Queen Mary University of London

Main Examination period 2023 – May/June – Semester B

MTH750U / MTH750P: Graphs and Networks

Duration: 3 hours

The exam is intended to be completed within **3 hours**. However, you will have a period of **4 hours** to complete the exam and submit your solutions.

You should attempt ALL questions. Marks available are shown next to the questions.

All work should be **handwritten** and should **include your student number**. Only one attempt is allowed – **once you have submitted your work, it is final**.

In completing this assessment:

- You may use books and notes.
- You may use calculators and computers, but you must show your working for any calculations you do.
- You may use the Internet as a resource, but not to ask for the solution to an exam question or to copy any solution you find.
- You must not seek or obtain help from anyone else.

When you have finished:

- scan your work, convert it to a **single PDF file**, and submit this file using the tool below the link to the exam;
- e-mail a copy to **maths@qmul.ac.uk** with your student number and the module code in the subject line;

Examiners: V. Latora, F. Fischer

Question 1 [40 marks].
Consider the graph G of $N = 5$ nodes described by the following incidence matrix:

$$B = \begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix}$$

(a) Write down the adjacency matrix A of the graph. Draw the graph. Is the graph directed or undirected? [**5**]

(b) How many weakly and how many strongly connected components are there in the graph? Which are the nodes in each weakly connected component? Which are the nodes in each strongly connected component? [**4**]

(c) Determine the in-degree distribution and the out-degree distribution. [**4**]

(d) Calculate the $N \times N$ matrix D whose entry d_{ij} indicates the length of the shortest paths from node i to node j. [**5**]

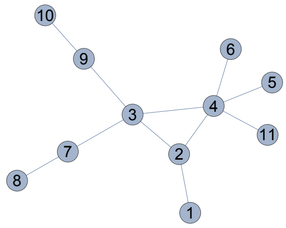
(e) Calculate the in-degree centrality $c_i^{D^{\text{in}}}$ of each node $i = 1, 2, \ldots, N$ of the graph, and rank the nodes based on their values of this centrality. [**4**]

(f) Find the eigenvalues and eigenvectors of matrix $A^\top$ (where the symbol $\top$ indicates the transpose of a matrix), and define and calculate the eigenvector centrality c_i^E of each node $i = 1, 2, \ldots, N$ of the graph G as the i-th component of the eigenvector associated with the eigenvalue $\lambda = 1$ of the matrix $A^\top$. Rank the nodes based on this centrality. Explain why, for this graph, we cannot define the node centralities from the components of the eigenvector associated to the largest (in absolute value) eigenvalue. [**6**]

(g) Write down the definition of the α-centrality, and calculate the α-centrality c_i^α of each node $i = 1, 2, \ldots, N$ of the graph. Which values of the parameter α are not allowed? Rank the nodes based on this centrality. [**7**]

(h) Write down the definition of the PageRank centrality, and calculate the PageRank centrality c_i^{PR} of each node $i = 1, 2, \ldots, N$ of the graph. Rank the nodes based on this centrality. [**5**]

Question 2 [30 marks].
Consider the graph G shown in the figure below.



(a) What is the probability of obtaining graph G by sampling a graph from an Erdős-Rényi random graph ensemble $G^{ER}_{N,K}$ with $N = 11$ and $K = 10$? What is the probability of obtaining graph G by sampling a graph from an Erdős-Rényi random graph ensemble $G^{ER}_{N,K}$, with $N = 11$ and $K = 11$? **[6]**

(b) Consider the Erdős-Rényi random graph ensemble $G^{ER}_{N,p}$ with $N = 11$. Find the value of p such that the ensemble $G^{ER}_{11,p}$ best describes graph G. (Hint: - *fix p so that the expectation value for the number of edges in $G^{ER}_{11,p}$ is equal to the number of edges of G*). **[5]**

(c) How is the number of links distributed over the graphs of the Erdős-Rényi random graph ensemble $G^{ER}_{11,p}$, with p being the value determined in part (b) above? What is the probability of finding a graph with 11 links in such ensemble? **[6]**

(d) What is the probability of obtaining graph G by sampling a graph from the Erdős-Rényi random graph ensemble $G^{ER}_{11,p}$, with p being the value determined in part (b)? Compare this probability to that found in part(a) for an the Erdős-Rényi random graph ensemble $G^{ER}_{N,K}$ with $N = 11$ and $K = 11$. **[6]**

(e) Prove the general result: let $G = (\mathcal{V}, \mathcal{E})$ be a graph with $N = |\mathcal{V}|$ nodes and $K = |\mathcal{E}|$ edges. Then for any $p \in (0, 1)$ the probability of obtaining G when sampling from $G^{ER}_{N,K}$ is higher than the probability of obtaining G when sampling from $G^{ER}_{N,p}$. **[7]**

Question 3 [30 marks].
Consider the following model to grow graphs. Given a positive integer N, the graph grows, starting at time $t = 0$ with $n_0 = 2$ nodes and $l_0 = 1$ links, and by iteratively repeating at time $t = 1, 2, \ldots, N - n_0$, the following steps:

(1) A link (i, j) is randomly chosen with a uniform probability:

$$\pi_{(i,j)} = \frac{a_{ij}}{l_{t-1}}$$

among the existing link in the graph. Here, $A = \{a_{ij}\}$ and l_{t-1} are respectively the adjacency matrix and the number of links in the graph at time $t - 1$.

(2) A new node, labelled by the index n, being $n = n_0 + t$, is added to the graph. The node arrives together with $m = 2$ links which are respectively connected to the two nodes i and j selected in step (1).

(a) Show that, at each time step t, the average number of links added to a node i is equal to:

$$\frac{2k_{i,t-1}}{\sum_{j=1}^{n-1} k_{j,t-1}}.$$

i.e. the graph growth follows a preferential attachment rule. **[4]**

(b) Find an expression for the number of nodes, n_t, and the number of links l_t, as a function of time t. What is the average degree $\langle k \rangle$ of the network at time t? What is the final average degree in the limit $N \to \infty$? **[4]**

(c) Use the result at part (a) to write down and solve the differential equation governing the time evolution $\overline{k}_i(t)$ of the average degree of a node i for $t \gg 1$ in the mean-field approximation. **[5]**

(d) Derive the degree distribution of the network at large times in the mean-field approximation. **[5]**

(e) Let $\overline{n}_{k,t}$ be the average number of nodes with degree k in the graph at time t (the average, as usual, is performed over infinite realizations of the growth process with the same parameters of the model). Write down the rate equation satisfied by $\overline{n}_{k,t}$. **[6]**

(f) Solve the rate equation derived in part (f) in the limit $N \to \infty$, finding the exact result for the degree distribution of the model. **[6]**

End of Paper.