Ahmed, Ammar; Hassan, Zohair Raza; Shabbir, Mudassir

Interpretable multi-scale graph descriptors via structural compression. (English)

Zbl 1459.68165
Inf. Sci. 533, 169-180 (2020).

Summary: Graph representations that preserve relevant topological information allow the use of a rich machine learning toolset for data-driven network analytics. Some notable graph representations in the literature are fruitful in their respective applications but they either lack interpretability or are unable to effectively encode a graph's structure at both local and global scale. In this work, we propose the Higher-Order Structure Descriptor (HOSD): an interpretable graph descriptor that captures information about the patterns in a graph at multiple scales. Scaling is achieved using a novel graph compression technique that reveals successive higher-order structures. The proposed descriptor is invariant to node permutations due to its graph-theoretic nature. We analyze the HOSD algorithm for time complexity and also prove the NP-completeness of three interesting graph compression problems. A faster version, HOSD-Lite, is also presented to approximate HOSD on dense graphs. We showcase the interpretability of our model by discussing structural patterns found within real-world datasets using HOSD. HOSD and HOSD-Lite are evaluated on benchmark datasets for applicability to classification problems; results demonstrate that a simple random forest setup based on our representations competes well with the current state-of-the-art graph embeddings.

MSC:

68T05 Learning and adaptive systems in artificial intelligence
68P30 Coding and information theory (compaction, compression, models of communication, encoding schemes, etc.) (aspects in computer science)
68R10 Graph theory (including graph drawing) in computer science
68T10 Pattern recognition, speech recognition

Keywords:

graph embeddings; graph compression; graph classification

Software:

NetLSD; AFGen

Full Text: DOI

References:

[1] R. Angles, C. Gutierrez, Survey of graph database models, ACM Computing Surveys (CSUR) 40(1) (2008) 1.

[2] Goyal, P.; Ferrara, E., Graph embedding techniques, applications, and performance: a survey, Knowl.-Based Syst., 151, 78-94 (2018)

[3] Cai, H.; Zheng, V. W.; Chang, K. C.-C., A comprehensive survey of graph embedding: problems, techniques, and applications, IEEE Trans. Knowl. Data Eng., 30, 9, 1616-1637 (2018)

[4] Zhang, M.; Cui, Z.; Neumann, M.; Chen, Y., An end-to-end deep learning architecture for graph classification, (Thirty-Second AAAI Conference on Artificial Intelligence (2018))

[5] R. Kondor, H. Pan, The multiscale laplacian graph kernel, in: Advances in Neural Information Processing Systems, 2016, pp. 2990-2998.

[6] N. Shervashidze, S. Vishwanathan, T. Petri, K. Mehlhorn, K. Borgwardt, Efficient graphlet kernels for large graph comparison, in: Artificial Intelligence and Statistics, 2009, pp. 488-495.

[7] Beg, M. A.; Ahmad, M.; Zaman, A.; Khan, I., Scalable approximation algorithm for graph summarization, (Pacific-Asia Conference on Knowledge Discovery and Data Mining (2018), Springer), 502-514

[8] Borgwardt, K. M.; Kriegel, H.-P., Shortest-path kernels on graphs, (Fifth IEEE International Conference on Data Mining (ICDM'05) (2005), IEEE), 8

[9] N. Shervashidze, P. Schweitzer, E.J.v. Leeuwen, K. Mehlhorn, K.M. Borgwardt, Weisfeiler-lehman graph kernels, J. Mach. Learn. Res. 12 (2011) 2539-2561. · Zbl 1280.68194

[10] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, P.S. Yu, A comprehensive survey on graph neural networks, arXiv preprint arXiv:1901.00596.

[11] Niepert, M.; Ahmed, M.; Kutzkov, K., Learning convolutional neural networks for graphs, (International Conference on Machine Learning (2016)), 2014-2023

[12] D.V. Tran, N. Navarin, A. Sperduti, On filter size in graph convolutional networks, in: 2018 IEEE Symposium Series on Computational Intelligence (SSCI), IEEE, 2018, pp. 1534-1541.

[13] K. Xu, W. Hu, J. Leskovec, S. Jegelka, How powerful are graph neural networks?, 2018.

[14] C. Morris, M. Ritzert, M. Fey, W.L. Hamilton, J.E. Lenssen, G. Rattan, M. Grohe, Weisfeiler and leman go neural: Higher-order graph neural networks, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, 2019, pp. 4602-4609.

[15] Micheli, A., Neural network for graphs: a contextual constructive approach, IEEE Trans. Neural Networks, 20, 3, 498-511 (2009)

[16] Pope, P. E.; Kolouri, S.; Rostami, M.; Martin, C. E.; Hoffmann, H., Explainability methods for graph convolutional neural networks, (Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2019)), 10772-10781

[17] Berlingerio, M.; Koutra, D.; Eliassi-Rad, T.; Faloutsos, C., Network similarity via multiple social theories, (Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (2013), ACM), 1439-1440

[18] A. Dutta, H. Sahbi, Stochastic graphlet embedding, IEEE Transactions on Neural Networks and Learning Systems.

[19] Tsitsulin, A.; Mottin, D.; Karras, P.; Bronstein, A.; Müller, E., Netlsd: hearing the shape of a graph, (Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery \& Data Mining (2018), ACM), 2347-2356

[20] S. Verma, Z.-L. Zhang, Hunt for the unique, stable, sparse and fast feature learning on graphs, in: Advances in Neural Information Processing Systems, 2017, pp. 88-98.

[21] Golumbic, M. C.; Hartman, I. B.-A., Graph Theory, Combinatorics and Algorithms: Interdisciplinary Applications, vol. 34 (2006), Springer Science \& Business Media

[22] Uehara, R.; Toda, S.; Nagoya, T., Graph isomorphism completeness for chordal bipartite graphs and strongly chordal graphs, Discrete Appl. Math., 145, 3, 479-482 (2005) · Zbl 1056.05099

[23] P. Yanardag, S. Vishwanathan, Deep graph kernels, in: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2015, pp. 1365-1374.

[24] Wale, N.; Watson, I. A.; Karypis, G., Comparison of descriptor spaces for chemical compound retrieval and classification, Knowl. Inf. Syst., 14, 3, 347-375 (2008)

[25] Helma, C.; King, R. D.; Kramer, S.; Srinivasan, A., The predictive toxicology challenge 2000-2001, Bioinformatics, 17, 1, 107-108 (2001)