Graph Theory Project (Fall 2021)

Instructor: Islem Rekik

Istanbul Technical University, Turkey

1 What problem will we solve?

The main goal of each project is to **investigate a particular relationship $\mathcal{R}$ between two graph-driven metrics or properties P_1 and P_2.** In the (P_1, P_2) space, plot simulated graphs from two *anchor* graphs G_1 and G_2 and derive conclusions and remarks.

Python Jupyter notebook presentation should cover the following key points:

- **I- Graph datasets G_1 and G_2 description.** ⇒ (1) Present your graphs G_1 and G_2 (number of nodes, edges, what a node and edge denote?). Notice that your choice of both graphs is related to the relationship $\mathcal{R}$ you intend to explore. Your choice should be well thought out. SNAP library contains many graph datasets[a]. (2) Display the adjacency matrices of both graphs. (3) Plot the node degree distributions of both graphs in an overlaid manner. [15%]
- **II- My graph theory principle: *name it here.*** ⇒ (1) Present the mathematical principle or concept to explore. The concept is the principle you are investigating (e.g., graph information diffusion, modularity, centrality, efficiency, etc). (2) Include a figure illustrating how it works on small graphs. (3) Formalize the principle mathematically and present your mathematical symbols and equations with comments and explanations. [25%]
- **III-The R space.** ⇒ (1) Define $\mathcal{R}$ between two graph-driven metrics or properties P_1 and P_2 to better investigate the graph theory principle/concept you selected in Section II. Justify your rationale for exploring such relationship. (2) Simulate $n_1 = 100$ random graphs derived from G_1 with same node strength distribution. (3) Simulate $n_2 = 100$ random graphs derived from G_2 with same node strength distribution. Detail the **simulation strategy**. (4) For each anchor graph G_i, plot its node strength distribution overlaid with that of the simulated graphs from G_i. (5) Plot the anchor graphs in the $\mathcal{R}$ space (as *star*) and the simulated graph populations in dots. For a better visualization, use two different colors. [40%]
- **VI-Result interpretation and discussion.** ⇒ Discuss and interpret the simulated graph behavior in the $\mathcal{R}$ space. Compare both simulated datasets and anchor graphs and derive conclusions. [10%]
- **Live Zoom group presentation.** ⇒ The length of the presentation should be about 5-8 min + 3 min Q&A session [10%]. During the presentation, each team members will be marked on how they address raised questions from the audience.

[a] https://snap.stanford.edu/data/

2 How to submit our project?

- Work in groups of 3 to 5 students. Individual projects will not be accepted .
- Register your group topic and full names using the following Excel sheet: https://docs.google.com/spreadsheets/d/1tN-cIY1b2UNbRZURjrkpxlpehYNJWsQNQAKeoCWbwgg/edit#gid=0
  Remark 1. Two groups at max can have the same graph theory principle/concept of work on in their projects. If I spot that more than two groups selected the same topic I will contact the third group who registered later so they change their topic . The rule we follow is 'first come first served' . **Feel free to explore a concept that we did not see in class.**
- Submit a zipped file entitled 'group_ID.zip' (ID denotes your group number) containing: **(1)** your graph datasets and any dependent files that are required to run your Jupyter Python notebook, **(2)** the saved .ipynb file displaying all results of different sections, and **(3)** upload a ".txt" file including the link to the pre-recorded video to download. Make sure that your prerecorded video lasts between 5 and 8 minutes. Otherwise there will be a strong penalty. .
  Partition your Jupyter Python notebook into 4 parts following the project chart.
  Remark 1. Make sure to well comment out your code.
  Remark 2. The code should be generic. Example: n_1 and n_2 are parameters that can be played with in different code runs.
  Penalty . There will be a penalty of -10 for codes that lack comments and are not generic.
  Penalty . During the day of the presentation, the camera should be on of all members.