

Due date: March 13, 2020

As part of the requirements of this course, the class project involves reading and understanding a study from the graph analytics literature. Select a paper or book chapter you would like to work on. The following items should be considered.

### Problem formulation and modeling

Explain the study of the selected study. In particular:

1. Explain why graph analytics offers new insights that helps solve the particular problem at hand.
2. Describe the graph introduced in the paper. Use the notation presented in class whenever possible.
3. Recreate the main results of the paper (through simulation). Use Networkx to analyze the graph and Neo4j to store it. Cypher queries should be done from Python.
4. Evaluate the robustness of an undirected version of the graph using Weighted Spectral Distribution (WSD) (e.g., to take into account data inaccuracies). See articles [1]-[2] for reference.

### Report

Make sure the report includes the following sections.

1. An abstract: A brief summary of the article that helps the reader quickly ascertain the project's purpose. It should be approximately half a page long and include in a very brief manner:
  - a. Statement of the problem being addressed.
  - b. Findings of the paper.
  - c. Main conclusions and recommendations.
2. The model: The formal representation of the phenomenon studied through graphs. Make sure to describe implicit and explicit assumptions being made by the type of graph that is being considered (e.g., simple, directed, pseudo-graphs, etc.).
3. Robustness: Present the results of applying robustness analysis based on WSD. In particular:
  - a. Proof that the trace of the normalized eigenvalues is equal to the sum of the eigenvalues of the normalized Laplacian.
  - b. Describe the approach and its equation.
4. Results: The analysis of your model and its implications (including simulation outcomes). Make sure to email me your code.
5. Conclusions: Present a summary that highlights the major points of your work. What did you learn from the project?

**Make sure your final report is not too long, but concise. Its length should not exceed 4 pages in double column. Oral presentations will be 15 minutes long (5 minutes for questions).**

### Grading

Report (your response to the items listed above) is worth 80%

Oral presentation - 20%

### References

- [1] Damien Fay, Hamed Haddadi, Steve Uhlig, Andrew W. Moore, Richard Mortier, Almerima Jamakovic. Weighted spectral distribution. Technical Report Number 729, 2008. Available at <https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-729.pdf>
- [2] B. Jiao, Y. Nie, C. Huang, J. Du, X. Pang and X. Yuan, "Calculating the weighted spectral distribution with 5-cycles," 2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Nanchang, 2016, pp. 192-197. Available at <https://ieeexplore.ieee.org/document/7565987>