Modelling and Analysis of Complex Networks

Exercise 4

Due: 13:00 on Oct. 18, 2024

The maximum score of this assignment is: **16 points**. Please submit the assignment in any readable data format (.txt, .doc, .pdf, .md ...) and submit the assignment before the deadline. If you have additional information concerning your answers, please also upload the document to the Moodle, or include the link to the document in your answers (e.g., link to your Github repository). Please indicate your team number in your submission.

Now let's explore more about the two datasets we have: Facebook-Ego and Twitter-Ego. The following questions can be answered with the help of Snap.py and NetworkX. You may also use other packages to deal with the problem. Please answer the following questions on both of the networks you have and submit your executable code. **($4 + 3 + 3 + 2 \times 3 = 16$ points)**

(a) Choose one of the datasets, calculate the betweenness closeness and PageRank mentioned in the lecture on all of the nodes. Please compare the results. What cause the difference between the results?

(b) Please compare the results in (a). What cause the difference between the results?

(c) Please plot the largest strong connected component in the Twitter-Ego dataset.

(d) Choose the results of one of the two algorithms mentioned in question (a), and choose the node with the highest value and the node with lowest value. Please plot the distributions of Jaccard similarity, cosine similarity and normalized Euclidean distance between either of the two nodes and the rest of the nodes in the graph.