View Reviews

Paper ID
2760

Paper Title
Breaking the Limit of Graph Neural Networks by Improving the Assortativity of Graphs with Local Mixing Patterns

Track Name
Research Track

Reviewer #2

Questions

1. What is the most accurate characterization of the paper
A novel perspective on an existing research topic of interest within data science

2. Is the paper properly contextualized, i.e., does the proposed work makes a non-trivial advance over relevant related work.
Agree

3. Please provide details on why you have selected the option to the above question
The conclusion from the analysis of the mixing pattern assortativity over GNN models can potentially benefit other research in the community.

4. Are the most important claims of the paper justified based on theoretical contribution(s) or experimental validation
Agree

5. Is the presentation of the paper (organization, writing style, grammar, formatting etc.) consistent with the high and well established SIGKDD standards
Strongly Agree

6. Please make a recommendation regarding acceptance of the paper
Accept

7. Please use this space to provide detailed feedback to the authors and justify your selection to the above questions. Please provide at least one paragraph of justification per question.
Summary:

This paper studies the relationships between the limits of prediction performance of GNN and different mixing patterns in a graph by using the assortativity metric. The authors show that the representation capability of a wide range of GNN models is highly correlated with the level of local assortative mixing in the graph, which sets a limit to the prediction performance for GNN models. They also proposed an algorithm to transform the input graph into a new one with a higher assortativity level and suitable for the deployment of GNN by leveraging both proximity and the local structural similarity of nodes at multiple scales. The experimental results validate the effectiveness of the proposed approach.

Strength：

- Using the assortativity to study the limits of prediction performance of GNNs is sound and interesting.
- The conclusion from the analysis of the mixing pattern assortativity over GNN models can potentially benefit other research in the community.
- The idea of transforming the input graph into a computation graph containing both proximity and structural

information is sound and novel.
- The experimental results validate the benefits of the proposed approach.

Weaknesses:

- The reason that the paper only considers semi-supervised node classification task need to be justified. In other words, the authors should explain why the proposed model shows better performance, particularly in the semi-supervised node classification task.
- The proposed WRGAT is a direct extension of multi-relational graphs from the GAT model. The WRGCN model seems to be equivalent to the R-GCN model [42]. From this point, the proposed approach is still lacking certain technical contributions. But I would suggest the authors test some other multi-relational GNNs with the computation graph, such as the Composition-GCN model (Composition-based Multi-Relational Graph Convolutional Networks, ICLR2020).
- Only the performances on the semi-supervised node classification are reported in this paper, so I might think that the proposed approach is still not able to outperform these baselines in other tasks, such as link prediction. However, regarding the semi-supervised node classification task, the labeled ratio for each dataset is unclear in this paper. Why the tested datasets were split under different ratios? In order to validate the effectiveness of the proposed model in the semi-supervised node classification task, I would like to see some performance comparison over different label ratios, and over some stronger baselines, such as the SCAN model (Semi-supervisedly Co-embedding Attributed Networks, NeurIPS2019).

Reviewer #3

Questions

1. What is the most accurate characterization of the paper
A new algorithm which improves the performance of an existing research problem

2. Is the paper properly contextualized, i.e., does the proposed work makes a non-trivial advance over relevant related work.
Agree

3. Please provide details on why you have selected the option to the above question
The authors proposed the node level assortativity to measure the level of mixing in the complex networks, which is related to KDD community in terms of modelling and understanding complex networks.

4. Are the most important claims of the paper justified based on theoretical contribution(s) or experimental validation
Agree

5. Is the presentation of the paper (organization, writing style, grammar, formatting etc.) consistent with the high and well established SIGKDD standards
Agree

6. Please make a recommendation regarding acceptance of the paper
Accept

7. Please use this space to provide detailed feedback to the authors and justify your selection to the above questions. Please provide at least one paragraph of justification per question.
The authors proposed node level assortativity to measure the level of mixing in the complex networks. Then they analysed the relations between the metric and performance of GNNs and found the limit of GNNs. To break the limit, they constructed a more assortative network by considering the structural patterns of the nodes' neighbours in a given network. Overall, this paper is well-written and organised. However, there are several problems in the

paper:

1. This is not an end-to-end model. The quality of the constructed network greatly affects the final result.

2. This study does not overcome the shortcomings of GNNs inherently. The GNNs are still only suitable for assortative networks. In a real scenario, we cannot know the type of real network, so we need to construct a new network for each one according to this study. It may be redundant and time-consuming for a network that is itself assortative.

3. The caption of Algorithm 2 does not match the content.

4. Geom-GCN is a baseline method, but it is not introduced in Section 6.1.

5. The authors did not compare the proposed method with Geom-GCN and H2GCN on the Air traffic and BGP dataset.

Reviewer #4

Questions

1. What is the most accurate characterization of the paper

A new algorithm which improves the performance of an existing research problem

2. Is the paper properly contextualized, i.e., does the proposed work makes a non-trivial advance over relevant related work.

Agree

3. Please provide details on why you have selected the option to the above question

This work creates a transformed computation graph encoding both graph structure and proximity information for each node in the heterogeneous graph. Specifically, heterogeneous graph mixing pattern information is injected into the computation graph with weighted edges added in the process. Various GNN models are then evaluated on this new graph.

4. Are the most important claims of the paper justified based on theoretical contribution(s) or experimental validation

Agree

5. Is the presentation of the paper (organization, writing style, grammar, formatting etc.) consistent with the high and well established SIGKDD standards

Agree

6. Please make a recommendation regarding acceptance of the paper

Accept

7. Please use this space to provide detailed feedback to the authors and justify your selection to the above questions. Please provide at least one paragraph of justification per question.

Overall, the paper is well-written. There is a good introduction to various heterogeneous graph mixing patterns e.g., disassortative mixing and assortative mixing and GNNs. The framework is novel.

One suggestion would be to more clearly explain how to determine a new link's weight in a computation graph if added due to assortativity.

Experiments need more baseline GNN models. Further, more discussion about hyperparameter settings and dataset statistics is required.

Reviewer #13

Questions

1. What is the most accurate characterization of the paper

A novel perspective on an existing research topic of interest within data science

2. Is the paper properly contextualized, i.e., does the proposed work makes a non-trivial advance over relevant related work.

Agree

3. Please provide details on why you have selected the option to the above question

This paper proposes a new metric of diversity in graphs that improves the classification performance of graph neural networks.

4. Are the most important claims of the paper justified based on theoretical contribution(s) or experimental validation

Agree

5. Is the presentation of the paper (organization, writing style, grammar, formatting etc.) consistent with the high and well established SIGKDD standards

Agree

6. Please make a recommendation regarding acceptance of the paper

Accept

7. Please use this space to provide detailed feedback to the authors and justify your selection to the above questions. Please provide at least one paragraph of justification per question.

This paper proposes an original new metrics of the diversity of graphs.

It carefully compares its approach to state-of-the-art competitors on several real-world datasets.

I appreciate the detailled analysis of the results and of the sensitivity to structure relations.

Moreover the paper is clearly written.