Spectral GNNs based on eigen-decomposition

Lexin Zhang

School of Computing

Australian National University

Overview

My research problem includes two parts. How to get useful extra structural information from a graph and how to use extra structural information to enhance MPNNs. I propose two spectral features and four integration methods and most of them showed success in evaluations.

Background: MPNN

Convolutional neural networks achieve great success on grid-like data like images or texts. Message passing neural graph neural networks is a family of networks that extend CNNs to graph data.

Message passing graph neural network (MPNN) can iteratively aggregate information from direct neighbors to form new node representation. This process can be described in four steps.

- **Message:** messages for each edge in the graph are computed as information to be passed between connected nodes. $m_{u\to v} = Message(h_u, h_v, h_{e_{uv}})$
- **Aggregation:** for each node, messages of all neighbors are combined and aggregated. The aggregation methods are usually $m_u = Aggregate(m_{v\to u} | v \in N(u))$
- **Update:** for each node, aggregated message is combined with previous node feature to form the new node feature. $h_u = Updata(h_u, m_u)$
- **Readout:** the readout function can produce the final representation of the whole graph using all node representations. $h_G = Readout(h_u, u \in G)$

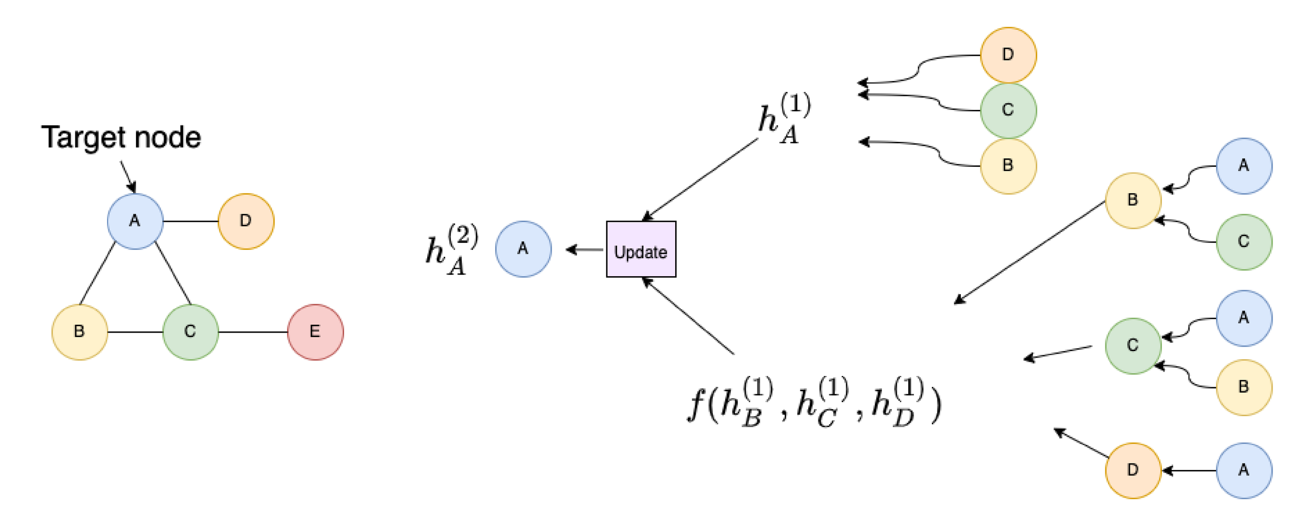

Figure 1. Illustration of a two layer MPNN

Motivation

The expressive power of any MPNN is limited by the first order Weisfeiler-Lehman test. When aggregating neighbors for node 1, the information that node 2 and node 3 have totally different neighborhood structure is not utilized.

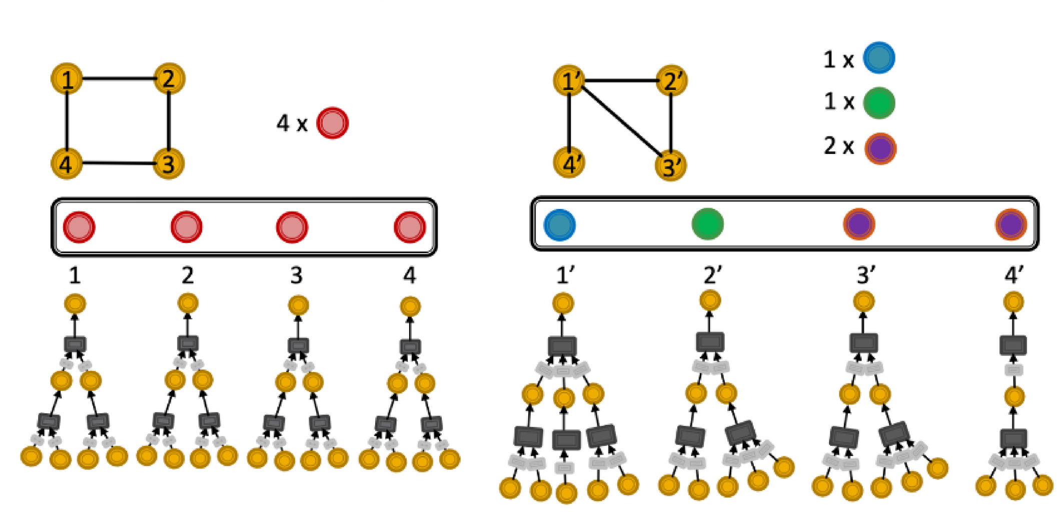

Figure 2. These two graphs are not distinguishable by MPNNs because they have the same computation tree. Image from [1]

How can we use extra structural information to enhance to expressive power of MPNNs?

Spectral feature

For a graph $G = (V, E)$, where $|V| = N$, we first calculate the adjacency spectrum and Laplacian spectrum of k-hop neighborhood of all nodes. For each type of spectrum, we have N vectors of different length, which contain the extra structural information that we're after, but the data is irregular and noisy.

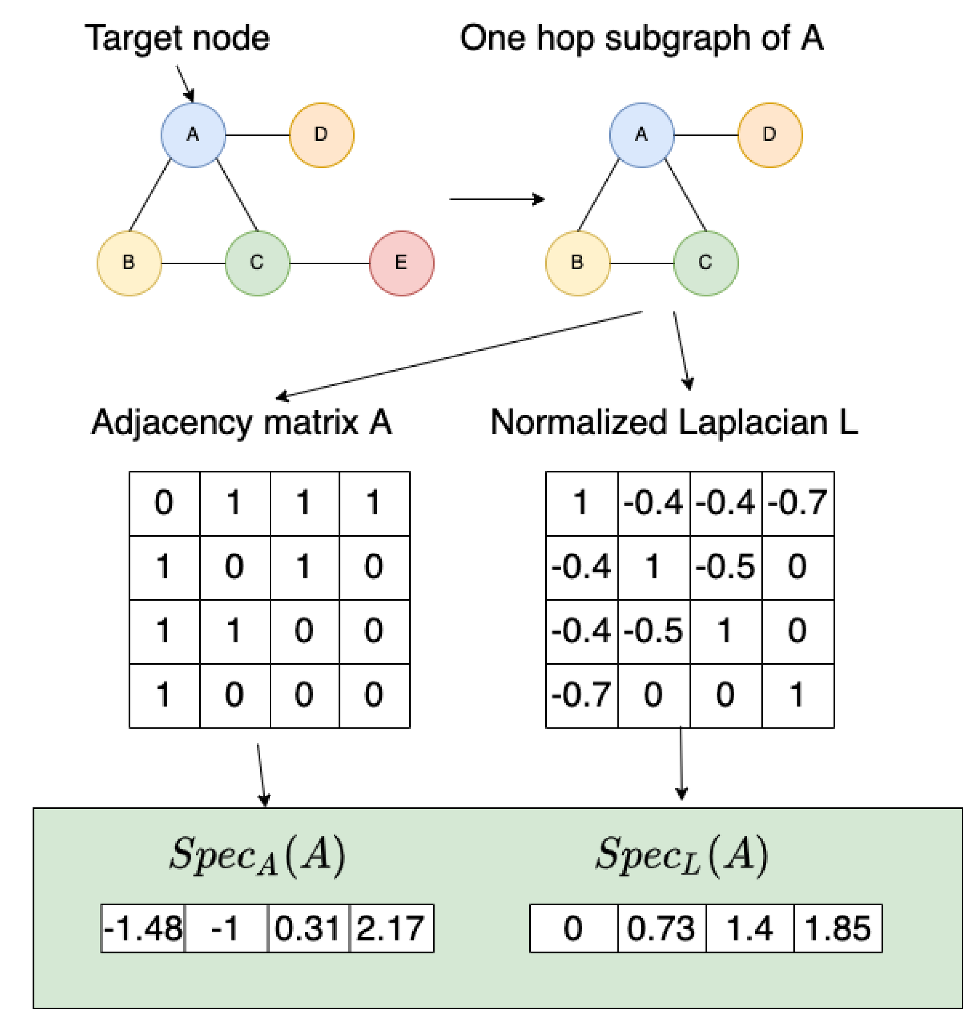

Figure 3. The (sorted) adjacency spectrum and Laplacian spectrum for 1-hop neighborhood of node A.

I propose two types of spectral features that can be obtained from this neighborhood spectra data.

- **Statistical spectral feature SSF:** for each vector(spectrum of a node), I use specific statistics to represent spectrum, including radius(largest spectrum), second-largest, multiplicity of 0, slope of spectrum between 0 and 1,... These statistics are well-studied in spectral graph theory to show strong relationship with property of graphs.
- **Distributional spectral feature DSF:** for each vector(spectrum of a node), given length l, I discretize the range of spectrum to equal connected intervals and then use the discrete empirical probability distribution to represent spectrum vector.

Integration into MPNNs-1

I propose four methods to incorporate spectral features into MPNNs for better expressive power.

Augmentation as node feature

The most intuitive way if to add SSF and DSF as node features before passing into MPNNs. Specifically, SSF and DSF are calculated and concatenated with original node features as a pre-processing step.

Augmentation as edge feature

SSF and DSF can also be computed based on subgraph of every edge in the graph, where we use the k-hop subgraph containing each edge. Then, SSF and DSF can be used as edge feature for MPNNs that support using multi-dimensional edge feature, for example GAT.

Integration into MPNNs-2

Parameterize aggregation

Basic aggregation functions treat all neighbors of a node equally. Here we propose spectral feature based aggregation. f' is either linear or three-layer MLP in evaluations.

$$h_u^{l+1} = f(h_u^l, AGG(\{h_v^l | v \in \mathcal{N}_u\}))$$

$$AGG(\{v_1, ..., v_n\}) = \sum c_i v_i$$

$$c_i \sim f'(DSF(v_i))$$

Adjust graph structure

It can be benefit for disconnected nodes with similar local structure can share information. We add edges between two nodes if the cosine similarity between their spectral features is above a certain threshold $0 < p < 1$

$$e_{ij} = 1, if \frac{SF(v_i) \cdot SF(v_j)}{\|SF(v_i)\| \|SF(v_j)\|} > p$$

Evaluation and results

Evaluations on combinations of spectral features and integration methods are conducted on the most famous MPNNs (GCN, GIN, GraphSAGE, GAT) for node classification task. Each result is averaged over 9 runs with random seeds 0 to 9.

	Integration-1		Integration-2		Integration-3	
Dataset	Best model	$\nearrow$	Best model	$\nearrow$	Best model	$\nearrow$
Cora	GCN-DSF2-A+L	1.79	GAT-DSF2-A+L	1.18	GCN-DSF2-A+L-MLP	-3.95
Citeseer	GAT-SSF2-A	2.81	GAT-DSF2-A	2.05	GIN-DSF1-A-MLP	-4.47
Pubmed	GCN-DSF2-L	0.60	GAT-DSF2-A	0.30	GIN-DSF1-A-MLP	4.05
Cornell	GAT-DSF2-A	5.68	GAT-SSF1-A+L	2.43	GGCN-SSF1-L-L	28.18
Texas	GraphSage-DSF1-A	2.97	GAT-SSF2-A+L	2.43	GIN-SSF1-L-MLP	18.79
Wisc.	GAT-DSF1-A+L	8.43	GAT-DSF2-L	7.65	GIN-DSF1-A+L-L	24.71

Table 1. Best variants and their improvement(or decrease) over baseline methods for the first three integration methods. GCN-DSF2-A+L means using DSF on ajacency spectrum and Laplacian spectrum of 2-hop neighborhood of each node.

For most variants, the test accuracy increased over baseline models. Only for integration-3, test accuracies for homophilic datasets decreased but test accuracies for heterophilic datasets increased.

Conclusion

1. I propose spectral features that contain meaningful local structural information that can not be captured by MPNNs. I explore four ways to incorporate spectral features into the working mechanism of MPNNs.
2. Through extensive experiments, I show the strength of proposed spectral features and integration methods and open the opportunity for utilizing local structural information from graph spectrum with more complexed techniques.

References

[1] Anil.
18 limitations of graph neural networks. Dec 2020.

Supervised by A/Prof. Qing Wang

u7239573@anu.edu.au