Aston University
BIRMINGHAM UK

College of Engineering & Physical Sciences

Assignment Brief

AM41NS – Network Science	Coursework – Part A
Roberto C. Alamino alaminrc@aston.ac.uk	

Assignment Brief/ Coursework Content:

- Theoretical and Computational Techniques
- This coursework assesses all learning outcomes on the first part of the module's content. (See *Assessment Rationale* on Module's Specification for more details.)

Descriptive details of Assignment:

- This is an individual work.
- **Justify every answer with all relevant details.** Include and mention explicitly all methods and references (when appropriate), including from every dataset you use. Failure to do so will lead to the loss of marks.
- The coursework should be submitted as one single file: a Jupyter notebook with all solutions and code in the correct order. The filename should be: [*Your_student_number*].ipynb.
- Any additional file will not be considered as part of the submission and will not be marked.
- Code in the format of a Python file instead of a Jupyter notebook will not be considered for marking.
- Your code should be properly commented, with all variables and functions clearly explained.
- The code on your notebook should run without problems or conflicts in any computer. Make sure that the libraries you use have no conflicts with other libraries. In case of doubt, stick to what has been used during the lectures.
- **All cells which do not run will be deleted**, which can have consequences to cells below them. **Therefore, be sure that all cells run without problems.**
- Text and analytical calculations should be presented as ***markdown***. Mathematical formulas should be rendered in LATEX (between $$) or typed and pasted as an image on the notebook.
- Whenever a question asks you to relate quantities to the **nature** of the

network, you should explain how they affect the functioning of the physical system represented by the network. Just describing the values without connection with the physical role of the network is not enough.

- Wherever third-party sources are used, you are obliged to include the appropriate reference. **Failure to do so might lead to a plagiarism process.**
- Analysing a network different from the one you were asked to will imply a penalty of 50% on the marks of the corresponding exercise after marking.
- Use the appropriate technical language taught during the module. Do not invent new terms, unless you carefully and rigorously define them beforehand. Do not use non-technical language to describe mathematical entities as it is not precise enough and prone to ambiguity.

Recommended reading/ online sources:

- Reading list, recorded lectures, lecture slides and any other material on the module's Blackboard page.

Key Dates:

Submission: 20/12/2024, 14:00

Feedback: 17/01/2025

Submission Details:

- Online on Blackboard

Declaration

I declare that I have personally prepared this assignment. The work is my own, carried out personally by me unless otherwise stated and has not been generated using paid for assessment writing services or Artificial Intelligence tools unless specified as a clearly stated approved component of the assessment brief. All sources of information, including quotations, are acknowledged by means of the appropriate citations and references. I declare that this work has not gained credit previously for another module at this or another University, save for permitted elements which formed part of an associated proposal linked directly to this submission. I understand that plagiarism, collusion, copying another student and commissioning (which for the avoidance of doubt includes the use of essay mills and other paid for assessment writing services, as well as unattributed use of work generated by Artificial Intelligence tools) are regarded as offences against the University's Assessment Regulations and may result in formal disciplinary proceedings.

I understand that by submitting this assessment, I declare myself fit to be able to undertake the assessment and accept the outcome of the assessment as valid.

Marking Rubric:

- The whole coursework is worth 50 marks: 40 - content, 10 - presentation
- This coursework is worth 50% of the module's total marks.

Marking Criteria for Presentation

10	Perfect presentation. Readable, well organised, professional looking.
9 - 7	Good presentation. Some issues with organisation and/or format.
6 - 4	Acceptable presentation. Several issues with organisations and/or format. Readability is somewhat compromised.
<4	Below-standards presentation. Too many issues of organisation and/or format. Difficult to read and follow.

General Advice

Presentation is a question of common sense. You should be critical about your work by putting yourself in the place of the marker and question what you would think of your presentation: how many marks would you assign to it if you were a professional educator marking it?

Points to Notice

Some particular mistakes are often repeated. The following non-exhaustive list provides advice on some of the most common points of concern:

- Solutions should be presented in the order they were asked. A different order not only compromises readability, but also shows lack of care.
- Graphs should be of professional standards: all axes should be appropriately labelled, captions should be informative, every annotation should be readable (pay attention to the size of the fonts!) and any provided graph should be referred to in the main text.
- When plotting, be sure that all relevant information is easily visualised. If necessary, change scales and adjust the range of axes.

AM41NS – Network Science

Roberto C. Alamino

Coursework – Part A

For this entire coursework:

- Use your student number as a seed when you generate random objects.
- If more than one random object is generated in the same item, for each subsequent object, add 1 to your seed.

1. (20 Marks)

For this exercise, use the network provided on the Blackboard page of this assignment.

a) (1.5 marks) Calculate the density of links in the network, defined as the number of links divided by the maximum number of possible links in a network of the same type without self or multilinks. Compared to a complete graph, would you say that this graph is well connected or not?

b) (2.5 marks) Calculate the average clustering coefficient of this graph (considered as undirected) and compare it to the average clustering coefficient taken over 200 $G(N, L)$ graphs with the same number of nodes and links. Which one is greater and what is a possible explanation for that?

c) (7 marks) Compare the in and out degree distributions of this network by plotting them and explaining their differences. Interpret your results by relating them to the nature of the network. Plot both of them and provide their average values.

d) (6 marks) Find the greatest weakly connected component of the network. For that component, plot the distribution of Katz centrality. Find how many nodes there are with the greatest and the smallest values of this centrality. Interpret your results by relating them to the nature of the network.

e) (1 mark) Explain why there are no strongly connected components of size greater than 1 in this network.

f) (2 marks) Calculate the value of assortativity by degree of the undirected version of this network and explain it according to the network's nature.

2. (20 Marks)

a) (6 marks) Show that $G(N, L)$ networks do not model well real networks. In order to do that, calculate and plot all quantities that are necessary to show it and explain the results. You will need to choose the parameters in such a way that the results you want to show are clearly identifiable.

b) (11 marks)

(i) (5 marks) Create a directed Barabasi-Albert network with 1000 nodes and $m = 2$ by modifying the original algorithm using the following rule: if a new link is from the new node i to an existing node j, the probability that this link is incoming to j is proportional to the in-degree of j.

(ii) (1 mark) Find the number of strongly and weakly connected components for the network you generated.

(iii) (5 marks) **Without using the Networkx pre-programmed function**, find the 10 most important authorities and the 10 most important hubs according to the HITS algorithm for the network's greatest strongly connected component.

c) (3 marks) Adjust a power law using linear regression to the in-degree distribution of the network from item (b) and provide the resulting value of γ writing down explicitly the resulting distribution. Only by visual inspection, does it look like a good fit?