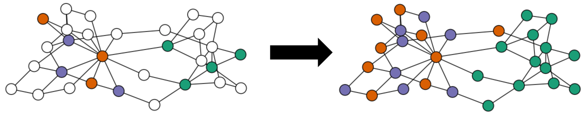
Beyond Homophily in Graph Neural Networks: Current Limitations and Effective Designs

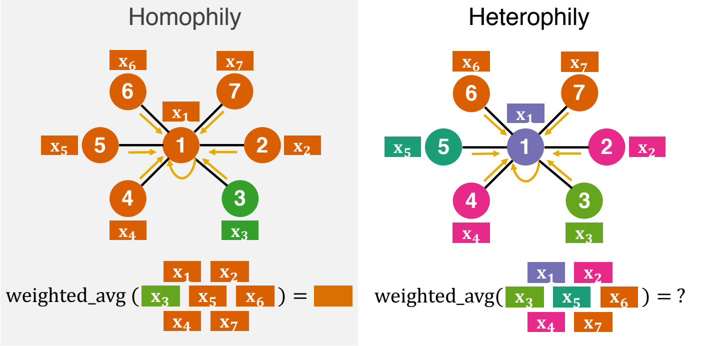
Jiong Zhu, Yujun Yan, Lingxiao Zhao, Mark Heimann, Leman Akoglu, Danai Koutra

Limitations of GNNs Beyond Homophily

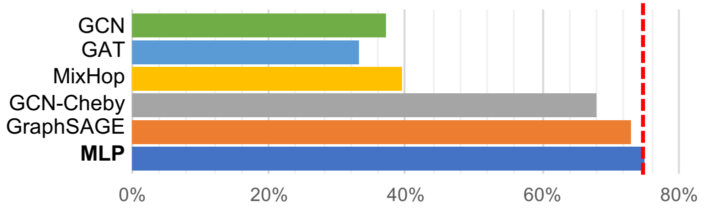
Task: semi-supervised node classification with node features



Problem: many popular GNN models (e.g. GCN) rely on assumed **homophily** and fail to generalize in **heterophily**.



Observation: In heterophily, existing methods have worse classification accuracy than graph-agnostic MLP.



References

M. Defferrard et al. 2016. Convolutional neural networks on graphs with fast localized spectral filtering. NeurIPS.

T. Kipf et al. (2017). "Semi-Supervised Classification with Graph Convolutional Networks". ICLR.

W. L. Hamilton, et al. (2017). "Inductive Representation Learning on Large Graphs". NeurIPS.

P. Veličković et al. (2018). "Graph Attention Networks". In ICLR.

K. Xu et al. (2018). "Representation Learning on Graphs with Jumping Knowledge Networks". ICML

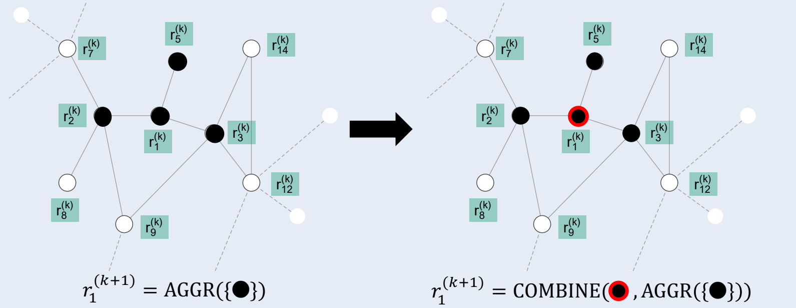
S. Abu-El-Haija et al. (2019). "MixHop: Higher-Order Graph Convolution Architectures via Sparsified Neighborhood Mixing". ICML

H. Pei et al. (2020). "Geom-GCN: Geometric Graph Convolutional Networks". ICLR.

Effective Designs for GNNs in Heterophily

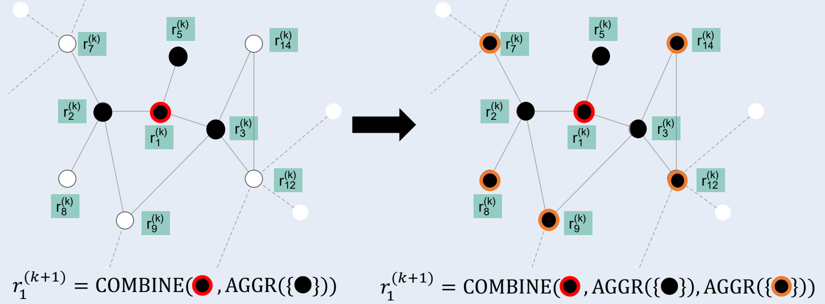
Design D1: Model the of ego- and neighbor-embeddings distinctly (per layer).

- In heterophily, neighbors may have information complementary to ego.



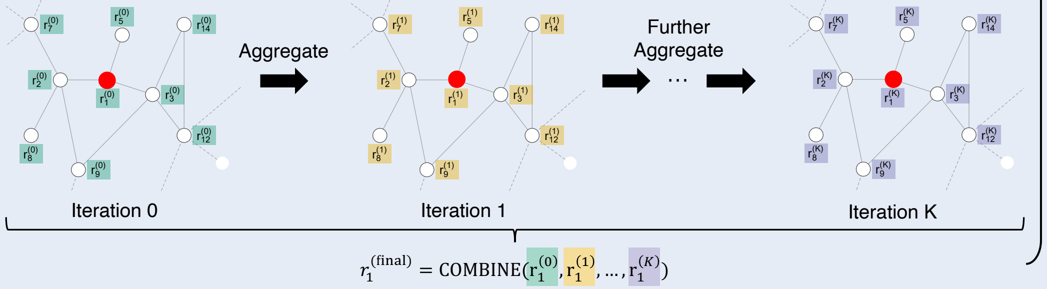
Design D2: Leverage representations of neighbors at different hops distinctly (per layer).

- Under heterophily, higher-order neighborhoods may still show homophily.



Design D3: Leverage the intermediate representations distinctly (at the final layer).

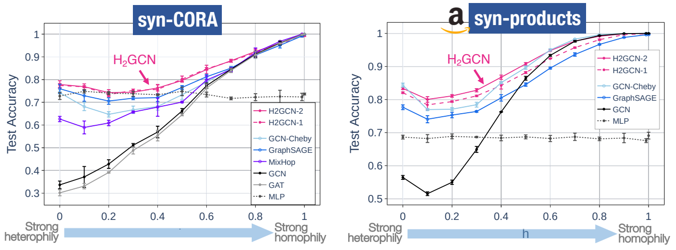
- Information with different locality contains different frequency components.



H2GCN

Empirical Analysis

Synthetic Benchmarks



Edge homophily ratio $h = \frac{\textit{intra-class edges}}{\textit{total edges}}$

- H_2GCN, our base model effectively combining all designs, has the best trend overall.
- Ablation study on H_2GCN shows effectiveness of each design, which results in up to 40% performance gain in heterophily.

Real Benchmarks

- In heterophily, models leveraging all or subsets of the designs perform significantly better than methods lacking them (e.g. GCN, GAT):
  - GraphSAGE (D1) vs. GCN: up to +23%
  - GCN-Cheby (D2) vs. GCN: up to +20%
  - GCN+JK (D3) vs. GCN: up to +14%

Method (Designs)	Average Rank Het.	Hom.	Overall
H_2GCN-1 (D1, D2, D3)	3.8	3.0	3.6
H_2GCN-2 (D1, D2, D3)	4.0	2.0	3.3
GraphSAGE (D1)	5.0	6.0	5.3
GCN-Cheby (D2)	7.0	6.3	6.8
MixHop (D2)	6.5	6.0	6.3
GraphSAGE+JK (D1, D3)	5.0	7.0	5.7
GCN-Cheby+JK (D2, D3)	3.7	7.7	5.0
GCN+JK (D3)	7.2	8.7	7.7
GCN	9.8	5.3	8.3
GAT	11.5	10.7	11.2
GEOM-GCN*	8.2	4.0	6.8
MLP	6.2	11.3	7.9

Detailed Results, Theorems & Code

GemsLab/H2GCN