THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Spring 2025 Exam

MY461/MY561
Social Network Analysis

Suitable for all candidates

Instructions to candidates

This paper contains five questions. Answer all questions. All questions will be given equal weight (20%). **Responses for each question should be a maximum of 400 words excluding tables and figures.** Please include a bibliography with any cited sources (this does not count towards the word count).

The exam is due on May 12, 2025 at 4pm.

Submission will be done through Moodle. Please submit your answers in a PDF file in a format similar to a report or an article. You will be evaluated based on your responses to the five prompts. You should include a subheading with the number and short title for each prompt but other than that, you should only include text, figures, tables, and references that substantiate your answer (do not include R code or unformatted output in the pdf; do not include Introduction, Conclusion, or any other additional text beyond your answers to the five prompts). To help us determine where any errors were made, you must additionally submit an annotated R or Rmd file that presents the code used to arrive at your responses.

Background information:

For this exam, we'll be focusing on a multilayer network representing relationships between lawyers in an American law firm. Every lawyer in the firm was asked to nominate people from a full list of their colleagues in response to three “name generator” questions:

1. Who did you go to for professional advice in the past year? (Advice)
2. Who do you socialise with outside of work? (Friend)
3. Who did you work with over the past year? (CoWork)

The responses to these questions form the basis of the edge lists provided. We also provide you with another file that includes information about each of the lawyers. There are 8 attributes:

1. ID: The ID for each lawyer
2. Status: (1 = partner (a more senior position); 2 = associate (a more junior position))
3. Gender: (1 = male; 2 = female)
4. Office: (1 = Boston; 2 = Hartford; 3 = Providence)
5. Seniority: number of years with the firm
6. Age
7. Practice: (1 = litigation; 2 = corporate)
8. LawSchool: (1= Harvard or Yale; 2 = University of Connecticut; 3 = Other)

Using these files, create four networks: one for each of the three relationships (advice, cowork, friend) and one aggregated network combining them together (with weighted, directed edges). With these networks in hand, answer the following questions:

1 Consider the overall metrics of the three relationship-specific networks (**advice, cowork, friend**): density, average path length, reciprocity, transitivity. How do these three networks compare to one another? Pick one network model to serve as the most suitable randomized baseline for comparison for these networks and justify your choice of model. Compare each of the three networks to its version generated by the model. What do these comparisons tell you about the nature and structure of these relationships among the lawyers? In your answer, make sure to give an intuitive interpretation for each metric and refer to important related concepts we introduced in the course.

2 What seems to dictate the structure of these networks? Calculate the assortativity for gender, age, and status in each of the **advice, cowork**, and **friend** networks. Report the results in a table. Choose a community detection algorithm to run on the **aggregated** network, justifying your choice. See how the resulting communities align (or not) with the vertex attributes of office, practice, and status. Include a plot of the aggregated network showing the results of the community detection algorithm. Discuss what the results of these two analyses imply about the nature of relationships among the lawyers.

3 What helps predict the existence of an advice-giving relationship between lawyers? Fit an ERG model on the **advice** network including terms for:

   a. the effect of age (overall activity, i.e., ignoring tie directionality)
   b. the effect of status on incoming ties
   c. the effect of status on outgoing ties
   d. gender homophily
   e. office homophily
   f. practice homophily

   Present the model results as output or in a table. Interpret each of the terms (except the `edges` term). Use the odds ratios in your substantive interpretation of each statistically significant term. Consider how well we are matching the structure of the actual network using the goodness of fit plots. What are we not getting well?

4 What helps predict the existence of a certain type of relationship between lawyers? Consider the ERG models for each of the three networks (**advice, cowork, friend**) shown below. Interpret and compare each term in the ERGMs (except `edges` and `odegree0`). Which effects differ between the different models and how? What in the nature of the relationship that is modelled could explain the difference?

Table 1. Estimates (and standard error in parentheses) of the coefficients for each term in three exponential random graph models that predict the log-odds of a tie in the (1) advice, (2) cowork, and (3) friend networks. The GWESP term has an alpha value of 0.7. The odegree0 term is a term accounting for the nodes with an out-degree of 0 in the friendship network (which can be ignored in your response). The reference category for status is 2 (associate, a more junior position).

	Advice	**CoWork**	**Friend**
Intercept (`edges`)	−5.188*** (0.244)	−5.109*** (0.296)	−4.715*** (0.216)
Age	−0.009*** (0.003)	−0.008** (0.002)	−0.007** (0.002)
Receiver's status = 1	1.080*** (0.080)	0.380*** (0.093)	0.278** (0.101)
Sender's status = 1	−0.361*** (0.092)	0.282** (0.092)	0.075 (0.104)
Same gender	0.271*** (0.071)	0.124 (0.063)	0.185** (0.064)
Same office	0.943*** (0.080)	0.799*** (0.072)	0.499*** (0.082)
Same practice	0.898*** (0.070)	0.821*** (0.064)	0.261*** (0.066)
Reciprocity	0.642*** (0.137)	2.441*** (0.127)	2.385*** (0.158)
Transitivity (`gwesp.fixed.0.7`)	1.069*** (0.083)	0.980*** (0.112)	0.945*** (0078)
`odegree0`			−0.525 (0.564)
AIC	3605	3816	2516
BIC	3664	3875	2581
Log Likelihood	−1794	−1899	−1248

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

5 Reflect on the differences, advantages, and limitations of the different analyses you conducted in the above four questions. Then, consolidate, summarize, and discuss the results from all your analyses: All things considered, what can you conclude about the network of interactions between the lawyers? Which findings extrapolate to other similar contexts? Which findings do you believe might be peculiar and specific to this company, profession, country, etc?