Multi-scale learning on graphs

Franz Rieger, Department of Informatics, Technical University of Munich, Munich, Germany

Abstract

This seminar paper presents several approaches for learning on graphs on different scales.

1. Introduction

Graphs are important: Spread of disease over population example from (Hammond et al., 2011),

learning on graphs is possible; in the recent years: GNNs successful (for an overview refer to (Zhang et al., 2020; Zhou et al., 2018; Wu et al., 2020; Ortega et al., 2018))

learning on different scales relevant

Learning: Different settings: Single graph (Semi-supervised node classification, finding fode embeddings) vs several (graph classification, graph embeddings); Link prediction; What else?

Scale:

- different sizes of graphs (molecules, citation networks, point clouds, 3D meshes, people in a social network, ...)
- hierarchy/clusters: nodes/ clusters of nodes; process information at different levels of abstraction on original graph

In the following, we present different approaches for learning on graphs on multiple scales. Many of these approaches are generalizations of or similar to existing methods on euclidean data (e.g. CNNs for images).

2. Notation

For later use we define the following notation: $G = (V, E, w, f)$, $|V| = n, E \subseteq V^2$, $w : E \to \mathbb{R}$, $f : V \to \mathbb{R}^m$

Adjacency Matrix: $A_{ij} = w((i, j))$

Laplacian $L = I - A$

Small intro to spectral graph theory: Eigenvalues of L (second smallest important)

Small intro to GNNs: $y(v) = activation(W_1 * x(v) + W_2 * aggregation_{winN(v)}(x(w)))$

Problem if repeated: oversmoothing

3. Spectral Methods

In this section, two methods which directly rely on spectral properties are presented.

3.1. Wavelets

Graph Wavelets (Hammond et al., 2011)

Older, Not learned, still useful

Convolution theorem:

Fourier Transformation:

Wavelets: Like Fourier Transform, just with another function instead of sin(x), have a scale

Defined for grid-like euclidean data (include formulas for this? Most papers do it), but can also be generalized for graphs:

Using spectral properties, Laplacian

Aim to preserve spectral properties (i.e. the rough structure can be reconstructed)

Put wavelets in GCN (Xu et al., 2019): node classification

3.2. Graph Scattering Transforms

Graph Scattering Transforms (Gama et al., 2019a;b) Wavelets used to show stability of output wrt. change in graphs (perturbations), scattering transforms robust, useful for transfer learning

Formula for definition:

4. Graph Convolution based Architectures

In this section, we present several modern architectures, which are based on GCNs. (todo: cite)

4.1. GraphZoom

GraphZoom (Deng et al., 2020)

generates node embeddings: scaleable + quality

figure: architecture overview

Steps: graph fusion -> coarsen graph -> find embeddings with existing method -> project + refine embeddings on original graph

4.2. DiffPool

Setting: Graph classification (+ link prediction)

figure: architecture overview

Several Layers, First: original graph, then vertices pooled together to clusters, clusters pooled to superclusters, superclusters pooled ...

In each layer: Embedding GCN (like SAGECONV, todo: cite) and seperate Assignment GCN Then: change embeddings, pool them (+edges) according to assignment GCN

Initially, small-scale neighbors, then bigger scale with each layer, making use of connections on different scales

Similar approach: Nodes merged according to edges: EdgePool (Diehl, 2019)

4.3. Multi-hop Hierarchical Graph Neural Networks

Multi-Hop (Xue et al., 2020) Concatenate Features from multiple message-passing hops Uses attention mechanisms, (weighiting nodes according to their predicted importance, as proved useful in (Lee et al., 2019) and in Computer Vision (Todo: citation) and Machine Translation (todo: citation))

4.4. Other mentions

Residual connections are useful for deeper GCNs (larger scale) (Li et al., 2019)

RecurrentGNNs (Ruiz et al., 2020) can be used for graphs modeling dynamic processes (like the initially mentioned spread of a novel disease)

Learning on small graphs (Simonovsky & Komodakis, 2018) vs large graphs (Hamilton et al., 2017)

One other way to prevent over-smoothing of deeper GCNs is using the idea of PageRank for GCNs (Klicpera et al., 2019).

The following are probably less relevant (only include if too much space left):

Simplified GNN architecture: (Wu et al., 2019)

Multiple Dimensions (nodes of different types) can be handled well (Ma et al., 2018; 2019)

5. Conclusion

We have seen several approaches, solving different learning tasks.

TODO: Update citations (remove arXiv preprints where possible)

References

Deng, C., Zhao, Z., Wang, Y., Zhang, Z., and Feng, Z. Graphzoom: A multi-level spectral approach for accurate and scalable graph embedding. In International Conference on Learning Representations, 2020. URL https://openreview.net/forum?id=r1lGO0EKDH.

Diehl, F. Edge contraction pooling for graph neural networks. arXiv preprint arXiv:1905.10990, 2019.

Gama, F., Estrach, J. B., and Ribeiro, A. Diffusion scattering transforms on graphs. In 7th International Conference on Learning Representations, ICLR 2019, 2019a.

Gama, F., Ribeiro, A., and Bruna, J. Stability of graph scattering transforms. In Advances in Neural Information Processing Systems, pp. 8036–8046, 2019b.

Hamilton, W., Ying, Z., and Leskovec, J. Inductive representation learning on large graphs. In Advances in neural information processing systems, pp. 1024–1034, 2017.

Hammond, D. K., Vandergheynst, P., and Gribonval, R. Wavelets on graphs via spectral graph theory. Applied and Computational Harmonic Analysis, 30 (2):129–150, 2011.

Klicpera, J., Bojchevski, A., and Günnemann, S. Combining neural networks with personalized pagerank for classification on graphs. In International Conference on Learning Representations, 2019. URL https://openreview.net/forum?id=H1gL-2A9Ym.

Lee, J., Lee, I., and Kang, J. Self-attention graph pooling. arXiv preprint arXiv:1904.08082, 2019.

Li, G., Müller, M., Thabet, A., and Ghanem, B. Deepgcns: Can gcns go as deep as cnns? In 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 9266–9275, 2019.

Ma, Y., Ren, Z., Jiang, Z., Tang, J., and Yin, D. Multi-dimensional network embedding with hierarchical structure. In Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, pp. 387–395, 2018.

Ma, Y., Wang, S., Aggarwal, C. C., Yin, D., and Tang, J. Multi-dimensional graph convolutional networks. In Proceedings of the 2019 SIAM International Conference on Data Mining, pp. 657–665. SIAM, 2019.

Ortega, A., Frossard, P., Kovačević, J., Moura, J. M., and Vandergheynst, P. Graph signal processing: Overview, challenges, and applications. Proceedings of the IEEE, 106(5):808–828, 2018.

Ruiz, L., Gama, F., and Ribeiro, A. Gated graph recurrent neural networks. arXiv preprint arXiv:2002.01038, 2020.

Simonovsky, M. and Komodakis, N. Graphvae: Towards generation of small graphs using variational autoencoders. In International Conference on Artificial Neural Networks, pp. 412–422. Springer, 2018.

Wu, F., Jr., A. H. S., Zhang, T., Fifty, C., Yu, T., and Weinberger, K. Q. Simplifying graph convolutional networks. In ICML, pp. 6861–6871, 2019. URL http://proceedings.mlr.press/v97/wu19e.html.

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Philip, S. Y. A comprehensive survey on graph neural networks. IEEE Transactions on Neural Networks and Learning Systems, 2020.

Xu, B., Shen, H., Cao, Q., Qiu, Y., and Cheng, X. Graph wavelet neural network. arXiv preprint arXiv:1904.07785, 2019.

Xue, H., Sun, X.-K., and Sun, W.-X. Multi-hop hierarchical graph neural networks. In 2020 IEEE International Conference on Big Data and Smart Computing (BigComp), pp. 82–89. IEEE, 2020.

Zhang, Z., Cui, P., and Zhu, W. Deep learning on graphs: A survey. IEEE Transactions on Knowledge and Data Engineering, 2020.

Zhou, J., Cui, G., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., and Sun, M. Graph neural networks: A review of methods and applications. arXiv preprint arXiv:1812.08434, 2018.

A. Appendix

Todo: Remove?