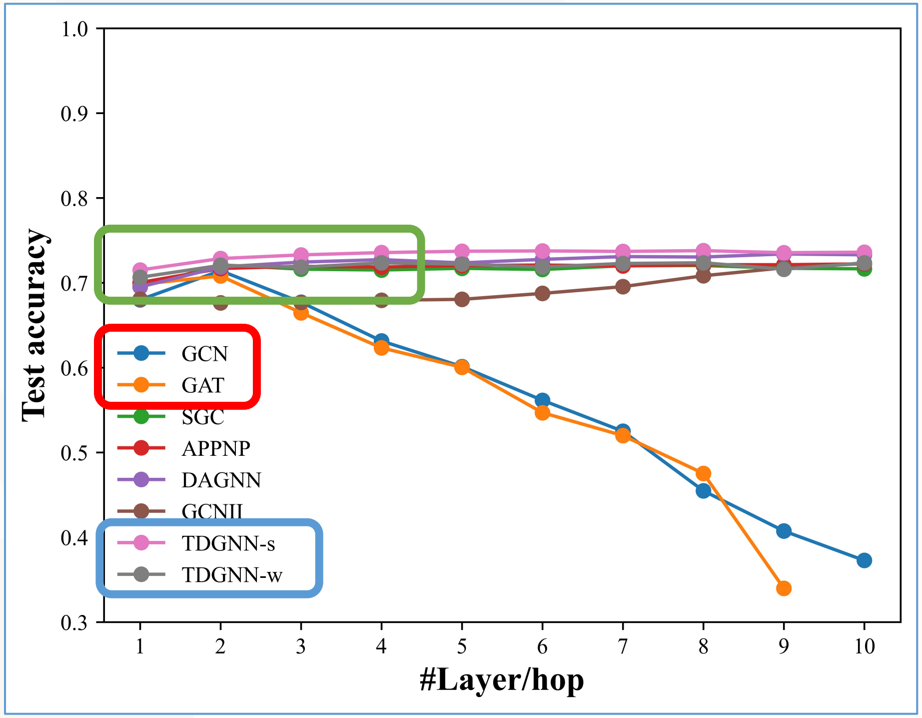
TDGNN-s, TDGNN-w leverage different layers of neighborhoods

Experiment – Further Probe

Alleviate over-smoothing

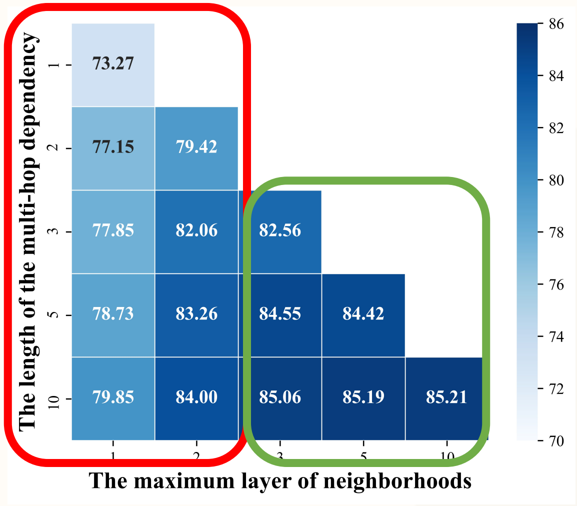


Cora

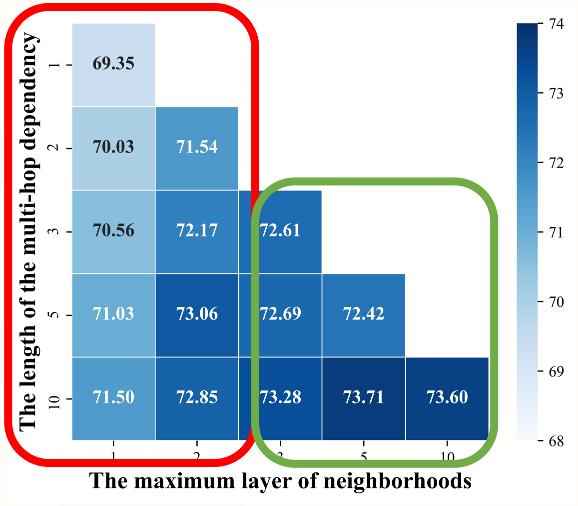


Citeseer

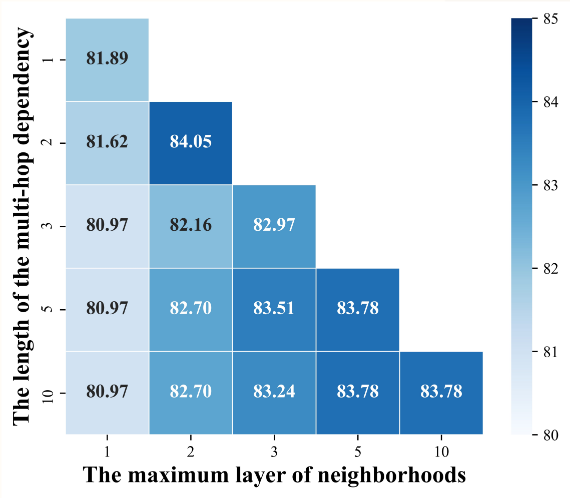
Experiment – Further Probe



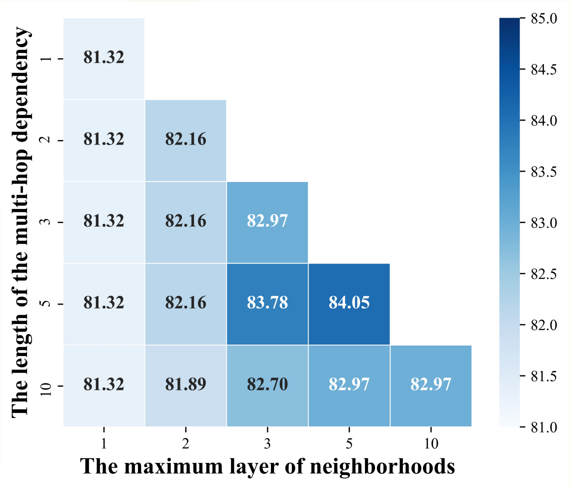
Cora



Citeseer



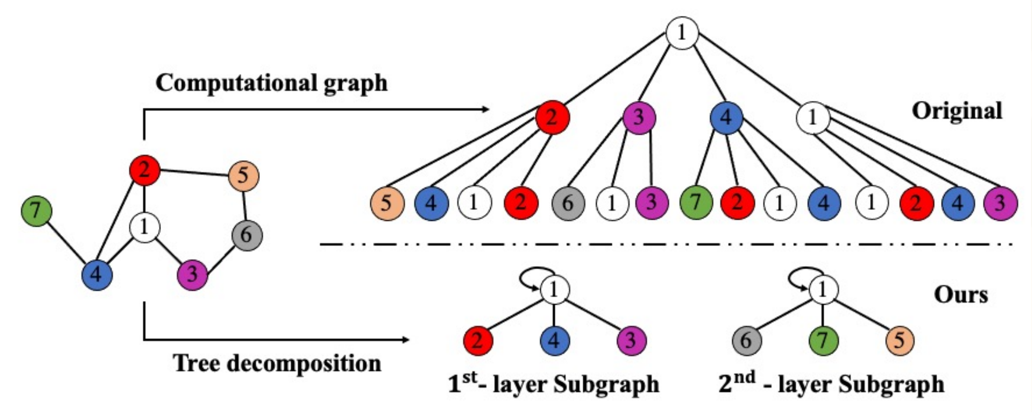
Texas



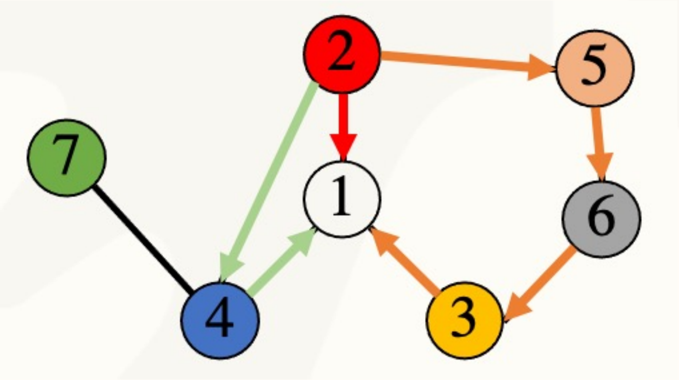
Wisconsin

Conclusion

Tree decomposition to alleviate over-smoothing between different layers



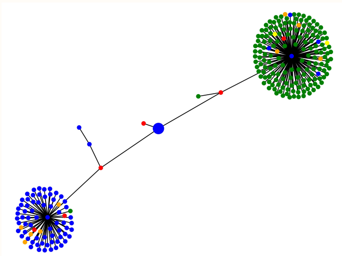
Graph diffusion to incorporate multi-hop dependencies

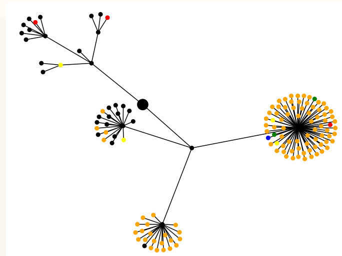


Width is also important compared with depth!

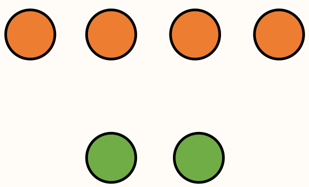
Future Directions

Layer adaptive -> node adaptive



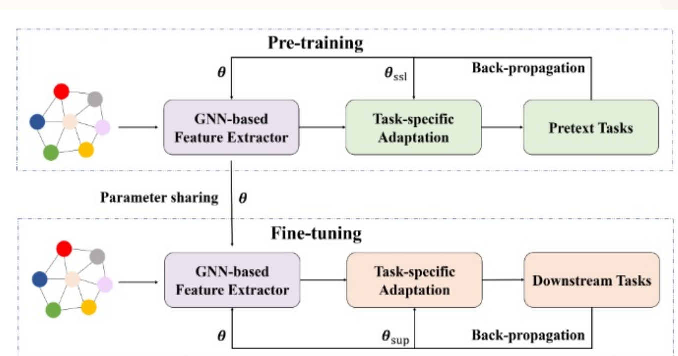


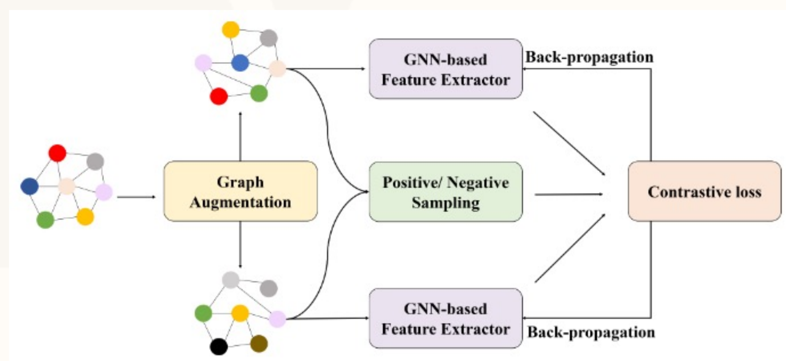
Node/Graph imbalance classification



$$\mathcal{L} = \mathcal{L}_{\text{min}} + \mathcal{L}_{\text{maj}}$$

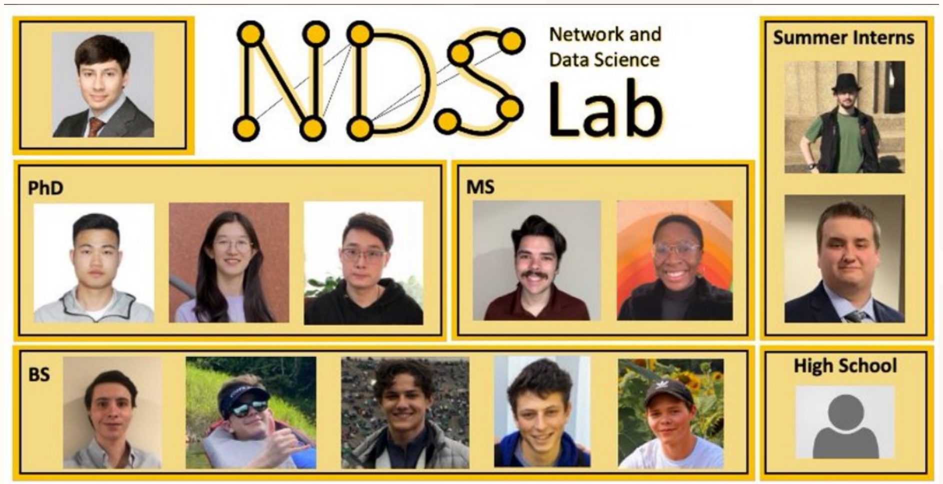
Incorporate self-supervised learning with deeper GNNs





Acknowledgement

Project webpage:
https://github.com/YuWVandy/TDGNN



Thank you!
Any question?

Please see my homepage for more details!

https://yuwvandy.github.io/