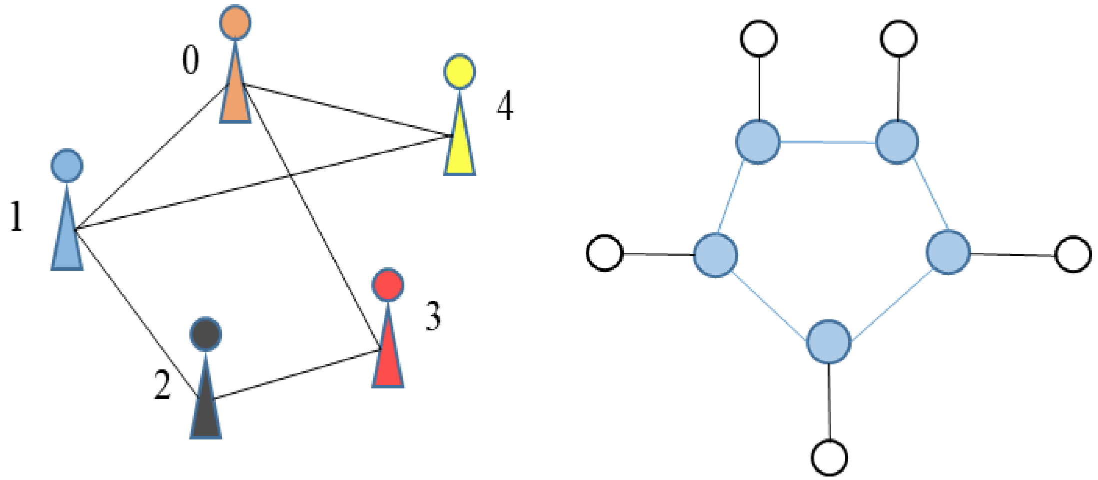
Unifying Homophily and Heterophily for Spectral Graph Neural Networks via Triple Filter Ensembles

Rui Duan, Mingjian Guang, Junli Wang, Chungang Yan, Hongda Qi, Wenkang Su, Can Tian, Haoran Yang



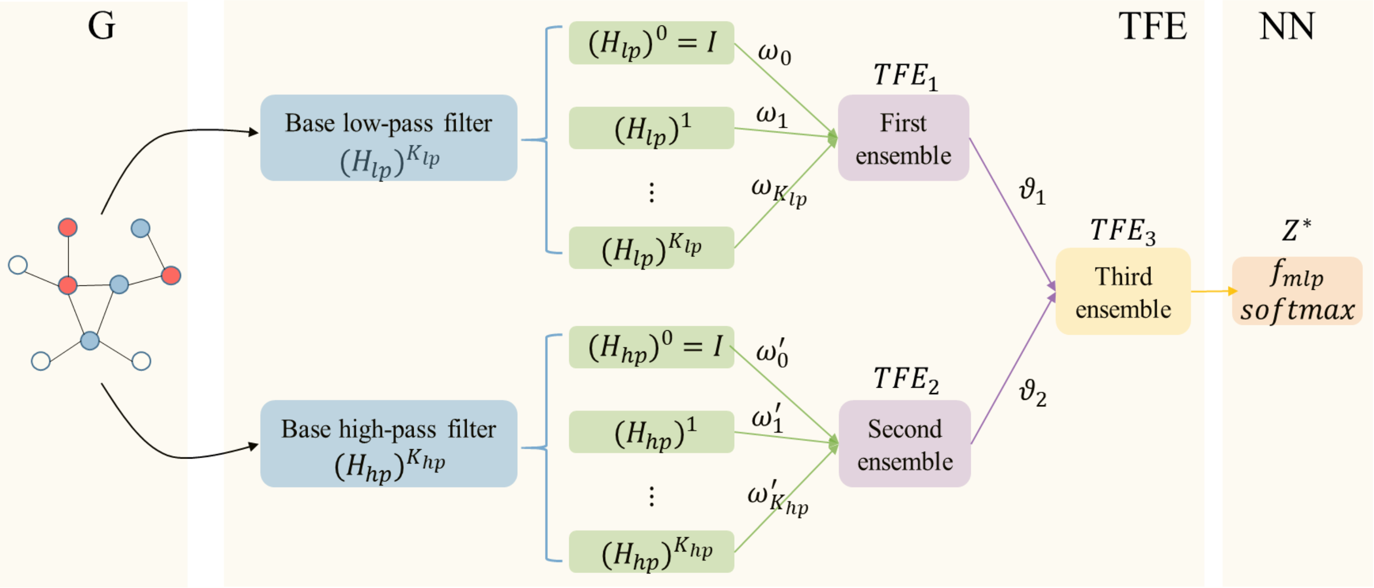
Polynomial-based learnable spectral graph neural networks

The source code of GEN is publicly available at https://github.com/graphNN/TFEGNN

Three Progressive Problems:

1. Some models use polynomials with better approximation for approximating filters, yet perform worse on real-world graphs.
2. Carefully crafted graph learning methods, sophisticated polynomial approximations, and refined coefficient constraints leaded to overfitting, which diminishes the generalization of the models.
3. How to design a model that retains the ability of polynomial-based spectral GNNs to approximate filters while it possesses higher generalization and performance?

TFE-GNN:



Motivations and Contributions:

Inspired by the following properties of ensemble learning: the strong classifier determined by the base classifiers can be more accurate than any of them if the base classifiers are accurate and diverse; and this strong classifier retains the characteristics of the base classifier to some extent. First, we combine a set of weak base low-pass f ilter to determine a strong low-pass filter that can extract homophily. Then, we use the same method to extract heterophily.

- We propose a spectral GNN with triple filter ensemble (TFE-GNN), which extracts ho mophily and heterophily from graphs with different levels of homophily adaptively while utilizing the initial features.
- The key difference between TFE-GNN and prior models is that TFE-GNN retains the ability of polynomial-based spectral GNNs while getting rid of polynomial computations, coefficient constraints, and specific scenarios.

Theoretical Analysis:

Theorem 1. *TFE-GNN and ChebNet can be transformed into each other under the following conditions, (1) learning the proper coefficients ω, ω', ϑ and ChebNet' coefficient θ, (2) the ensemble methods EM_1, EM_2 and EM_3 take* ***summation****, the base high-pass filter H_{hp} takes the symmetric normalized Laplacian L_{sym} and the base low-pass filter H_{lp} takes the affinity (transition) matrix of L_{sym}, and (3) $K_{hp} = K_{lp} = K - 1$. Thus, TFE-GNN can also learn arbitrary filters.*

Theorem 2. *TFE-GNN can be rewritten in the following form, with certain conditions to be satisfied, which is a combination of two polynomial-based learnable spectral GNNs: $Z^* = softmax(f_{mlp}(\vartheta_1 \sum_{k=0}^{K'} \bar{\theta}_k^1 P_k^1(\bar{H}_{gf}^1)^k X \bigoplus \vartheta_2 \sum_{k=0}^{K''} \bar{\theta}_k^2 P_k^2(\bar{H}_{gf}^2)^k X))$, where P_k denote polynomials used for approximation, $\bar{\theta}$ are the learnable coefficients, $\bar{H}_{gf}$ denote graph filters, and $\bigoplus$ denotes EM_3. Conditions are (1) learning the proper coefficients ω, ω', ϑ, $\bar{\theta}^1$, and $\bar{\theta}^2$, (2) the ensemble methods EM_1, EM_2 take* ***summation*** *and EM_3 takes ensemble method capable of preserving the properties of the model, such as* ***summation*** *and* ***concatenation****, the base high-pass filter H_{hp} takes $\bar{H}_{gf}^2$ and the base low-pass filter H_{lp} takes $\bar{H}_{gf}^1$, and (3) $K_{lp} = K'$ and $K_{hp} = K''$. Thus, TFE-GNN can match various graph structures adaptively.*

Experiments:

Results:

Datasets	Cora	Citeseer	Pubmed	CS	Physics
ehr	0.81	0.74	0.80	0.81	0.93
MLP	76.89±0.97	76.52±0.89	86.14±0.25	94.76±0.51	96.52±0.66
GCNs	87.18±1.12	79.85±0.78	86.79±0.31	93.11±0.19	96.66±0.74
ARMA	87.13±0.80	80.04±0.55	86.93±0.24	92.14±0.35	95.11±0.19
APPNP	88.16±0.74	80.47±0.73	88.13±0.33	92.61±0.28	95.81±0.11
ChebNet	87.32±0.92	79.33±0.57	87.82±0.24	91.63±0.39	94.21±0.26
GPR-GNN	88.54±0.67	80.13±0.84	88.46±0.31	95.67±0.16	96.80±0.08
BernNet	88.51±0.92	80.08±0.75	88.51±0.39	95.81±0.13	96.81±0.07
ChebNetII	88.71±0.93	80.53±0.79	88.93±0.29	96.03±0.11	97.23±0.07
SPECFORMER	88.57±1.01	81.49±0.94	89.13±0.35	95.92±0.19	97.44±0.08
EvenNet	87.25±1.42	78.65±0.96	89.52±0.31	94.66±0.23	95.59±0.11
Favard	89.35±1.09	81.89±0.63	90.90±0.27	95.77±0.15	97.58±0.08
PCNet	90.02±0.62	81.76±0.78	91.30±0.38	96.33±0.15	97.62±0.08
Half-Hop	88.73±1.22	80.33±0.66	89.86±0.36	95.13±0.21	95.75±0.13
GCNII	88.46±0.82	79.97±0.65	89.94±0.31	96.58±0.07	97.27±0.12
TWIRLS	88.57±0.91	80.07±0.94	88.87±0.43	95.43±0.04	97.17±0.07
PDE-GCN	88.62±1.03	79.98±0.97	89.92±0.38	95.35±0.19	96.89±0.08
EGNN	87.47±1.33	80.51±0.93	88.74±0.46	95.22±0.20	96.61±0.08
TFE-GNN$_{con}$	90.11±1.27	82.39±0.96	90.94±0.29	93.55±1.56	97.62±0.23
TFE-GNN$_{sum}$	**90.73±1.11**	**82.83±1.24**	**91.66±0.51**	**96.96±0.17**	**98.85±0.13**
TFE-GNN$\setminus TFE_1$	89.49±1.30	61.00±1.19	90.80±0.37	96.96±0.17	97.24±0.10
TFE-GNN$\setminus TFE_2$	90.15±1.75	82.83±1.24	91.66±0.51	96.57±0.16	97.19±0.19
TFE-GNN$_{rw+sum}$	89.57±1.26	81.92±1.14	90.96±0.49	95.70±0.53	97.73±0.16

Datasets	Cora-Full	Chameleon	Squirrel	Wisconsin	Texas	Cornell
ehr	0.57	0.23	0.22	0.21	0.06	0.30
MLP	52.45±0.64	46.59±1.84	31.01±1.18	86.55±2.36	86.81±2.24	84.15±3.05
GCNs	66.04±0.38	60.81±2.95	45.87±0.88	74.19±3.15	76.97±3.97	65.78±4.16
ARMA	63.53±0.66	60.21±1.00	36.27±0.62	87.25±1.63	83.97±3.77	85.62±2.13
APPNP	59.85±0.54	52.15±1.79	35.71±0.78	91.08±1.79	90.64±1.70	91.52±1.81
ChebNet	58.65±0.74	59.51±1.25	40.81±0.42	84.19±2.58	86.28±2.62	83.91±2.17
GPR-GNN	71.86±0.29	67.49±1.38	50.43±1.89	91.71±1.62	92.91±1.32	91.57±1.96
BernNet	72.01±0.26	68.53±1.68	51.39±0.92	92.45±1.22	92.62±1.37	92.13±1.64
ChebNetII	72.11±0.24	71.37±1.01	57.72±0.59	93.72±1.27	93.28±1.47	92.30±1.48
SPECFORMER	71.84±0.26	74.72±0.19	64.64±0.81	92.98±1.84	92.77±2.37	91.86±2.69
EvenNet	70.04±0.47	67.57±1.52	50.36±0.93	93.55±1.68	93.77±1.73	92.13±1.71
Favard	72.39±0.34	74.26±0.74	63.62±0.76	93.33±1.95	91.87±3.11	92.06±2.96
PCNet	72.35±0.26	73.55±1.26	63.53±0.26	94.26±1.85	92.78±1.80	93.83±1.91
Half-Hop	72.55±0.31	62.98±3.35	45.25±1.52	87.59±1.77	85.95±6.42	74.60±6.06
GCNII	66.70±0.85	63.44±0.85	41.96±1.02	85.66±1.95	80.46±5.91	84.26±2.13
TWIRLS	68.88±0.22	50.21±2.97	39.63±1.02	91.53±2.81	91.31±3.36	89.83±2.29
PDE-GCN	71.37±0.35	66.01±1.56	48.73±1.06	92.85±1.67	93.24±2.03	89.73±1.35
EGNN	71.51±0.27	51.55±1.73	35.81±0.91	83.76±1.64	81.34±1.56	82.09±1.16
TFE-GNN$_{con}$	**74.12±0.40**	77.03±1.47	71.47±1.15	96.00±1.73	93.11±4.26	**94.26±2.77**
TFE-GNN$_{sum}$	73.60±0.21	**77.16±1.41**	**72.27±1.32**	**97.38±1.42**	**94.87±2.66**	93.11±2.96
TFE-GNN$\backslash TFE_1$	64.09±0.41	76.63±2.20	57.06±1.03	92.97±1.38	93.44±1.80	93.11±2.96
TFE-GNN$\backslash TFE_2$	73.60±0.21	61.05±2.45	41.91±0.69	81.62±8.40	68.36±7.10	93.11±2.96
TFE-GNN$_{rw+sum}$	72.81±0.51	69.26±4.84	56.93±1.04	96.00±1.34	93.28±2.69	90.98±3.38

Generalization:

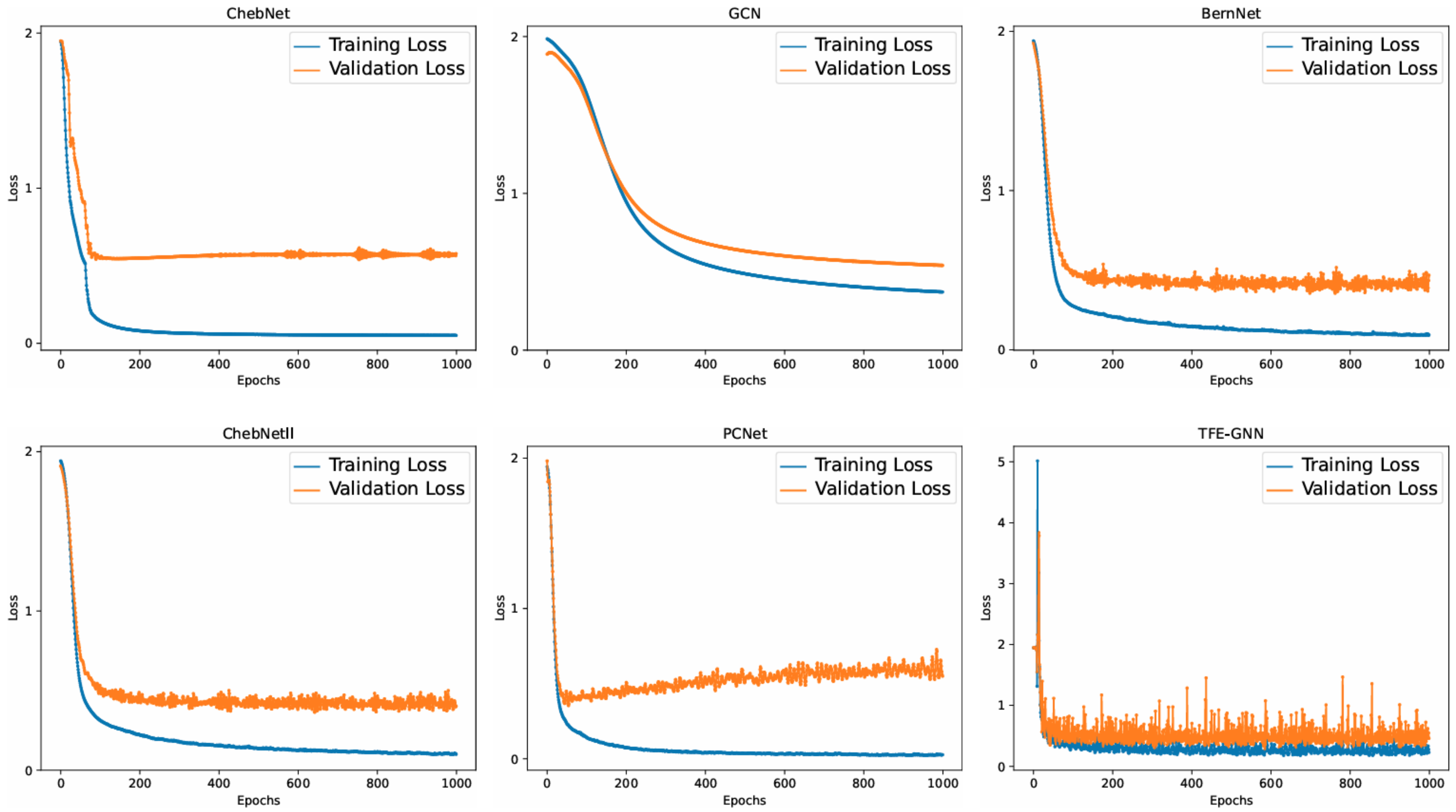


Figure 2: Generalization on Cora.

Thanks