**Advanced Social Network Analysis 2024-2025**

Filip Agneessens, University of Trento

(version 27 Nov 2024)

**Assignment 1. Nodal level analysis**

This is the first of **two** assignments, and this assignment counts for 50% of your grade (the other assignment counts for the other 50%).

**General information**

You are asked to use the Borgatti\_Scientists504 ("Borgatti\_Scientists504.RDA") collaboration network dataset and attribute file. (See the RDA file in Moodle under Assignment 1).

- The collaboration network captures the amount of time two scientists worked together.

- The attribute dataset consists of:

1) NodeName Names of the nodes

2) Years Number of years employed

3) Sex 1=Female (N=177), 2=Male (N=327)

4) DeptID Functional department - reflects discipline

Department ID N

Behavioural scientists 1 80

Economics 2 74

Health, education, welfare 3 36

International studies 4 23

Management science 5 74

Applied sciences 6 84

Technical applications 7 108

Other 10,20,30 25

**AIM**

You are asked to calculate two “measures of position” (which could be some centrality measure or some structural holes measure that we saw in class). You can decide which two measures you want to use, but it is important to justify your choice.

The aim is to 1) explore the network, and 2) correlate the two measures of position with an assigned attribute to check if there is a relation between the position in the collaboration network and the specific attribute (note that an attribute could be binary, but nevertheless you can use the correlation coefficient to test your hypotheses). You will also need to correlate the two measures of position with each other to explore whether they might capture something similar/are overlapping.

**Your assignment**

*Part 1.* Draw the network at the **specific cutoff** (>2, >3 or >4). This is decided by the lecturer. You might be asked to only focus on a part of the network. See IMPORTANT INFORMATION below.

*Part 2.* Draw the network at the **specific cutoff** (>2, >3 or >4) with the **specific nodal attribute** decided by the lecturer. See IMPORTANT INFORMATION below.

Explain briefly what you did and how the different values for the attribute are represented. Do you see anything interesting?

*Part 3.* Calculate the two measures of position of choice using the specific cutoff (>2, >3 or >4). Justify *why* you focus on these two measures.

-Find the node(s) with the highest value for the respective measures of position. What is its/their value and explain this value.

*Part 4.* Correlation the two chosen “measures of position” with each other. What do you conclude?

*Part 5.* Correlation the respective two “measures of position” with the attribute value. What do you conclude?

**IMPORTANT INFORMATION**

You will be asked to perform the analysis at a specific cutoff (>2, >3 or >4).

You are also assigned a specific attribute.

You might also be asked to only consider a subset of whole dataset (part of the 504 scientists).

***How to get the dataset?***

The dataset is available on Moodle (Under Assignment 1)

***How to know which cutoff and attribute to use?***

To get information about which cutoff and attribute you are asked to use, as well as whether you should consider a subset of the data, please email me (the lecturer) your group composition ([filip.agneessens@unitn.it](mailto:filip.agneessens@unitn.it)) with all names and student numbers. Identify a coordinator. Please use “Advanced Social Network Analysis” in the title of the email.

**Some useful starting script**

#define your working directory:

setwd("C:/Users/Filip/Desktop/…")

#importing the data

load("Borgatti\_Scientists504.RDA")

#the network

Borgatti\_Scientists504$Collaboration

#the attribute file

Borgatti\_Scientists504$Attributes