Community detection, the process of identifying groups within a network, is a core problem in network science, social network analysis, and data analysis using networks [1], [2]. Among the most common approaches for detecting communities are the modularity maximization algorithms [3]–[5], which are designed to maximize a utility function, *modularity* [6], across all possible ways that a network can be broken down into communities. Three common methods for modularity maximization are the Clauset-Newman-Moore (CNM) algorithm [7], the Louvain algorithm [8], and the Leiden algorithm [9], which are all heuristic algorithms.

Generate synthetic networks based on Lancichinetti-Fortunato-Radicchi (LFR) benchmarks [10] and use them as experimentation data for comparing the performance of five modularity maximization algorithms including the CNM algorithm, the Louvain algorithm, and the Leiden algorithm (and two modularity-maximization algorithms of your own choosing) based on their running time and the modularity of their output communities evaluated on a wide range of LFR benchmark networks.

Email me the following for this assignment (by 7 Dec at 23:59 EST):

1) Link to a public GitHub repository which stores a Jupyter notebook file documenting all your computational experiments
2) A summary in pdf format (600 words excluding references) with one table (optional) and one illustrative figure (optional) explaining your experiments and findings. Include your GitHub repository link also in the first line of the pdf file. Use this format for your file: Firstname_Lastname_summary.pdf

Hints:

- The Python package ‘networkX’ is probably useful for you. It has functions for the CNM algorithm and for generating LFR benchmarks.
- The Python packages ‘igraph’, ‘louvain’, and ‘leidenalg’ will probably be useful as well.
- If you have any questions about this assignment, please consult online resources and references.
- If any part of this assignment is not fully determined in the above instructions, it is an opportunity for you to make your own choices and justify them in your summary.

References

[1] A. Lancichinetti and S. Fortunato, “Community detection algorithms: A comparative analysis,” *Phys. Rev. E*, vol. 80, no. 5, p. 056117, Nov. 2009, doi: 10.1103/PhysRevE.80.056117.

[2] Z. Yang, R. Algesheimer, and C. J. Tessone, “A Comparative Analysis of Community Detection Algorithms on Artificial Networks,” *Sci. Rep.*, vol. 6, no. 1, p. 30750, Aug. 2016, doi: 10.1038/srep30750.

[3] M. Chen, K. Kuzmin, and B. K. Szymanski, “Community detection via maximization of modularity and its variants,” *IEEE Trans. Comput. Soc. Syst.*, vol. 1, no. 1, pp. 46–65, 2014.

[4] A. Lancichinetti and S. Fortunato, “Limits of modularity maximization in community detection,” *Phys. Rev. E*, vol. 84, no. 6, p. 066122, Dec. 2011, doi: 10.1103/PhysRevE.84.066122.

[5] M. E. J. Newman, “Equivalence between modularity optimization and maximum likelihood methods for community detection,” *Phys. Rev. E*, vol. 94, no. 5, p. 052315, Nov. 2016, doi: 10.1103/PhysRevE.94.052315.
[6] M. E. J. Newman and M. Girvan, “Finding and evaluating community structure in networks,” *Phys. Rev. E*, vol. 69, no. 2, p. 026113, Feb. 2004, doi: 10.1103/PhysRevE.69.026113.
[7] A. Clauset, M. E. Newman, and C. Moore, “Finding community structure in very large networks,” *Phys. Rev. E*, vol. 70, no. 6, p. 066111, 2004.
[8] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, “Fast unfolding of communities in large networks,” *J. Stat. Mech. Theory Exp.*, vol. 2008, no. 10, p. P10008, 2008.
[9] V. A. Traag, L. Waltman, and N. J. van Eck, “From Louvain to Leiden: guaranteeing well-connected communities,” *Sci. Rep.*, vol. 9, no. 1, p. 5233, Mar. 2019, doi: 10.1038/s41598-019-41695-z.
[10] A. Lancichinetti, S. Fortunato, and F. Radicchi, “Benchmark graphs for testing community detection algorithms,” *Phys. Rev. E*, vol. 78, no. 4, p. 046110, Oct. 2008, doi: 10.1103/PhysRevE.78.046110.