Homework 1

Due date is **3th December 2020, 23:59!**

Homework Policy

- Use comments whenever necessary to explain your code.
- You will use Python as the programming language.
- IMPORTANT: This is an individual assignment! You are expected to act according to Student Code of Conduct, which forbids all ways of cheating and plagiarism. It is okay to discuss the homework with others, but it is strictly forbidden to use all or parts of code from other students' codes or online sources and let others do all or part of your homework.
- Only electronic submissions through Ninova will be accepted. You only need to submit your Jupyter Notebook file. Any comments and discussions should be included in the same file.
- If you have any questions, you can contact me at kamard@itu.edu.tr. Do NOT hesitate to send an e-mail if you are confused.

Jupyter Notebook Installation

You need to have Python and Pip installed in your computer to install Jupyter Notebook.

Windows

On command prompt (cmd.exe with admin mode):

```
C:\**path**> python -m pip install jupyter
```

After changing your current folder to the folder which you want to work on (see 'cd' command):

```
C:\**your_working_folder**> jupyter notebook
```

Then it will be launched on your default browser

Ubuntu/Linux/Unix/Mac

On Terminal:

```
$ pip install notebook
```

Then launch with:

```
$ jupyter notebook
```

For more information: https://jupyter.org/install

Introduction

You will carry out all the tasks below using **Ipython Notebook**. Simply add all your work to the provided template file **HW1_template.ipynb** using jupyter notebook.
In this homework, you will code up several experiments in Oldham et al. paper[3] (You can click anywhere on this sentence instead of a small, hard-to-click word "here" to find the paper).To this aim, first simulate 40 networks: 20 unconstrained weighted using Erdos-Renyi generative model [1] and 20 constrained unweighted using Maslov-Sneppen algorithm [2]. For more details, check the attached supplementary material. For this task, you can use ready-made pieces of code. But all needs to be commented out. The number of nodes in each network category (e.g., ER) should equal to 200.

Part A Simulate random weighted and unweighted networks (25 Points, 5 points each question)

1. Briefly explain how Erdos-Renyi generative model works.
2. What are the key properties of Erdos-Renyi graphs?
3. Briefly explain how Maslov-Sneppen algorithm works.
4. What are the key properties of Maslov-Sneppen graphs?
5. Visualize two random graphs you simulated (ER and MS).

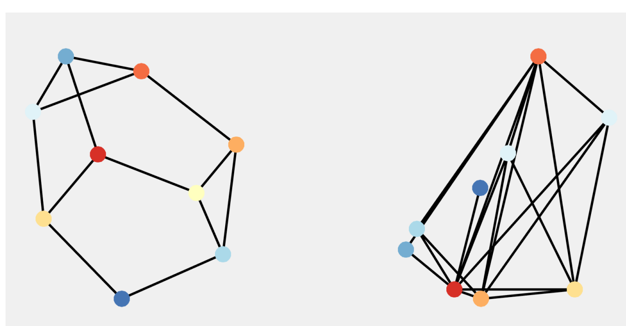

Figure 1: Example visualization of two graphs.

Part B: Analyzing Erdos-Renyi and Maslov-Sneppen graphs using centrality measures (50 Points)

1. **[15p]** Code up the necessary steps to reproduce plots A and B in Figure 2 using the Centrality Measure Correlation(CMC)s of the 40 networks you simulated: 20 weighted Erdos-Renyi and 20 unweighted Maslov-Sneppen networks.

2. **[5p]** What conclusions can you derive from the plots?

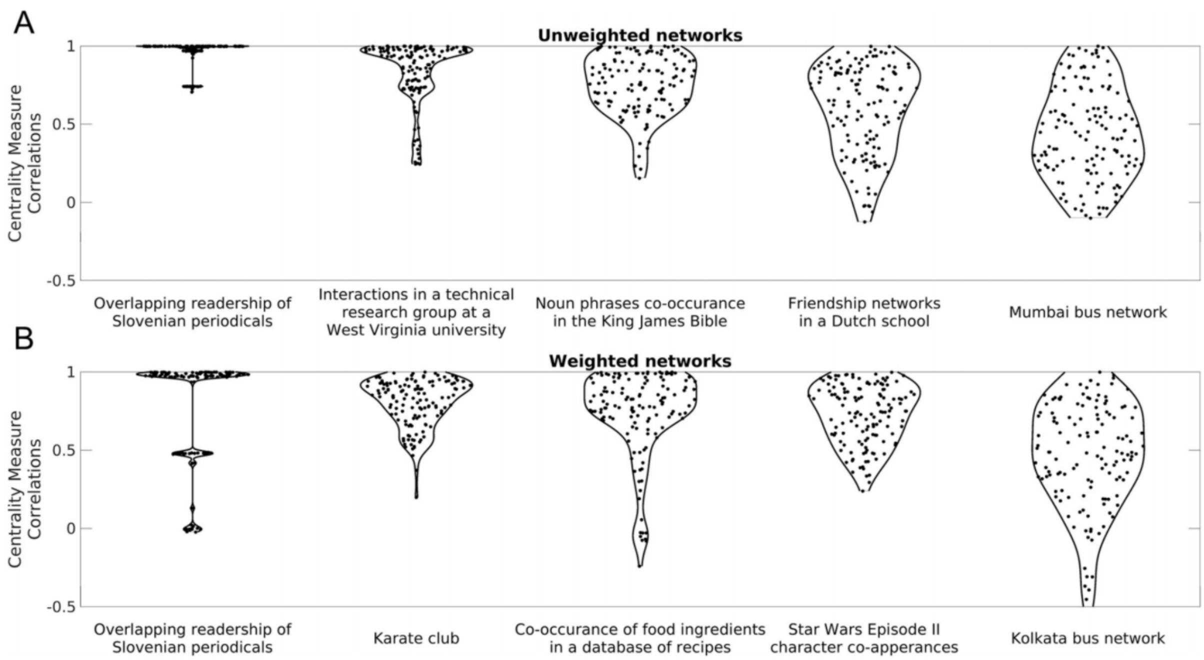

Figure 2: **Distributions of Centrality Measure Correlations (CMCs) for example unweighted and weighted networks.** Distributions of CMCs for every pair of centrality measures for five example unweighted (panel A); and weighted networks (panel B). Networks have been ordered from highest (left) to lowest (right) median CMC. **Both the figure and the caption is from [3]**

3. **[15p]** Code up the necessary steps to reproduce Figure 3 (A to D) using the between-network CMCs of the 40 networks you simulated: 20 weighted Erdos-Renyi and 20 unweighted Maslov-Sneppen networks.

4. **[5p]** What conclusions can you derive from the plots?

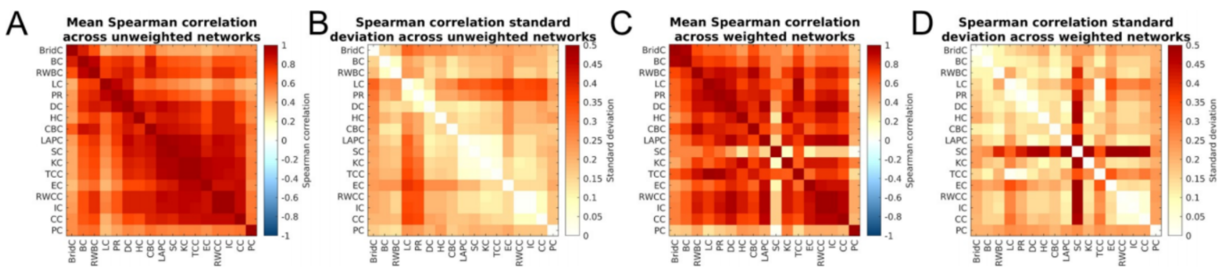

Figure 3: **Mean and standard deviation of between-network CMCs.** Panels A and B show the between-network CMC mean and standard deviation for unweighted measures, respectively. Panels C and D show the between-network CMCs mean and standard deviation for weighted measures, respectively. **Both the figure and the caption is from [3]**

5. **[10p]** Plot the CMC distributions using bar plots as in the Figure 4 for your 20 unweighted and 20 weighted networks. What do you notice?

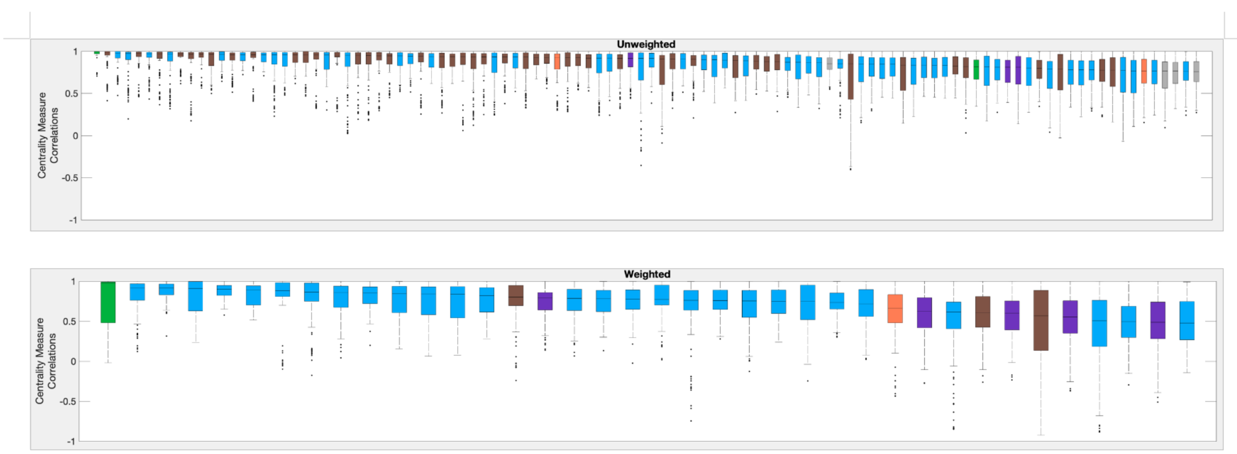

Figure 4: Example bar plot of CMC distributions.

Part C: Association between mean within-network Centrality Measure Correlation(CMC) and network properties (25 Points)

1. **[20p]** Choose two different network properties (e.g., modularity and density) and regenerate similar plots to Figure 5 by plotting the mean within-network CMC against the selected network property. Erdos-Renyi networks should have the same color (e.g., transparent blue) and Maslov-Sneppen networks can be assigned a different color (e.g., transparent red). You can use existing codes to compute the network properties.

2. **[5p]** Interpret your plots. What conclusions can you derive?

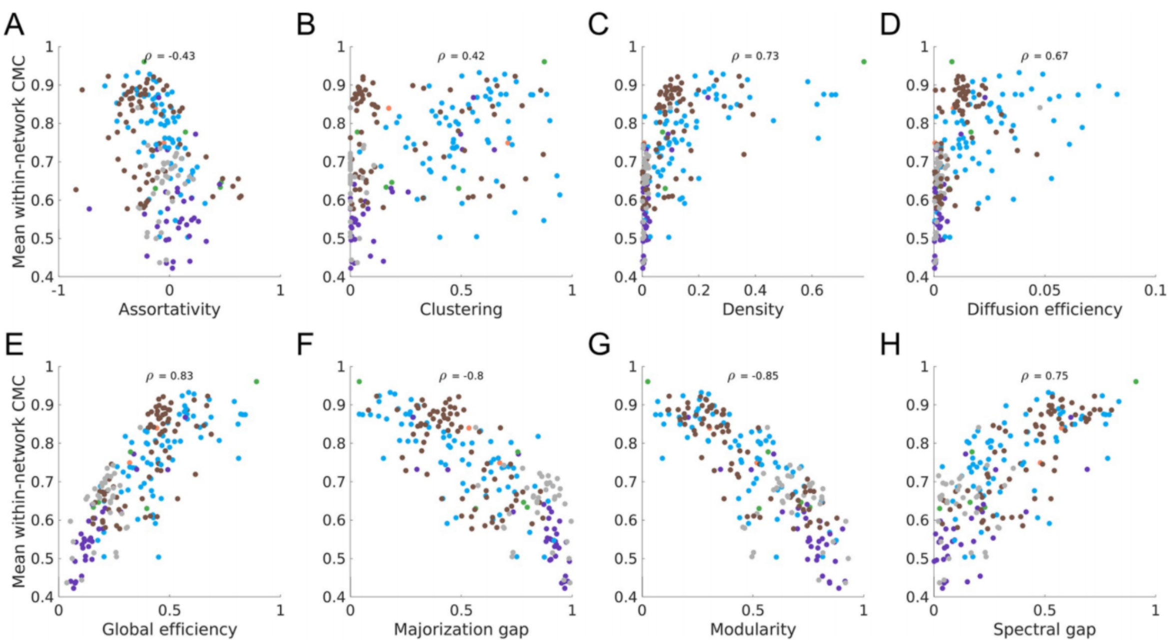

Figure 5: **Association between mean within-network CMC and network properties in unweighted networks.** The association between the mean within-network CMC (the average of all CMCs within a single network) and each of the global topological properties. Networks are coloured by their natural category (blue = social, grey = technological, brown = biological, orange = informational, purple = transportation; green = economic). **Both the figure and the caption is from [3]**

Important note: The Matlab code of all experiments of the paper [3] is available as well as the output figures and data. Feel free to get inspired from those.

Penalty of [-10 points]: Your code should be generic. The number of networks to simulate should be defined in the beginning as a hyper-parameter to fix as well as the size of the networks. If the code cannot be executed when changing the number of networks to simulate or their size (i.e., number of nodes), there will be a penalty of 10 points.

References

[1] Bela Bollobss. The evolution of random graphs. *Transactions of the American Mathematical Society*, 286(1):257–274, 1984.

[2] Sergei Maslov and Kim Sneppen. Specificity and stability in topology of protein networks. *Science*, 296(5569):910–913, 2002.

[3] Stuart Oldham, Ben Fulcher, Linden Parkes, Aurina Arnatkeviciute, Chao Suo, and Alex Fornito. Consistency and differences between centrality metrics across distinct classes of networks. 05 2018.