Tree Decomposed Graph Neural Network


Yu Wang


Tyler Derr

Network and Data Science Lab, Department of Computer Science, Vanderbilt University

Network and Data Science Lab

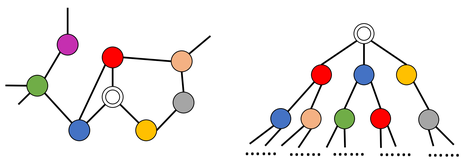
Acknowledgements

We thank SIGIR for their support through the student travel award.

SIGIR Special Interest Group on Information Retrieval

Motivation

Over-smoothing among different layers



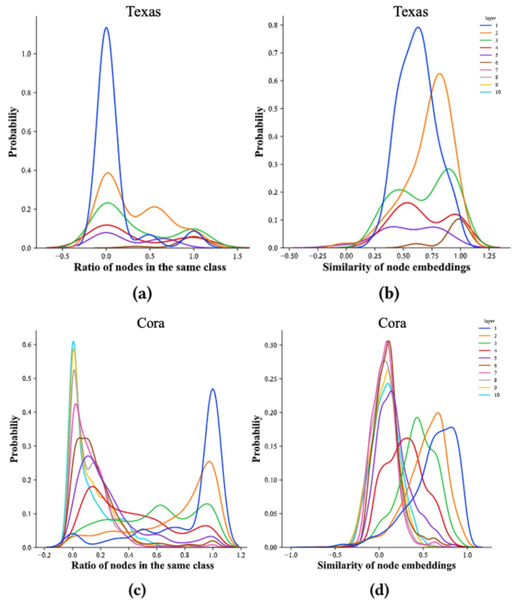
$$\hat{\mathbf{h}}_i^l = \text{AGGREGATE}^l\left(\mathbf{h}_i^{l-1}, \{\mathbf{h}_j^{l-1} | j \in \mathcal{N}_i\}\right),$$
$$\mathbf{h}_i^l = \text{TRANSFORMATION}^l(\hat{\mathbf{h}}_i^l)$$

$$\mathbf{h}_i^2 = \sigma(\mathbf{h}_i^0\mathbf{W}^0)\mathbf{W}^1\left(\frac{1}{(d_i+1)^2} + \sum_{j\in\mathcal{N}_i}\frac{1}{(d_i+1)(d_j+1)}\right)$$ 0th-layer feature (self)

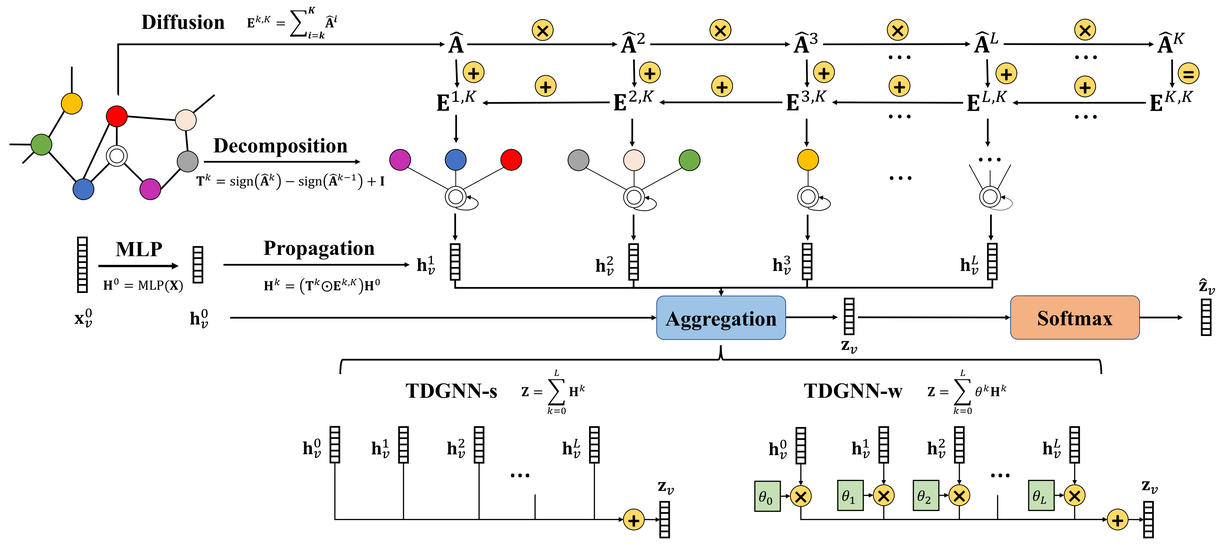
$$+\sum_{j\in\mathcal{N}_i}\sigma(\mathbf{h}_i^0\mathbf{W}^0)\mathbf{W}^1\left(\frac{d_i+d_j+2}{(d_i+1)^{1.5}(d_j+1)^{1.5}}\right)$$ 1st-layer neighborhood features

$$+\sum_{j\in\mathcal{N}_i}\sum_{k\in\{\mathcal{N}_j\cap\mathcal{N}_i\}}\sigma(\mathbf{h}_k^0\mathbf{W}^0)\mathbf{W}^1\frac{1}{\sqrt{d_i+1}\sqrt{d_k+1}(d_j+1)}$$ 1st-layer neighborhood features

$$+\sum_{j\in\mathcal{N}_i}\sum_{k\in\{\mathcal{N}_j\cap\mathcal{N}_i^2\}}\sigma(\mathbf{h}_k^0\mathbf{W}^0)\mathbf{W}^1\frac{1}{\sqrt{d_i+1}\sqrt{d_k+1}(d_j+1)}$$ 2nd-layer neighborhood features



Tree Decomposed Graph Neural Network

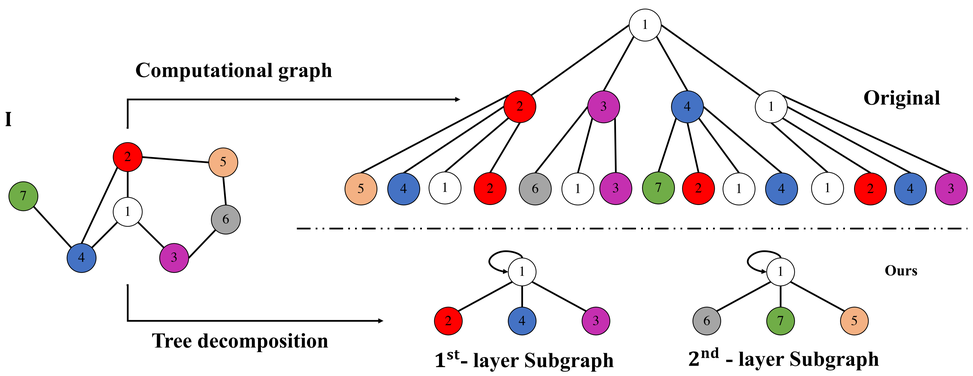


Tree Decomposition and Multi-hop Dependency

Tree Decomposition

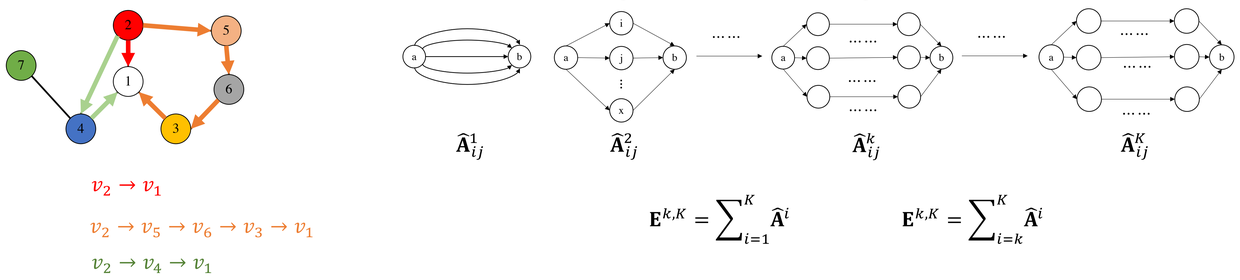
$$\mathbf{T}^k = \text{sign}(\hat{\mathbf{A}}^k) - \text{sign}(\hat{\mathbf{A}}^{k-1}) + \mathbf{I}$$

$$\text{sign}(\hat{\mathbf{A}}^k)_{ij} = \begin{cases} 1, \text{if } \hat{\mathbf{A}}_{ij}^k > 0 \\ 0, \text{if } \hat{\mathbf{A}}_{ij}^k = 0 \end{cases}$$



Multi-hop Dependency

Width of GNNs



$v_2 \to v_1$

$v_2 \to v_5 \to v_6 \to v_3 \to v_1$

$v_2 \to v_4 \to v_1$

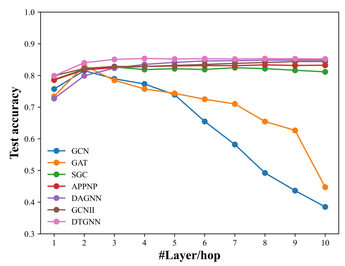
$$\mathbf{E}^{k,K} = \sum_{i=1}^{K}\hat{\mathbf{A}}^i \qquad \mathbf{E}^{k,K} = \sum_{i=k}^{K}\hat{\mathbf{A}}^i$$

Real World Graph Result

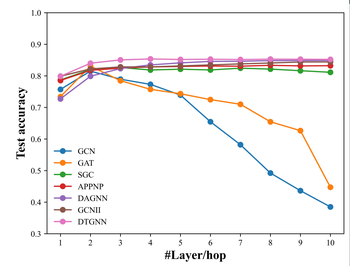
Model	Cora	Cite.	Pub.	Corn.	Tex.	Wisc.	Act.
GCN	86.97	76.37	88.19	58.57	58.68	53.14	28.65
GAT	87.30	76.55	85.33	61.89	58.38	55.29	28.45
SGC	87.07	76.01	85.11	58.68	60.43	53.49	27.46
Geom-GCN	85.35	78.02	89.95	60.54	66.76	64.51	31.63
APPNP	86.76	77.08	88.45	74.59	74.30	81.10	34.36
DAGNN	87.26	76.47	87.49	80.97	81.32	85.38	36.60
GCNII	**88.27**	77.06	**90.26**	76.70	77.08	80.94	**35**.18
DTGNN-s	88.26	76.64	89.13	80.97	82.95	85.47	36.70
DTGNN-w	88.01	76.58	89.22	**82.92**	**83.00**	**85.57**	**37.11**

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	7.17
GAT	83.10	80.80	70.80	68.90	79.10	77.80	7.00
SGC	82.63	80.18	72.10	69.33	79.12	76.74	6.83
APPNP	83.34	82.26	72.22	70.53	80.14	79.54	3.83
DAGNN	84.88	83.47	73.39	70.87	**80.51**	79.52	2.33
GCNII	**85.57**	82.58	73.24	70.04	80.00	79.03	3.83
TDGNN-s	**8**5.35	**83.84**	**73.78**	**71.27**	80.20	**80.01**	**1.33**
TDGNN-w	84.42	83.43	72.14	70.32	80.12	79.77	3.67

Cora



Citeseer



Shallow layers with long range dependencies achieves the comparable performance