

Take-home assignment

Tasks

- A.** Read these three papers, think about them, and try to understand them to the best of your ability.
1. Lambiotte, Renaud, Jean-Charles Delvenne, and Mauricio Barahona. "Random walks, Markov processes and the multiscale modular organization of complex networks." *IEEE Transactions on Network Science and Engineering* 1.2 (2015): 76-90. [Pay careful attention to section 6.1]
 2. Aref, Samin, Mahdi Mostajabdaveh, and Hriday Chheda. "Bayan algorithm: Detecting communities in networks through exact and approximate optimization of modularity." *Physical Review E* 110.4 (2024): 044315. <https://arxiv.org/pdf/2209.04562>
 3. Arnaudon, Alexis, et al. "Algorithm 1044: PyGenStability, a multiscale community detection framework with generalized markov stability." *ACM Transactions on Mathematical Software* 50.2 (2024): 1-8.
- B.** Write a summary of the three papers without using any AI tools. In your summary, you can put more focus on Section 6.1 of the first paper. Your summary must answer this question "*why Markov stability can be expressed as modularity?*" Include formal citations for referring to the papers in your summary.
- C.** Reproduce the results from Table 2 of paper 2 above to quantify the extent to which you can get the same results when you run the same computational experiments.
1. Use Python and fix the random seed numbers so that you consistently get the same results if you rerun your experiments.
 2. For reproducing the results of Table 2, you can use the FigShare dataset <https://doi.org/10.6084/m9.figshare.22442785> for accessing the 500 ABCD networks
 3. Make sure to use the same versions of libraries PyGenStability, graph-tool, networkX, and CDlib as described in Section VI of paper 2 above. If not possible, try to use the closest versions.
 4. From the 30 algorithms of Table 2, you should use at least 8 of them in your experiments. The must include these algorithms: Bayan, MV, and MR.
- D.** Compare your results with the results in Table 2 of paper 2 and describe any discrepancies that you have observed.

Submission Requirements

Please email me aref@mie.utoronto.ca the following for this assignment (by Sep 21th at 5pm):

1. **GitHub Repository:** Provide a link to a public GitHub repository containing a Jupyter notebook file that documents all your algorithms and code.
2. **Summary Report:** Submit a summary in PDF format (maximum 1000 words, excluding references). The summary should include one table (your reproduced results from task C), a few paragraphs for task B, and a few paragraphs for task D. Include the GitHub repository link in the first line of the PDF file. Use the following format for your file: Firstname_Lastname_summary.pdf.

Hints

- Take your time. This is not meant to be an easy or small assignment.
- If you have any questions about this assignment, feel free to consult online resources and references. You can check other articles on <https://scholar.google.com/> or the references of the three papers whenever something is unclear to you.
- If any part of this assignment is not fully defined in the above instructions, take it as an opportunity to make your own choices and justify them in your summary.