Mixed-Integer Optimisation of Graph Neural Networks for Computer-Aided Molecular Design (Article Abstract)*

Tom McDonald[1], Calvin Tsay[2], Artur M. Schweidtmann[1], and Neil Yorke-Smith[1]

[1] Delft University of Technology, The Netherlands
[2] Imperial College London, United Kingdom
n.yorke-smith@tudelft.nl

Abstract. The full article published in Computers and Chemical Engineering proposes a bilinear formulation for Graph Convolutional Networks, a form of Graph Neural Network (GNN), and a mixed integer linear programming (MILP) formulation for GraphSAGE GNNs. These formulations enable solving optimisation problems with trained GNNs embedded to global optimality. The article applies the novel optimisation approach to an illustrative computer-aided molecular design (CAMD) case study where the formulations of the trained GNNs are used to design molecules with optimal boiling points.

1 Motivation

The full article by McDonald et al. [6] recognises that the attraction of using GNN in CAMD is that molecules can naturally be represented as graphs. Every atom in a molecule is represented by a node, and the properties of this molecule are stored in the feature vectors associated with the atom-representing nodes. Moreover, GNNs preserve invariance of the graph structure, e.g., rotating a molecule does not affect the prediction. There have been multiple studies where GNNs have been used to predict properties of molecules [9].

There has been progress on exact formulations of non-graph neural networks as mixed integer (non-)linear programs [1, 2, 7, 8]. The work reported was to our knowledge the first MILP formulation of a trained GNN presented in the literature. The importance of such a model is that MILP formulations for GNNs can be used in CAMD, where properties of molecules can be modelled using GNNs and then optimised using MILP formulations of these trained GNNs. Recent work by Zhang et al. [10] developed symmetry-breaking constraints that can reduce the search space for MILP or other optimisation strategies. Furthermore, there are broader applications, namely the use of MILP formulations of GNNs for similar applications as MLPs [4], e.g., verification of GNNs, lossless compression of GNNs, and using GNNs as surrogate models in optimisation problems.

* This is an extended abstract of McDonald et al. [6].

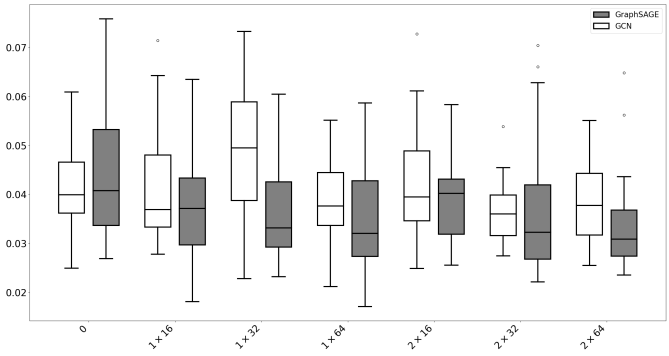

Fig. 1: Box-plots of the MSE of the validation data, for the GCN and GraphSAGE models for different layer depths and node width, after independently running the models 20 times for each configuration. The box-plots indicate the median, the lower and upper quartile and the lowest and largest MSE.

The paper considers two GNN architectures. The first is the Graph Convolutional Neural Network by Kipf and Welling [5]. This neural network is one of the earliest GNN and is used often in GNN applications. The second is the GraphSAGE network by Hamilton et al. [3], which learns properties of large graph data by sampling the neighbourhood of nodes instead of using information of all neighbouring nodes.

2 Results

Empirical evidence obtained suggests that the MILP formulation of the GraphSAGE model solves to optimality faster than the MIQCP formulation of the GCN model, with similar model accuracy. This is because the trained model accuracy is about the same and sometimes better for the GraphSAGE model compared to the GCN model, for models with similar hidden layers and number of nodes, combined with the fact that the GraphSAGE model often solves to optimality faster with similar configurations.

Both MI(N)LP formulations of trained GNNs were applied to a case study of optimising the boiling points of molecules. The case study successfully derived a set of optimal molecules, given constraints on the design space. Of the 20 molecules derived, 12 were found were experimentally observed. Of the other eight, the literature notes two as hypothetical molecules. These were able to be synthesised under very high pressure or were an unstable molecule of molecular reaction. The remaining six molecules appear to be novel; their chemical feasibility in practice would be tested in vitro studies.

Acknowledgements Thanks to the C&CE reviewers. This research was partially supported EU Horizon 2020 programme grant number 952215 (TAILOR).

References

[1] Anderson, R., Huchette, J., Ma, W., Tjandraatmadja, C., Vielma, J.P.: Strong mixed-integer programming formulations for trained neural networks. Mathematical Programming **183**(1), 3–39 (2020)

[2] Fischetti, M., Jo, J.: Deep neural networks and mixed integer linear optimization. Constraints **23**(3), 296–309 (2018)

[3] Hamilton, W., Ying, Z., Leskovec, J.: Inductive representation learning on large graphs. Advances in Neural Information Processing Systems **30** (2017)

[4] Huchette, J., Muñoz, G., Serra, T., Tsay, C.: When deep learning meets polyhedral theory: A survey. arXiv preprint arXiv:2305.00241 (2023)

[5] Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. In: Proc. of 5th International Conference on Learning Representations (ICLR'17) (2017), URL https://openreview.net/forum?id=SJU4ayYgl

[6] McDonald, T., Tsay, C., Schweidtmann, A.M., Yorke-Smith, N.: Mixed-integer optimisation of graph neural networks for computer-aided molecular design. Comput. Chem. Eng. **185**, 108660 (2024), https://doi.org/10.1016/J.COMPCHEMENG.2024.108660

[7] Thorbjarnarson, T., Yorke-Smith, N.: Optimal training of integer-valued neural networks with mixed integer programming. PLoS ONE **18**(2), e0261029 (2023)

[8] Tsay, C., Kronqvist, J., Thebelt, A., Misener, R.: Partition-based formulations for mixed-integer optimization of trained ReLU neural networks. Advances in Neural Information Processing Systems **34**, 3068–3080 (2021)

[9] Wieder, O., Kohlbacher, S., Kuenemann, M., Garon, A., Ducrot, P., Seidel, T., Langer, T.: A compact review of molecular property prediction with graph neural networks. Drug Discovery Today: Technologies **37**, 1–12 (2020)

[10] Zhang, S., Campos, J.S., Feldmann, C., Walz, D., Sandfort, F., Mathea, M., Tsay, C., Misener, R.: Optimizing over trained GNNs via symmetry breaking. In: Proc. of NeurIPS'23 (2023), URL http://papers.nips.cc/paper_files/paper/2023/hash/8c8cd1b78cdae751265c88efc136e5bd-Abstract-Conference.html