Deep Learning Attention Model For Supervised and Unsupervised Network Community Detection

Stanislav Sobolevsky[1][0000−0001−6281−0656]

Center for Urban Science and Progress, New York University, Brooklyn, NY, USA
Department of Mathematics and Statistics and Institute of Law and Technology, Masaryk University, Brno, Czech Republic
sobolevsky@nyu.edu

Abstract. Network community detection is a complex problem that has to utilize heuristic approaches. It often relies on optimizing partition quality functions, such as modularity, description length, stochastic block-model likelihood etc. However, direct application of the traditional optimization methods has limited efficiency in finding the global maxima in such tasks. This paper proposes a novel bi-partite attention graph neural network model for supervised and unsupervised network community detection, suitable for unsupervised optimization of arbitrary partition quality functions, as well as for minimization of a loss function against the provided partition in a supervised setting. The model is demonstrated to be helpful in the unsupervised improvement of suboptimal partitions previously obtained by other known methods like Louvain algorithm for some of the classic and synthetic networks. It is also shown to be efficient in supervised learning of the provided community structure for a set of classic and synthetic networks. Furthermore, the paper serves as a proof-of-concept for the broader application of graph neural network models to unsupervised network optimization.

Keywords: Complex networks · Community detection · Deep Learning · Graph Neural Networks

1 Introduction

The network community detection saw a wide range of applications, including social science [36], biology[21], and economics [35]. In particular, partitioning the networks of human mobility and interactions is broadly applied to regional delineation [37], [7], [44], [1], [24], [43], [3], [20], and urban zoning [42], [28], [27].

Community detection is a complex problem, and multiple algorithms have been proposed to address it. Some of them are straightforward, such as hierarchical clustering [23] or the Girvan-Newman [16] algorithm, while the majority rely on optimization techniques for various objective functions. The most well-known partition quality function is modularity [34, 33] assessing the relative strength of edges and quantifying the cumulative strength of the intra-community links. A large number of modularity optimization strategies have been suggested over the

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_51

last two decades [31], [12], [34], [33], [45], [8], [22], [19], [13], [29], [41]. Comprehensive overviews are presented in [14], [15] and later surveys [26], [25].

As the problem is known to be NP-hard [9]), there is no efficient algorithm that could guarantee finding the optimal partition. Therefore, optimization has to rely on heuristic algorithms, which often fail to reach the optimal partition, and, therefore, may require further fine-tuning. Although in some cases, an algorithmic optimality proof of the partition is possible [40], [5], [6], [4].

The rise of deep learning and graph neural networks in particular, offer new opportunities. Recently graph neural networks (GNNs) have become increasingly popular for supervised classifications and unsupervised embedding of the graph nodes with diverse applications in text classification, recommendation systems, traffic prediction, computer vision etc [47]. And GNNs were already attempted to be applied for community detection, including supervised learning of the ground-truth community structure [11] as well as some unsupervised learning of the node features enabling representation modeling of the network, including stochastic blockmodel [10] and other probabilistic models with overlapping communities [38] or more complex self-expressive representation [2]. However, existing GNN applications for unsupervised community detection has been limited so far, and largely overlook unsupervised modularity optimization.

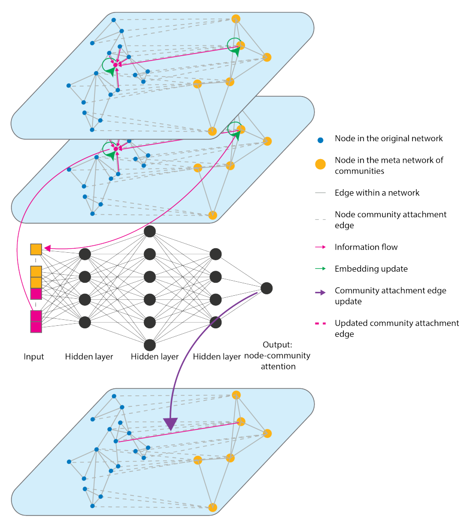

Fig. 1. A deep learning framework for network community detection.

This work proposes a suitable network augmentation with an additional layer of community meta-nodes, and a novel deep learning model over such a network for supervised and unsupervised community detection, capable of optimizing arbitrary quality functions for the network partition, including modularity.

2 Methodology

In our recent paper [39], we proposed a simple recurrent GNN-inspired algorithm to serve as a proof-of-concept for unsupervised modularity optimization. The algorithm tunes node community attachment through an iteration of the GNN-style transformations. It reached the best-known partitions for some of the

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_51

classical networks, and provided a scheme for fine-tuning the network community structure with a flexible trade-off between quality and speed.

Table 1. Performance on the proposed deep learning algorithm improving partition modularity from Louvain algorithm for some classic and synthetic network examples

Network	Louvain	Improvement
Word adjacency network in David Copperfield [32]	0.305052	0.309279
Amazon co-purchases of political books, orgnet.com	0.527082	0.527236
LFR 500 nodes	0.837450	0.837501
LFR 1000 nodes	0.888029	0.888615
LFR 2000 nodes	0.901624	0.902277

In order to further improve the efficiency of the approach, we propose a new model which consists of a two-layer bi-partite convolutional graph neural network stacked with a fully connected attention vanilla neural network 1. The model takes certain initial network node embedding as the input, such as the personalized Pagerank probability vectors (for each source node defining the stationary probability distribution of a Markov chain that, with probability $\alpha = 0.15$, randomly transitions following the link structure of the network, and with a probability $1 - \alpha$ teleports to a source node [17]), further reduced in dimensionality using a linear principal component method. The edges between the network nodes and community meta-nodes are initially defined with respect to a certain initial node community attachment (either a known one to be improved if used for partition fine-tuning, or based on clustering initial node embedding), and can be further updated as the model updates the node community attachment. The two layers of a graph neural network propagating the initial node embedding over such a bi-partite network generates the final embedding for both - the original network nodes as well as the community meta-nodes. Finally, the vanilla neural network (in the experiments below, the configuration with five hidden layers and batch normalization has been evaluated) for each pair of the original network node and the community meta-node takes the stacked vectors of those node embedding, generated by the graph neural network, and computes the relative attention score between the two nodes. The resulting node community attachment (a "fuzzy" probabilistic one rather than discrete) is redefined proportional to those attention scores. The weights of both - the final vanilla neural network as well as the graph neural network layers are trained together, within the backpropagation framework, aiming to optimize the final objective function - either the quality function of the resulting network partition in the unsupervised setting (like modularity or stochastic block-model likelihood) or the loss function between this resulting partition and the known partition in the supervised setting (like categorical cross-entropy). In order to improve the convergence stability and final performance of the model in the unsupervised set-

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_51

ting, it can be initially pre-trained to reconstruct a previously known community structure prior to the final unsupervised fine-tuning phase.

3 Results

The approach turned out to be efficient in fine-tuning the results of other algorithms, e.g. a popular Louvain algorithm [8]. The table 1 provides examples of such improvement reached by the Python 3.7 implementation of the proposed model for several classic and Lancichinetti-Fortunato-Radicchi (LFR) synthetic networks (table 2). And while for the two classic networks - Amazon political books and Word Adjacency network in David Copperfield - other known efficient algorithms like Combo were also capable of improving the partition, for the three provided cases of LFR networks, the fine-tuned partition, provided by the proposed deep learning algorithm, is the best partition known to us.

Table 2. Out-of-sample performance on the proposed deep learning algorithm in supervised learning of the best-known or given partition for some classic network examples (community reconstruction accuracy for the 40% randomly masked nodes) in comparison with the label propagation baseline

Network	Accuracy	Baseline
Amazon.com co-purchases of political books, www.orgnet.com	97.22%	94.44%
Dolphins' Social Network [30]	95.00%	95.00%
Network of Jazz Musicians [18]	93.10%	88.51%
Neural network of C. Elegans [46]	86.96%	75.65%
Metabolic network of C. Elegans [13]	64.33%	57.31%

The proposed approach can also be applied to other quality functions, such as a block-model likelihood or description length.

Furthermore, the model can perform supervised community detection, extrapolating the community structure provided for a certain part of the network nodes to the rest of them. The out-of-sample reconstruction accuracy for the best-known partition often ranges within 90-99% for a number of classic (table 2) and LFR synthetic networks (table 3).

For comparison, the out-of-sample accuracy of the label propagation baseline algorithm (nodes with unknown community attachments get attached according to the majority of their neighbors with known attachments) for most of the provided networks falls noticeably short of the accuracy achieved by the proposed deep learning approach.

Those cases represent initial proof-of-concept results, while fine-tuning of the model's configuration could further help improve the performance. Also, evaluation of the approach on a broader range of examples and comparison against

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_51

known state-of-the-art/baseline supervised community detection approaches remains the subject of future work.

Table 3. Out-of-sample performance on the proposed deep learning algorithm in supervised learning of the best-known or given partition for LFR synthetic networks (community reconstruction accuracy for the 40% randomly masked nodes) in comparison with the label propagation baseline

Network	Size	Accuracy	Baseline
1	500	93.65%	91.53%
2	500	99.47%	93.65%
3	500	94.18%	91.53%
4	500	95.24%	89.42%
5	500	98.94%	96.30%
6	1000	98.61%	93.98%
7	1000	96.99%	91.67%
8	1000	95.83%	89.35%
Average		96.61± 2.23%	92.18± 2.36%

Finally, as the deep learning model configuration does not depend on the dimensionality of the network or the number of network communities but only on the selected dimensionality of the node embedding, it makes it possible to consider transferring the pre-trained model architectures and parameters between the networks. And similarly to [39], iterating an ensemble of partition fine-tuning models (pre-trained over select sample networks) over the target network partition may provide the best practical results.

4 Conclusions

To summarize, the novel bi-partite attention graph neural network has been proposed for supervised and unsupervised network community detection. The model augments the original network with the meta-nodes representing the network communities and learns the node embedding as well as the relevance links between the two types of network nodes.

It was proven useful for the supervised reconstruction of the network community structure for both - classic and synthetic networks, consistently outperforming a baseline label propagation algorithm. In an unsupervised setting, we found the model helpful in fine-tuning the suboptimal network partitions obtained for some of the classic and synthetic networks by other known community detection algorithms, like Louvain.

While the presented results serve as a proof of concept of the proposed deep learning model's utility for supervised and unsupervised community detection, its further fine-tuning and extensive evaluation, as well as exploring the potential of transfer learning between the networks, is the subject of our future research.

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_51

Acknowledgements This research was supported by the MUNI Award in Science and Humanities (MASH Belarus) of the Grant Agency of Masaryk University under the Digital City project (MUNI/J/0008/2021). The work of Stanislav Sobolevsky was also partially supported by ERDF "CyberSecurity, CyberCrime and Critical Information Infrastructures Center of Excellence" (No. CZ.02.1.01/0.0/0.0/16_019/0000822). The author thanks Mingyi He for valuable help with visualization, and Dr. Alexander Belyi for insightful discussions.

References

1. Amini, A., Kung, K., Kang, C., Sobolevsky, S., Ratti, C.: The impact of social segregation on human mobility in developing and industrialized regions. EPJ Data Science **3**(1), 6 (2014)
2. Bandyopadhyay, S., Peter, V.: Self-expressive graph neural network for unsupervised community detection. arXiv preprint arXiv:2011.14078 (2020)
3. Belyi, A., Bojic, I., Sobolevsky, S., Sitko, I., Hawelka, B., Rudikova, L., Kurbatski, A., Ratti, C.: Global multi-layer network of human mobility. International Journal of Geographical Information Science **31**(7), 1381–1402 (2017)
4. Belyi, A., Sobolevsky, S.: Network size reduction preserving optimal modularity and clique partition. In: Gervasi, O., Murgante, B., Hendrix, E.M.T., Taniar, D., Apduhan, B.O. (eds.) Computational Science and Its Applications – ICCSA 2022. pp. 19–33. Springer International Publishing, Cham (2022). https://doi.org/10.1007/978-3-031-10522-7_2
5. Belyi, A., Sobolevsky, S., Kurbatski, A., Ratti, C.: Improved upper bounds in clique partitioning problem. J. Belarusian State Univ. Math. Informatics **2019**(3), 93–104 (2019). https://doi.org/10.33581/2520-6508-2019-3-93-104
6. Belyi, A., Sobolevsky, S., Kurbatski, A., Ratti, C.: Subnetwork constraints for tighter upper bounds and exact solution of the clique partitioning problem. arXiv preprint arXiv:2110.05627 (2021)
7. Blondel, V., Krings, G., Thomas, I.: Regions and borders of mobile telephony in belgium and in the brussels metropolitan zone. Brussels Studies. La revue scientifique électronique pour les recherches sur Bruxelles/Het elektronisch wetenschappelijk tijdschrift voor onderzoek over Brussel/The e-journal for academic research on Brussels (2010)
8. Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment **2008**(10), P10008 (2008)
9. Brandes, U., Delling, D., Gaertler, M., Görke, R., Hoefer, M., Nikoloski, Z., Wagner, D.: Maximizing modularity is hard. arXiv preprint physics/0608255 (2006)
10. Bruna, J., Li, X.: Community detection with graph neural networks. stat **1050**, 27 (2017)
11. Chen, Z., Li, X., Bruna, J.: Supervised community detection with line graph neural networks. arXiv preprint arXiv:1705.08415 (2017)
12. Clauset, A., Newman, M.E.J., Moore, C.: Finding community structure in very large networks. Phys. Rev. E **70**, 066111 (Dec 2004). https://doi.org/10.1103/PhysRevE.70.066111, http://link.aps.org/doi/10.1103/PhysRevE.70.066111

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_51

13. Duch, J., Arenas, A.: Community detection in complex networks using extremal optimization. Phys. Rev. E **72**, 027104 (Aug 2005). https://doi.org/10.1103/PhysRevE.72.027104, http://link.aps.org/doi/10.1103/PhysRevE.72.027104
14. Fortunato, S.: Community detection in graphs. Physics Report **486**, 75–174 (2010)
15. Fortunato, S., Hric, D.: Community detection in networks: A user guide. Physics reports **659**, 1–44 (2016)
16. Girvan, M., Newman, M.: Community structure in social and biological networks. Proc. Natl. Acad. Sci. USA **99 (12)**, 7821–7826 (2002)
17. Gleich, D.F.: Pagerank beyond the web. siam REVIEW **57**(3), 321–363 (2015)
18. Gleiser, P.M., Danon, L.: Community structure in jazz. Advances in Complex Systems **06**(04), 565–573 (2003). https://doi.org/10.1142/S0219525903001067, http://www.worldscientific.com/doi/abs/10.1142/S0219525903001067
19. Good, B.H., de Montjoye, Y.A., Clauset, A.: Performance of modularity maximization in practical contexts. Phys. Rev. E **81**, 046106 (Apr 2010). https://doi.org/10.1103/PhysRevE.81.046106, http://link.aps.org/doi/10.1103/PhysRevE.81.046106
20. Grauwin, S., Szell, M., Sobolevsky, S., Hövel, P., Simini, F., Vanhoof, M., Smoreda, Z., Barabási, A.L., Ratti, C.: Identifying and modeling the structural discontinuities of human interactions. Scientific Reports **7** (2017)
21. Guimerà, R., Nunes Amaral, L.A.: Functional cartography of complex metabolic networks. Nature **433**(7028), 895–900 (Feb 2005). https://doi.org/10.1038/nature03288, http://dx.doi.org/10.1038/nature03288
22. Guimera, R., Sales-Pardo, M., Amaral, L.A.N.: Modularity from fluctuations in random graphs and complex networks. Physical Review E **70**(2), 025101 (2004)
23. Hastie, T.: The elements of statistical learning : data mining, inference, and prediction : with 200 full-color illustrations. Springer, New York (2001)
24. Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., Ratti, C.: Geolocated twitter as proxy for global mobility patterns. Cartography and Geographic Information Science **41**(3), 260–271 (2014)
25. Javed, M.A., Younis, M.S., Latif, S., Qadir, J., Baig, A.: Community detection in networks: A multidisciplinary review. Journal of Network and Computer Applications **108**, 87–111 (2018)
26. Khan, B.S., Niazi, M.A.: Network community detection: A review and visual survey. arXiv preprint arXiv:1708.00977 (2017)
27. Landsman, D., Kats, P., Nenko, A., Kudinov, S., Sobolevsky, S.: Social activity networks shaping st. petersburg. In: Proceedings of the 54th Hawaii International Conference on System Sciences. p. 1149 (2021)
28. Landsman, D., Kats, P., Nenko, A., Sobolevsky, S.: Zoning of st. petersburg through the prism of social activity networks. Procedia Computer Science **178**, 125–133 (2020)
29. Lee, J., Gross, S.P., Lee, J.: Modularity optimization by conformational space annealing. Phys. Rev. E **85**, 056702 (May 2012). https://doi.org/10.1103/PhysRevE.85.056702, http://link.aps.org/doi/10.1103/PhysRevE.85.056702
30. Lusseau, D., Schneider, K., Boisseau, O.J., Haase, P., Slooten, E., Dawson, S.M.: The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations. Behavioral Ecology and Sociobiology **54**(4), 396–405 (2003). https://doi.org/10.1007/s00265-003-0651-y, http://dx.doi.org/10.1007/s00265-003-0651-y

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_51

31. Newman, M.E.J.: Fast algorithm for detecting community structure in networks. Phys. Rev. E **69**, 066133 (Jun 2004). https://doi.org/10.1103/PhysRevE.69.066133, http://link.aps.org/doi/10.1103/PhysRevE.69.066133
32. Newman, M.E.: Finding community structure in networks using the eigenvectors of matrices. Physical review E **74**(3), 036104 (2006)
33. Newman, M.: Modularity and community structure in networks. Proceedings of the National Academy of Sciences **103**(23), 8577–8582 (2006)
34. Newman, M., Girvan, M.: Finding and evaluating community structure in networks. Phys. Rev. E **69 (2)**, 026113 (2004)
35. Piccardi, C., Tajoli, L.: Existence and significance of communities in the world trade web. Phys. Rev. E **85**, 066119 (Jun 2012). https://doi.org/10.1103/PhysRevE.85.066119, http://link.aps.org/doi/10.1103/PhysRevE.85.066119
36. Plantié, M., Crampes, M.: Survey on social community detection. In: Social media retrieval, pp. 65–85. Springer (2013)
37. Ratti, C., Sobolevsky, S., Calabrese, F., Andris, C., Reades, J., Martino, M., Claxton, R., Strogatz, S.H.: Redrawing the map of great britain from a network of human interactions. PLoS ONE **5**(12), e14248 (12 2010). https://doi.org/10.1371/journal.pone.0014248,
38. Shchur, O., Günnemann, S.: Overlapping community detection with graph neural networks. arXiv preprint arXiv:1909.12201 (2019)
39. Sobolevsky, S., Belyi, A.: Graph neural network inspired algorithm for unsupervised network community detection. Applied Network Science **7**(1), 1–19 (2022)
40. Sobolevsky, S., Belyi, A., Ratti, C.: Optimality of community structure in complex networks. arXiv preprint arXiv:1712.05110 (2017)
41. Sobolevsky, S., Campari, R., Belyi, A., Ratti, C.: General optimization technique for high-quality community detection in complex networks. Physical Review E **90**(1), 012811 (2014)
42. Sobolevsky, S., Kats, P., Malinchik, S., Hoffman, M., Kettler, B., Kontokosta, C.: Twitter connections shaping new york city. In: Proceedings of the 51st Hawaii International Conference on System Sciences (2018)
43. Sobolevsky, S., Sitko, I., Des Combes, R.T., Hawelka, B., Arias, J.M., Ratti, C.: Money on the move: Big data of bank card transactions as the new proxy for human mobility patterns and regional delineation. the case of residents and foreign visitors in spain. In: Big Data (BigData Congress), 2014 IEEE International Congress on. pp. 136–143. IEEE (2014)
44. Sobolevsky, S., Szell, M., Campari, R., Couronné, T., Smoreda, Z., Ratti, C.: Delineating geographical regions with networks of human interactions in an extensive set of countries. PloS ONE **8**(12), e81707 (2013)
45. Sun, Y., Danila, B., Josić, K., Bassler, K.E.: Improved community structure detection using a modified fine-tuning strategy. EPL (Europhysics Letters) **86**(2), 28004 (2009), http://stacks.iop.org/0295-5075/86/i=2/a=28004
46. White, J.G., Southgate, E., Thomson, J.N., Brenner, S.: The structure of the nervous system of the nematode caenorhabditis elegans. Philosophical Transactions of the Royal Society of London. B, Biological Sciences **314**(1165), 1–340 (1986). https://doi.org/10.1098/rstb.1986.0056, http://rstb.royalsocietypublishing.org/content/314/1165/1.abstract
47. Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., Philip, S.Y.: A comprehensive survey on graph neural networks. IEEE transactions on neural networks and learning systems (2020)

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_51