Beyond Homophily: Structure-aware Path Aggregation Graph Neural Network

Yifei Sun[1], Haoran Deng[1], Yang Yang[1*], Chunping Wang[2], Jiarong Xu[3], Renhong Huang[1], Linfeng Cao[1], Yang Wang[2] and Lei Chen[2]
[1]College of Computer Science and Technology, Zhejiang University
[2]Finvolution Group
[3]Department of Information Management and Business Intelligence, Fudan University
{yifeisun, denghaoran, yangya, renh2}@zju.edu.cn, jiarongxu@fudan.edu.cn, {wangchunping02, wangyang09, chenlei04}@xinye.com, linfengcao1996@gmail.com

IJCAI 2022
Code:None

2022.09.07 • ChongQing

Leibniz-Institut für Sozialwissenschaften

Reported by Chenghong Li

Introduction

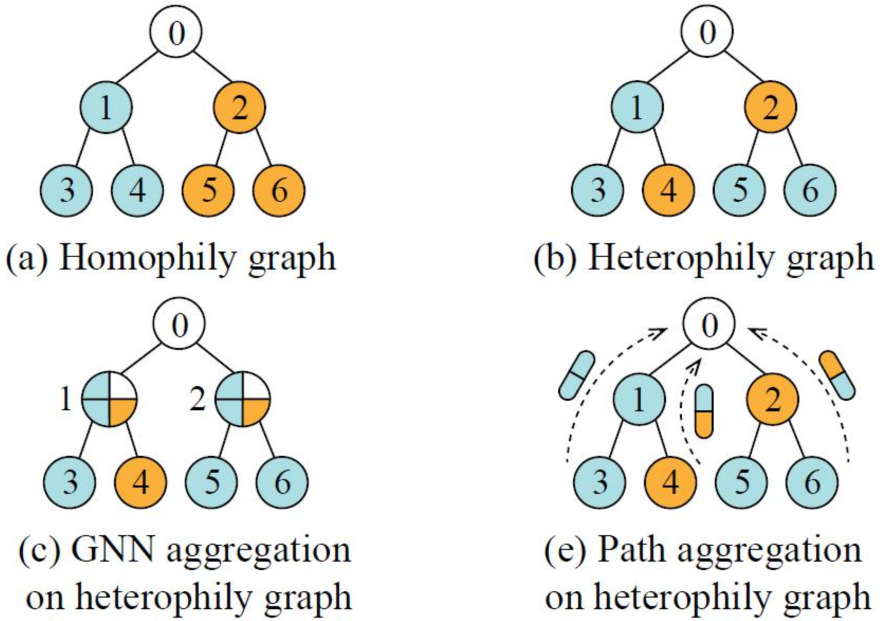

Figure 1: (a) & (b) show patterns of homophily and heterophily graphs. (c) demonstrates the GNN aggregation process for node 0 of (b). (d) our proposed path aggregation compared with (c).

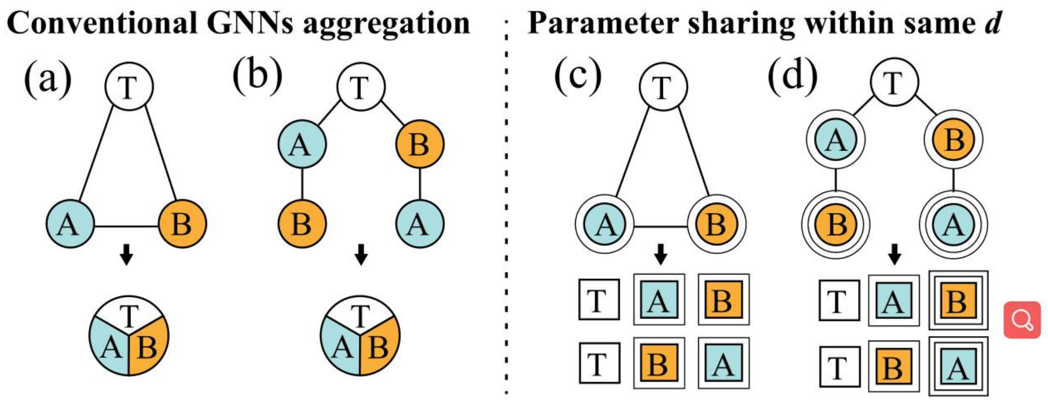

Figure 3: Intuitive case where GNNs are not able to distinguish the target nodes in (a) & (b), while our path-based aggregation with parameter sharing mechanism can capture the topological information and make the embedding distinguishable (c) & (d). (Nodes with similar characteristic represented as the same notation.)

Method

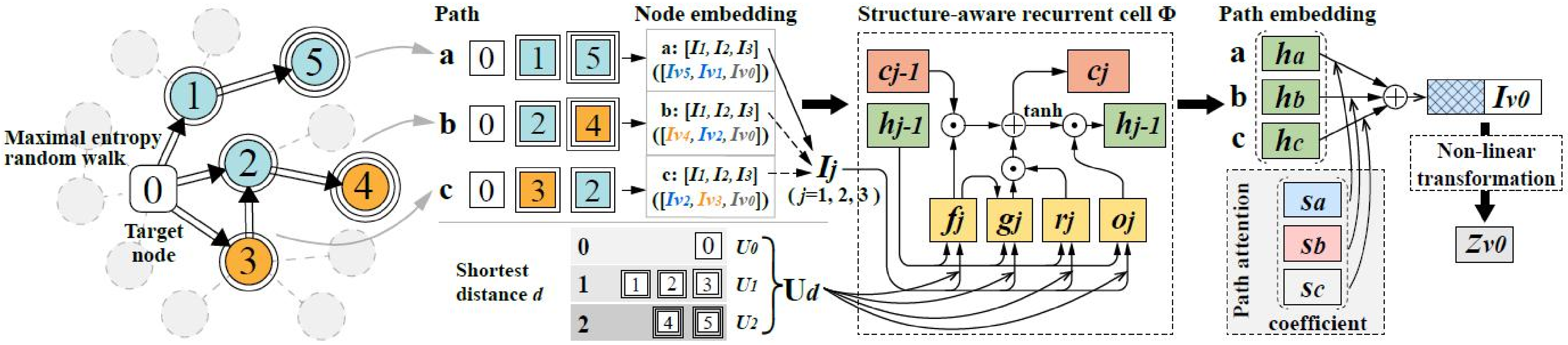

Figure 2: The workflow of PathNet for node classification of node 0 with the walk length $k = 3$. The color of node stands for label and the number stands for node index.

Method

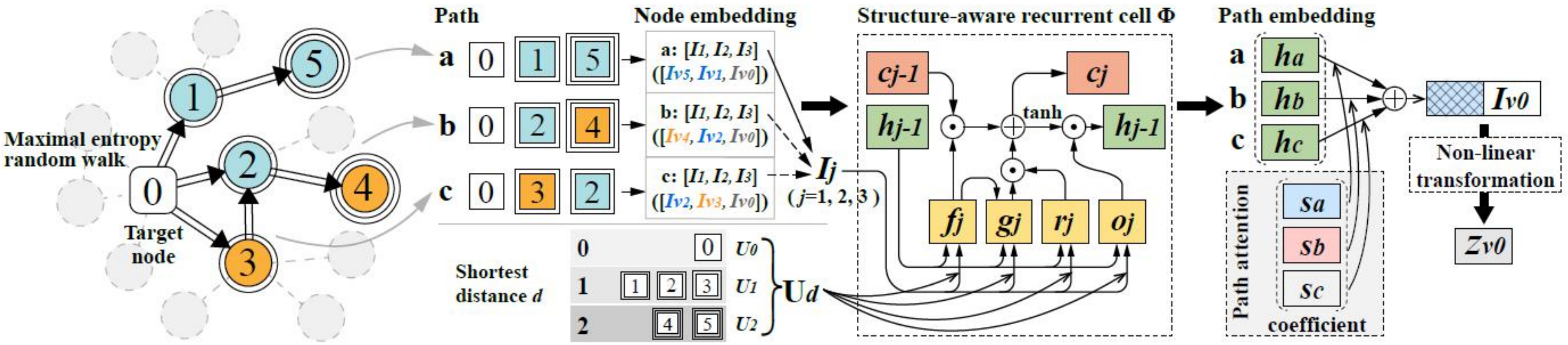

Figure 2: The workflow of PathNet for node classification of node 0 with the walk length $k = 3$. The color of node stands for label and the number stands for node index.

$$\eta = -\sum_i \pi_i \sum_j p_{ij} \ln p_{ij}. \tag{1}$$

$$\mathbf{P}_u = \frac{\mathbf{D}_u^{-1}\mathbf{A}\mathbf{D}_u}{\lambda}, \tag{2}$$

Method

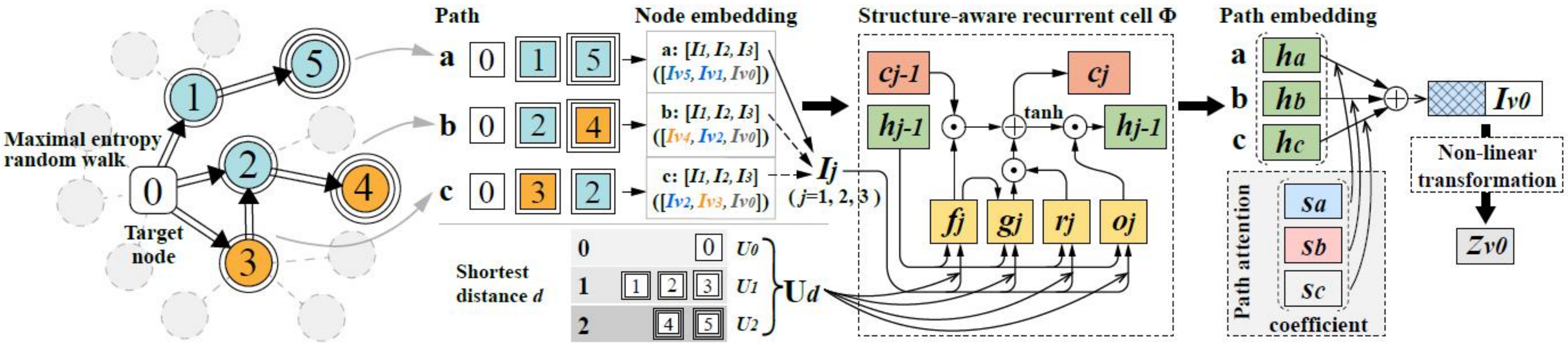

Figure 2: The workflow of PathNet for node classification of node 0 with the walk length $k = 3$. The color of node stands for label and the number stands for node index.

$$\mathbf{I}=\sigma\left(\mathbf{W_{in}}\mathbf{X}+\mathbf{b_{in}}\right), \tag{3}$$

$$\begin{aligned}
\mathbf{r}_j &= \sigma\left(\mathbf{W_r}\cdot h_{j-1}+\mathbf{U_d}\cdot\mathbf{I}_j\right),\\
\mathbf{f}_j &= \sigma\left(\mathbf{W_f}\cdot h_{j-1}+\mathbf{U_d}\cdot\mathbf{I}_j\right),\\
\mathbf{o}_j &= \sigma\left(\mathbf{W_o}\cdot h_{j-1}+\mathbf{U_d}\cdot\mathbf{I}_j\right),\\
\mathbf{g}_j &= \tanh\left(\mathbf{W_g}\cdot h_{j-1}+\mathbf{U_d}\cdot\mathbf{I}_j\right),\\
\mathbf{c}_j &= \mathbf{f}_j\odot\mathbf{c}_{j-1}+\mathbf{r}_j\odot\mathbf{g}_j,\\
\mathbf{h}_j &= \mathbf{o}_t\odot\tanh\left(\mathbf{c}_j\right),
\end{aligned} \tag{4}$$

$$\boldsymbol{s}_{v,p}=\mathrm{SOFTMAX}\left(\delta\left(\boldsymbol{a}\left(\mathbf{I}_v\|\mathbf{h}_p\right)\right)\right), \tag{5}$$

$$\mathbf{z}_v=\sigma\left(\mathbf{W_{out}}\left(\mathbf{I}_v\|\sum_{j\in p}\boldsymbol{s}_{v,p}\mathbf{h}_p\right)+\mathbf{b_{out}}\right). \tag{6}$$

Experiments

		Cora	Pubmed	Citeseer	Cornell	NBA	BGP	Electronics
#Hom. ratio		0.81	0.80	0.74	0.30	0.39	0.37	0.25
Baselines	MLP	74.75± 2.22	86.65±0.35	72.41±2.18	81.08±6.37	59.21±6.92	63.39±0.34	75.03±0.08
	GIN	84.97±1.51	86.97±0.53	72.19±1.74	58.10±5.70	65.47±6.85	OOM	OOM
	GAT	82.68±1.80	84.68±0.44	75.46±1.72	58.92±3.32	67.19±1.04	62.25±0.90	64.64±0.27
	GraphSage	**86.90±1.04**	88.45±0.50	76.04±1.30	75.95±5.01	61.70±2.40	61.71±0.85	74.92±0.19
	MixHop	85.41±1.61	86.38±0.46	75.43±1.89	72.51±6.36	68.89±5.95	64.80±0.83	67.84±0.50
	H2GCN	86.21±0.98	87.86±0.19	76.73±1.48	81.27±4.63	66.67±7.02	65.13±1.01	73.92±0.52
	GPRGNN	86.00±2.46	86.56±0.29	78.45±0.27	50.82±3.28	48.25±4.97	61.49±0.40	75.79±0.16
	FAGCN	86.30±1.74	88.50±0.27	76.20±1.45	72.70±4.50	63.49±3.89	64.48±0.55	71.10±2.02
	P-GNN	68.05±1.30	84.97±0.38	64.81±1.29	58.65±3.21	58.41±7.40	54.04±3.81	57.25±2.78
	GeniePath	85.15±0.65	86.50±0.34	76.46±1.42	59.19±4.43	68.73±5.41	63.15±2.94	73.39±0.35
	SPAGAN	86.12±0.54	85.10±0.19	77.41±0.82	55.41±2.18	53.65±7.23	52.59±0.67	53.93±5.08
Ablation	PathNet-MLP	82.89±2.84	87.86±0.07	75.78±1.50	90.54±1.35	69.05±6.08	64.36±0.54	75.81±0.58
	PathNet-GRU	84.76±1.52	87.89±0.12	76.57±1.08	90.74±1.81	69.52±7.16	64.46±0.76	76.16±0.52
	PathNet-LSTM	84.39±2.77	87.89±0.14	76.44±2.86	91.35±1.62	69.37±6.27	65.19±0.79	76.12±0.39
	RW-PathNet	85.08±1.17	87.84±0.34	78.54±2.13	90.27±2.16	71.27±5.65	64.92±0.61	76.31±0.45
	PathNet-Mean	83.46±2.36	88.18±3.94	76.59±1.61	91.08±2.43	70.16±6.18	64.81±0.77	76.85±0.55
	PathNet-Sum	84.21±1.43	86.93±0.27	74.80±2.44	89.19±2.70	67.70±6.44	65.39±0.83	75.06±0.44
	PathNet	85.76±2.67	**88.92±0.21**	**77.98±2.40**	**91.35±2.91**	**71.69±4.83**	**65.72±0.66**	**76.97±0.84**

Table 1: Mean accuracy and standard deviation of PathNet on node classification compared with baselines and ablation study.

Experiments

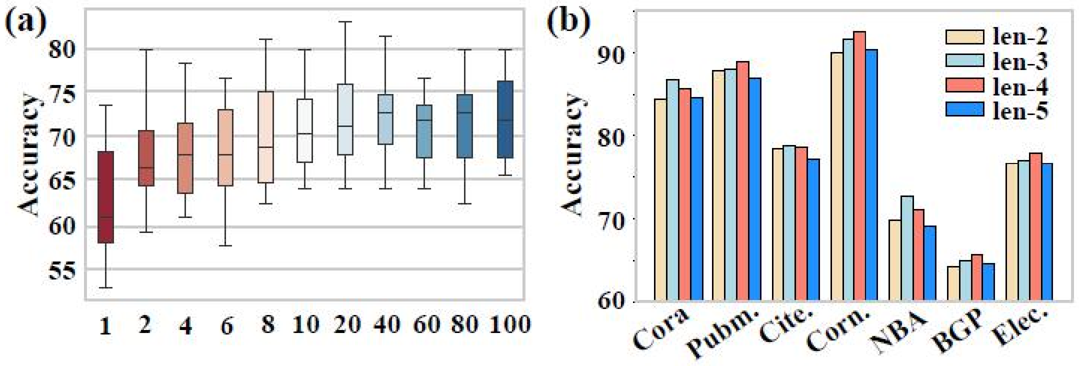

Figure 4: Model variants of different number (a) and length (b) of sampled paths.

	Syn-Cora	Syn-Citeseer
#Hom. ratio $\mathcal{H}_\mathcal{G}$	0.37	0.39
GIN	51.40±1.55	59.09±2.71
GAT	36.96±1.60	47.44±1.76
GPRGNN	43.62±1.69	53.95±1.87
PathNet	**57.59±1.54**	**71.42±1.15**

Table 2: Node classification on synthetic datasets.

Thanks