Queen Mary
University of London

Main Examination period 2018

MTHM750 / MTH750U / MTH750P: Graphs and Networks

Duration: 3 hours

Apart from this page, you are not permitted to read the contents of this question paper until instructed to do so by an invigilator.

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

Calculators are not permitted in this examination. The unauthorised use of a calculator constitutes an examination offence.

Complete all rough work in the answer book and cross through any work that is not to be assessed.

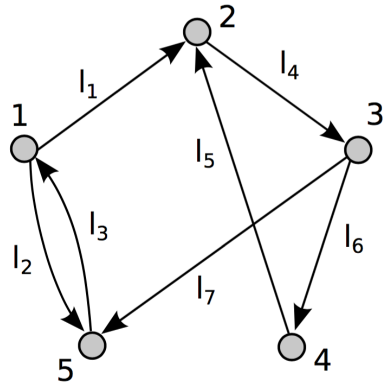
Possession of unauthorised material at any time when under examination conditions is an assessment offence and can lead to expulsion from QMUL. Check now to ensure you do not have any notes, mobile phones, smartwatches or unauthorised electronic devices on your person. If you do, raise your hand and give them to an invigilator immediately.

It is also an offence to have any writing of any kind on your person, including on your body. If you are found to have hidden unauthorised material elsewhere, including toilets and cloakrooms, it shall be treated as being found in your possession. Unauthorised material found on your mobile phone or other electronic device will be considered the same as being in possession of paper notes. A mobile phone that causes a disruption in the exam is also an assessment offence.

Exam papers must not be removed from the examination room.

Examiners: V. Latora, G. Bianconi

Question 1. [41 marks]
Consider the directed graph G_1, with $N = 5$ nodes and $K_1 = 7$ arcs, shown below:



(a) Write down the arc list, the adjacency matrix, A_1, and the incidence matrix, B_1, of the graph. Can you say that the graph is directed by only looking at its adjacency matrix? **[9]**

(b) Is graph G_1 weakly connected? Is G_1 strongly connected? Find the matrix D_1 of distances between the nodes of the graph. **[5]**

(c) Consider now the undirected graph G_2 obtained from G_1 by connecting a pair of nodes i and j in G_2 if at least one of the two arcs, from i to j or from j to i, exists in G_1. Write down the adjacency matrix, A_2, of graph G_2, the number of links in G_2, and the matrix of node distances, D_2. Write down the degree sequences and the degree distributions of both G_1 and G_2. **[11]**

(d) State the definitions of node clustering coefficients and that of clustering coefficient of a graph. State the definition of characteristic path length of a graph. Evaluate the clustering coefficient of graph G_2 and its characteristic path length. **[8]**

(e) State the definition of node closeness centrality and that of node betweenness centrality. Evaluate the normalised closeness centrality and the normalised betweenness centrality of node 3 both in graph G_1 and in graph G_2. **[8]**

Question 2. [29 marks]
Consider an ensemble of Erdős-Renyí (ER) random graphs $G^{\mathrm{ER}}_{N,p}$, where N is the number of nodes in the graphs and p is the probability to connect each pair of nodes. Assume $N \to \infty$ and that the connection probability scales, as a function of N, as $p(N) = zN^{-1}$, where z is a non negative real constant.

(a) Determine the expression for the node degree distribution p_k of the graphs in the ensemble. Approximate the degree distribution by a Poisson distribution, and express the values of the average node degree $\langle k \rangle$, and the standard deviations, σ_k, of node degrees (averages are performed over graph nodes) as functions of z **[6]**

(b) Find an expression for the average number of cycles of length three in the graphs of the ensemble as a function of z. **[5]**

(c) Define $S = \overline{s_1}/N$, where $\overline{s_1}$ is the expected size of the largest component in the graphs of the ensemble (the average is performed over the graphs in the ensemble). Show that S satisfies the self-consistency equation:

$$S = 1 - e^{-zS}$$

[5]

(d) Solve graphically the equation above, and prove that $S \neq 0$, i.e. the system has a non-zero giant component, when the average node degree is larger than a critical value $\langle k \rangle_c = 1$. **[7]**

(e) State the Molloy-Reed criterion. Show that the criterion returns the same result as in point (d), when the degree distribution considered is Poisson. **[6]**

Question 3. [30 marks]
Consider the following non-linear modification to the Barabási-Albert model. Given three positive integers $N \gg 1$, n_0 and m (with $m \leqslant n_0$), and a real number $\beta \in [-1,1]$, grow a graph, by starting at time $t = 0$ with a complete graph with n_0 nodes, and by iteratively repeating at time $t = 1,2,3,\ldots,N-n_0$, the following two steps:

(1) A new node, labeled by the index n, being $n = n_0 + t$, is added to the graph. The node arrives together with m edges.

(2) The m edges link the new node to m different nodes already present in the system. The probability $\Pi_{n\to i}$ that a new edge links the new node n to node i (with $i = 1,2,\cdots,n-1$) is:

$$\Pi_{n\to i} = \frac{k_{i,t-1}^{\beta}}{\sum_{l=1}^{n-1} k_{l,t-1}^{\beta}}$$

where $k_{i,t}$ is the degree of node i at time t.

(a) Write down the expression for the number of nodes, n_t, the number of links, l_t, and the average node degree as functions of time t. **[6]**

(b) Find the average node degree $\langle k \rangle$ when $N \to \infty$ **[4]**

(c) From now on, fix $n_0 = 3$ and $m = 3$. Write down the rate equations of the model, i.e. the equations for $\overline{n}_{k,t}$, where $\overline{n}_{k,t}$ denotes the average number of nodes with degree k present in the graph at time t. Averages are performed over infinite realizations of the growth process with the same parameters N, n_0, m and β. **[6]**

(d) In the case $\beta = 1$, determine the stationary degree distribution p_k (i.e. the limit of $p_{k,t} = \overline{n}_{k,t}/n_t$ when $t \to \infty$), and show that it is scale free with an exponent γ equal to 3. **[9]**

(e) In the case $\beta = 1$, write down and solve the differential equation governing the time evolution of the average degree $\overline{k_i}(t)$ of a node i introduced at time $t_i = 100$, in the so-called mean-field approximation. **[5]**

End of Paper.