Social Computing, Evaluation 2

CS60081, Autumn 2021

12:10 pm to 12:50 pm, 7th October 2021
Full marks: 40
Answer ALL questions

IMPORTANT INSTRUCTIONS

Taking the exam: You need to log into zoom, keep your video on during taking the test (so that we can monitor you during the exam). You will use pen and paper to write the exam,
Decorum: Throughout the examination, you are strictly expected to have their cameras on, directing towards their workspace including themselves. Arrange your laptops/desktops/mobiles beforehand to save time during the examination. Disconnecting video for a long duration will be grounds for suspecting malpractice.

You need to keep your workplace, your hands and your mobiles visible to us. We are trying to avoid the visibility of your answers in the papers to the rest of them. Once you open your question paper, refrain from using your PC/laptop from searching for anything or typing during the exam.

Tip: Install Adobe scan and, MS Teams on your phone to make the whole process easier. In that case, your laptop acts as a camera, while you are using your mobile for checking the questions, scanning and uploading the answers.

Submission: You can do either of two things (i) take pictures of your answer script pages, name the pictures page1.jpg, page2.jpg, page3.jpg etc., zip the pictures and upload the zipped file via CSE Moodle. (ii) Put all the pages sequentially in a pdf file and upload the pdf to KHARAGPUR Moodle. YOU HAVE TO USE PEN AND PAPER TO GIVE THE EXAM.

- Name your zip file as <your roll no>.zip

Policies: Note that, if we face problems with your answer script e.g., cannot open your submitted zipped file, cannot read the text in pictures (due to bad resolution), cannot determine the page order from the file names (or the pages in the pdf is jumbled up), or we find you copying, it will affect your marks.

Malpractice: If any group of students is found to have similar work in their answer sheets, all of them will receive the maximum penalty with no grace. We expect you to not take help from the internet, your copies, textbooks, slides or video recordings during the exam. Note that this is not an open-book exam. If found otherwise, you will be penalized.

PLEASE WRITE YOUR NAME AND ROLL NO. ON THE TOP OF THE FIRST PAGE OF YOUR ANSWER SCRIPT. WE WILL NOT EVALUATE YOUR ANSWER SCRIPT WITHOUT IT.

Question 1. You are collecting data from Facebook for the purpose of network analysis of an individual user. Which of the friendship network(s) can you potentially collect? Please explain your answer (1 – 3 sentences). No marks without explanation.

a) Complete 1-degree egocentric network
b) 1.5-degree egocentric
c) 2-degree egocentric
d) None

[3 marks]

Question 2. Let X, Y and Z are three social media platforms where X, Y are more popular than Z. Let's define a metric k which is the ratio of E_u (expected engagements for unreliable posts) and E_r (expected engagements for reliable posts). The values of k for X, Y, Z are 100%, 40% and 250%. Today we are seeing many misinformation getting propagated through social media. In your view which of these 3 social media platforms will be the most susceptible to misinformation propagation and why? No marks without explanation

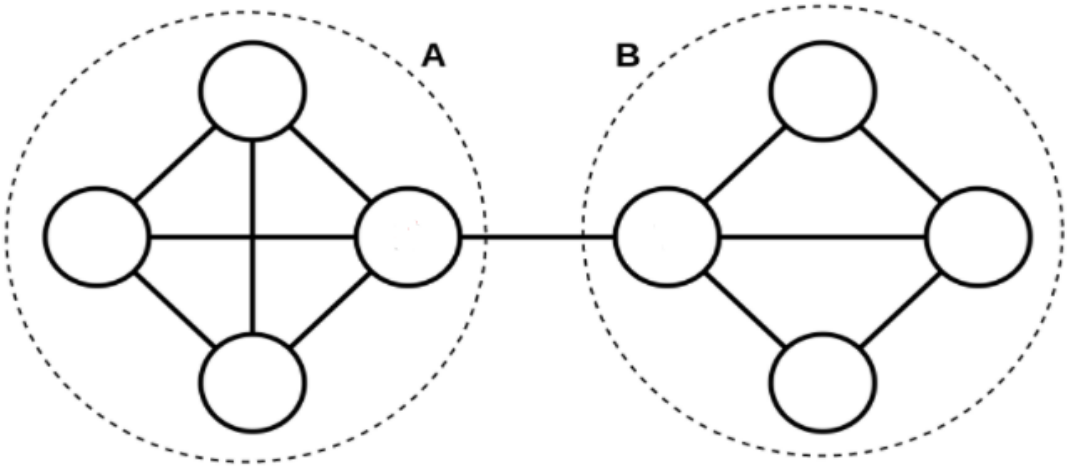
[5 marks]

Question 3. Consider a complete graph with n nodes. Its clustering coefficient can be represented in the form of $(an^2 + bn + c)$. What is the value of $(a + 2b + 3c)$? (Give an Integer). No marks without showing calculation/logic.

[4 marks]

Question 4. Calculate the modularity of the network given below with two communities A and B.

[6 marks]



Question 5. Draw examples of directed graphs (number of nodes > 6) which satisfy the following properties and justify your answers briefly (no marks without justification):

[3 + 3 = 6 marks]

A. There is no difference between the degree centrality and page-rank centrality of the nodes in the graph.

B. There is an obvious difference between the degree centrality and page-rank centrality of the nodes in the graph.

Question 6. Imagine a situation where you are tasked to create an algorithm to detect fake accounts on Facebook. You have the knowledge that fake accounts are spread throughout the Facebook graph and even if their goal is to make friends with real user accounts, to increase credibility, fake accounts often connect with other fake accounts in different parts of the Facebook graph (e.g., Brazilian Fake accounts connect with other Brazilian fake accounts). Now, thanks to the Facebook user-reports you know details of 2,000 fake accounts. Please answer the following questions

[5 + 3 + 2 = 10 marks]

A. Write an algorithm to effectively detect rest of the fake accounts in the network. Please write your algorithm in a pseudo code format as discussed in class for Page rank or HITs algorithms.

B. Give reasons for why you believe your proposed scheme would be successful in identifying more fake accounts.

C. Briefly describe 2 shortcomings of your scheme in correctly identifying more fake accounts.

Question 7. Imagine that you have a small-world network. Answer the following questions and justify the answers (no marks without justification)

[2 + 4 = 6 marks]

A. Would you use a divisive or an agglomerative community detection algorithm on this network to find communities? Why?

B. Imagine you have a dendrogram of hierarchical communities for this network. How would you use a community quality metric to choose a level to cut this dendrogram to obtain a set of communities? Which community quality metric would you choose? Why?