Explanatory Essay Requirements

Yushun Dong (yd24f@fsu.edu)

Please write a one-page explanatory essay about the research topic below. You will be able to complete this explanatory essay without digging into much technical details of related topics.

Background. In an online social network[1], a user account (e.g., a Twitter account) can be represented as a node, and the connections between users (e.g., following relationship and friend connections) can be considered as the topological structure over these nodes. For example, when you have collected lots of Twitter accounts and their following connections[2], you have a network. To take a step further, an attributed network refers to the network with the associated attribute information. For example, for your previously collected network, if you can further collect the information associated with each user and encode the information of each user as a vector, then you have an attributed network. We will focus on such a type of attributed network, where we have nodes, one attribute vector per node, and the topological structure between nodes. Such data is widely used by many graph learning models (e.g., Graph Neural Networks) for a plethora of learning tasks (e.g., node classification).

Let us assume that we have an attributed network (as introduced above), and we have a 10-dimensional attribute vector per node. Imagine the topology of this network is drawn on a piece of paper, and the darkness of the color for each node is determined by the value of the first dimension in its attribute vector. Then, you may see a snapshot of "darkness waves" across the network topology. Such a snapshot of darkness waves may be very messy, but it can be "decomposed" (*Graph Fourier Transform*) as a series of snapshots of waves with pre-defined appearances (*Graph Fourier Basis*) [1]. In these snapshots of waves with pre-defined appearances, the darkness levels of some snapshots of waves could fluctuate "quickly": even the darkness levels of neighboring nodes change a lot (high-frequency components); the darkness levels of other snapshots of waves could fluctuate "slowly": the darkness levels on lots of nodes that are topologically close to each other look similar (low-frequency components).

We now introduce a mainstream of Graph Neural Networks (GNNs), spectral GNNs [6], about how they typically process an attributed network and return predictions. Specifically, when spectral GNNs are given an attributed network, a simplified process for them is to (1) perform decomposition of the snapshots of waves; (2) "arbitrarily" throw away those snapshots of waves whose darkness fluctuates very quickly (*Low-Pass Filtering*) [5, 3]; (3) add the rest snapshots together as a snapshot that looks a bit different from the vanilla snapshot; (4) repeat the first three steps for every dimension of the attributes of nodes; and (5) use a deep learning model (e.g., a multilayer perceptron) to generate output. Notably, in this process, spectral GNNs only keep those snapshots whose darkness fluctuates slowly (connected nodes have similar color darkness levels), and they further make predictions based on the "summation" of these snapshots. Does this operation always help? Well, it depends.

Observation. Surprisingly, some researchers found that spectral GNNs seem to be very helpful in performing predictions in certain applications. For example, in a citation network where nodes are papers and connections represent the citation relationship, spectral GNNs are found to predict the multi-categorical research area of papers much more accurately compared with directly using step (5) in the previous paragraph to process node attributes (bag-of-word representations in a citation network) and obtain predictions [2]. This is a node-level task, i.e., we have one prediction (which categorical research area a paper belongs to) for each node. In network-level tasks, i.e., we have one prediction for each network[3], researchers also found spectral GNNs help a lot. For example, spectral GNNs are found to be helpful for the classification of molecular properties [4], where each molecule is considered as a network (with atoms being nodes and chemical bounds being connections).

Question. You will write a one-page essay to explain the observation above: why do spectral GNNs help a lot when they "arbitrarily" throw away those snapshots of waves whose darkness fluctuates very quickly? You are recommended to answer this question in two perspectives: *node-level tasks* (e.g., node classification) and *graph-level tasks* (e.g., graph classification). You will need to survey related literature to complete this essay (maybe start from the references below), and you are required to include necessary references to support key claims. You are not restricted from using tools based on generative AI but not recommended, since they constantly return inaccurate claims and fabricated references.

References

[1] Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

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

[3] Hoang Nt and Takanori Maehara. Revisiting graph neural networks: All we have is low-pass filters. *arXiv preprint arXiv:1905.09550*, 2019.

[4] Oliver Wieder, Stefan Kohlbacher, Mélaine Kuenemann, Arthur Garon, Pierre Ducrot, Thomas Seidel, and Thierry Langer. A compact review of molecular property prediction with graph neural networks. *Drug Discovery Today: Technologies*, 37:1–12, 2020.

[5] Felix Wu, Amauri Souza, Tianyi Zhang, Christopher Fifty, Tao Yu, and Kilian Weinberger. Simplifying graph convolutional networks. In *International conference on machine learning*, pages 6861–6871. PMLR, 2019.

[6] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and S Yu Philip. A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, 32(1):4–24, 2020.

[1] We use the word "network" and "graph" interchangeably.

[2] For simplicity, we consider all connections to be undirected, and two nodes can only have a maximum of one connection between them.

[3] Here, we still have one node attribute vector per node in the network.