EXPLORING ADAPTIVE STRUCTURE LEARNING FOR HETEROPHILIC GRAPHS

Anonymous authors
Paper under double-blind review

ABSTRACT

Graph Convolutional Networks (GCNs) gained traction for graph representation learning, with recent attention on improving performance on heterophilic graphs for various real-world applications. The localized feature aggregation in a typical message-passing paradigm hinders the capturing of long-range dependencies between non-local nodes of the same class. The inherent connectivity structure in heterophilic graphs often conflicts with information sharing between distant nodes of same class. We propose structure learning to rewire edges in shallow GCNs itself to avoid performance degradation in downstream discriminative tasks due to oversmoothing. Parameterizing the adjacency matrix to learn connections between non-local nodes and extend the hop span of shallow GCNs facilitates the capturing of long-range dependencies. However, our method is not generalizable across heterophilic graphs and performs inconsistently on node classification task contingent to the graph structure. [1]

1 INTRODUCTION

The application of Graph Convolutional Networks (GCNs) (Kipf & Welling, 2016) (Chen et al., 2020) (Song et al., 2023) (Yu et al., 2024) gained momentum for their efficacy in learning from graph-structured data. Shallow GCNs are typically fail to harness information from distant nodes through conventional message-passing, and the trivial solution to stack numerous convolution layers to incrementally aggregate multi-hop information leads to oversmoothing (Li et al., 2018) consequently hurting performance on downstream tasks. In the past there have been attempts (Yan et al., 2021) (Zheng et al., 2022) to develop more expressive GNN architectures for heterophilic graph datasets. However Structure Learning (SL) (Wu et al., 2023) emerges as a potential alternative to deep GCNs, utilizing a transformed adjacency matrix that adapts graph structure to the downstream tasks thus tackling the problem of unexplored non-local topology as explained in Zheng et al. (2022). Empirical evaluations reveal a correspondence between the performance of GCN and the graph properties.

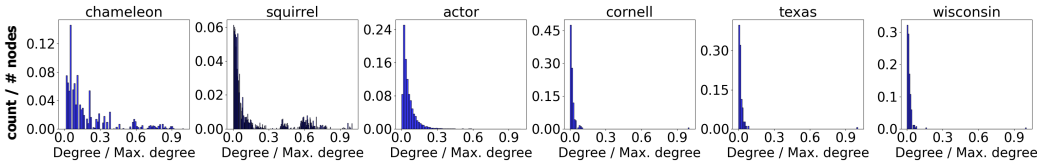

Figure 1: Distribution of the node degrees are presented for 6 heterophilic graphs

2 PROPOSED METHOD

Assume G is a heterophilic graph with node feature matrix $\boldsymbol{X} \in \mathbb{R}^{n \times d}$ and adjacency matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ where n and d are the number of nodes and feature dimension respectively. To apply

[1] Code is available at: `https://github.com/GARV-k/GSL`

Structure Learning on the graph's adjacency matrix we define a transformation on $\boldsymbol{A}$ as follows :

$$\hat{\boldsymbol{A}} = \sigma(\boldsymbol{A}\tilde{\boldsymbol{W}} + \boldsymbol{b}), \tag{1}$$

where $\tilde{\boldsymbol{W}} \in \mathbb{R}^{n \times n}$ and $\boldsymbol{b} \in \mathbb{R}^{n}$ are end to end trainable parameters , stagnant for all layers of the proposed network and $\hat{\boldsymbol{A}}$ denotes the transformed adjacency matrix with $\sigma(.)$ as ReLu activation. We term our method as **Graph Structure Learning (GSL)** and suggest to feed the transformed adjacency $\hat{\boldsymbol{A}}$ to the GCN model, as follows,

$$\boldsymbol{X}^{(l+1)} = \text{GCN}(\hat{\boldsymbol{A}}, \boldsymbol{X}^{(l)}, \boldsymbol{W}^{(l+1)}), \tag{2}$$

where $\boldsymbol{X}^{l}$ is the updated node features at the l^{th} layer and $\boldsymbol{W}^{(l+1)}$ is the trainable parameter of the $(l+1)^{th}$ layer of the GCN. It is important to note that $\tilde{\boldsymbol{W}}$ and $\boldsymbol{W}^{(l+1)}$ are trained simultaneously in contrast to Zheng et al. (2020), solely by backpropagation from the conventional cross entropy loss applied on the predicted and ground truth labels for the nodes. This startegy offers a notion of the edge rewiring (Barbero et al., 2024) of the original graph connections. The transformed adjacency matrix introduces continuous-valued edge weights enhancing the expressiveness of our framework.

Table 1: Properties of six heterophilic graphs are presented

Properties	**Chameleon**	**Squirrel**	**Actor**	**Cornell**	**Texas**	**Wisconsin**
Edge Homophily	0.23	0.22	0.22	0.3	0.11	0.21
# Isolated Nodes(%)	0	0	8.37	47.54	39.89	32.27
Avg. Degree	15.85	41.74	3.95	1.63	1.78	2.05
Edge density	0.0139	0.016	0.001	0.0178	0.0195	0.0164

3 EXPERIMENTS

We compare our results to GCN (Kipf & Welling, 2016) and GCNII (Chen et al., 2020) on 6 heterophilic datasets (Table 1) as our proposed method has a backbone of GCN. Figure 2 demonstrates that our method performs well on heterophilic graphs with high average degree and edge density such as Chameleon and Squirrel. However, it exhibits poor test accuracy on other datasets with skewed degree distributions (Figure 1).

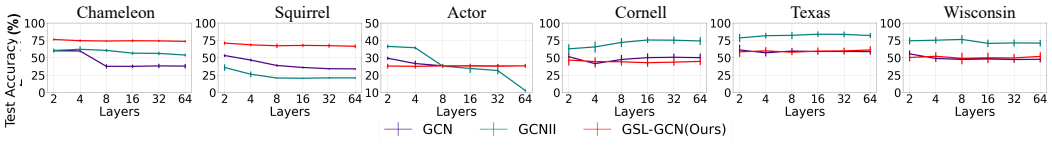

Figure 2: Performance of 6 datasets with 10-fold cross-validation.

4 CONCLUSION

Despite the intuitive hypothesis of adaptive adjacency matrix's ability to extend the hop span of shallow GCNs to circumvent the adverse effects of oversmoothing in deeper GCNs, such a structure learning approach does not result in consistent performance across heterophilic graphs.

REFERENCES

Federico Barbero, Ameya Velingker, Amin Saberi, Michael M. Bronstein, and Francesco Di Giovanni. Locality-aware graph rewiring in GNNs. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=4Ua4hKiAJX.

Ming Chen, Zhewei Wei, Zengfeng Huang, Bolin Ding, and Yaliang Li. Simple and deep graph convolutional networks. In *International Conference on Machine Learning*, pp. 1725–1735. PMLR, 2020.

Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

Qimai Li, Zhichao Han, and Xiao-Ming Wu. Deeper insights into graph convolutional networks for semi-supervised learning. In *Thirty-Second AAAI conference on artificial intelligence*, 2018.

Yunchong Song, Chenghu Zhou, Xinbing Wang, and Zhouhan Lin. Ordered GNN: Ordering message passing to deal with heterophily and over-smoothing. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=wKPmPBHSnT6.

Lirong Wu, Cheng Tan, Zihan Liu, Zhangyang Gao, Haitao Lin, and Stan Z Li. Learning to augment graph structure for both homophily and heterophily graphs. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 3–18. Springer, 2023.

Yujun Yan, Milad Hashemi, Kevin Swersky, Yaoqing Yang, and Danai Koutra. Two sides of the same coin: Heterophily and oversmoothing in graph convolutional neural networks. *CoRR*, abs/2102.06462, 2021. URL https://arxiv.org/abs/2102.06462.

Zhizhi Yu, Bin Feng, Dongxiao He, Zizhen Wang, Yuxiao Huang, and Zhiyong Feng. Lg-gnn: Local-global adaptive graph neural network for modeling both homophily and heterophily. In Kate Larson (ed.), *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*, pp. 2515–2523. International Joint Conferences on Artificial Intelligence Organization, 8 2024. doi: 10.24963/ijcai.2024/278. URL https://doi.org/10.24963/ijcai.2024/278. Main Track.

Cheng Zheng, Bo Zong, Wei Cheng, Dongjin Song, Jingchao Ni, Wenchao Yu, Haifeng Chen, and Wei Wang. Robust graph representation learning via neural sparsification. In Hal Daumé III and Aarti Singh (eds.), *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 11458–11468. PMLR, 13–18 Jul 2020. URL https://proceedings.mlr.press/v119/zheng20d.html.

Xin Zheng, Yixin Liu, Shirui Pan, Miao Zhang, Di Jin, and Philip Yu. Graph neural networks for graphs with heterophily: A survey. 02 2022. doi: 10.48550/arXiv.2202.07082.