FRTN30 Network Dynamics
Hand-in 1

Due: 2024-04-14

Instructions:

- You may implement your solutions in any language you see fit, but the TAs can only guarantee you support with MATLAB/Octave. Your code should be written in a quite general manner, i.e., if a question is slightly modified, it should only require slight modifications in your code as well. Upload a PDF of your solutions together with your code in a separate, runnable file to the hand-in under *Assignments* on the Canvas course page.
- The PDF should present a detailed solution, including a description of what you are doing, a quick summary of the theory used, proper presentation of results (including readable figures with axis labels) with interpretations. Also include your name and Lucat-id on the first page.
- Comment your code well. Clarity is more important than efficiency.
- Late submission is discouraged: you get 1 points in your exam (out of 25) for each on-time submission.
- Collaboration policy: collaboration such as exchange of ideas among students is encouraged. However, every student has to write and submit her/his own manuscript (in pdf format) and code, and specify whom she/he has collaborated with and on what particular part of the work.
- Up to five of the best hand-ins may be rewarded with one extra point.

1 Centrality in Input-Output Network of Goods

When a financial crises arises, it is important for governmental institutions to know which sector they should give financial aid. In this part of the hand-in, we will analyze the centralities of different sectors, e.g., "Food products", "Health and social work" and "Wholesale and retail trade". In our model, we let each vertex in $\mathcal{V}$ represent such a sector. The directed

links model the flow of economic activity between the sectors. The flows are only *intermediate inputs*, i.e., they only contain the sales of goods or services that are directly consumed in the production processes. Hence the input will not equal the output for each sector, since the final demands such as private consumption and net exports are not included, neither are e.g. salaries to the workers.

This part of the hand-in is based on the data from [1]. Be aware that they are using different centrality measures, so their results do not necessarily coincide with the ones you will get when applying the centrality measures discussed in the course.

The data is structured as follows: In `IOdownload.mat` there are three cells. The cell `econ` contains a list of 39 countries and which year the data is obtained. The cell `name` contains the list of the 47 sectors that are studied. The struct `io` contains the 39 input-matrices, one for each country in `econ`. For example, to get the sales from "Food products" to "Wholesale & retail trade; repairs" in France during year 2000, we type:

```
io.fra2000(4, 31)
```

Consider the countries Sweden (`swe2000`) and Indonesia (`idn2000`), find the three most central sectors for each country with respect to the following centrality measures:

a) The in-degree and out-degree centrality;

b) The eigenvector centrality on the largest connected component;

   Hint: check out the `conncomp()` matlab function

c) The Katz centrality, with $\beta = 0.15$ and with two different values on μ, namely $\mu = \mathbb{1}$ and when $\mu_i = 1$ for the "Wholesale & retail trade; repairs" sector and zero for all other sectors.

Comment your results briefly and feel free to do other investigations of the data that you find interesting.

2 Influence on Twitter

For this second part of the hand-in, we have extracted a subgraph of the X (formerly Twitter) network for you to work with. To obtain the data, we started to crawl the network from a subset of one user's followers, then the followers' followers and so on[1]. The file `twitter.mat` contains a three-column table with links, where the first two columns represent the tail and the head of the link, and the third column the weight of the link. Here a link (i, j) represents that i follows j. The adjacency matrix can be loaded into MATLAB using the following code:

[1]We stopped after approximately 6000 users, and as you might discover, this crawling method has its draw-backs when determining the PageRank.

```
load -ascii twitter.mat
W = spconvert(twitter);
```

The command `spconvert()` creates a *sparse* matrix from an input table, which is beneficial when working with larger matrices with many zero elements. It is recommended to continue using sparse matrices in this part, by e.g. using the `sparse()` command when creating new matrices.
The file `users.mat` contains on the ith row, the Twitter user id of node i (this can then be converted to the corresponding username by e.g. `http://tweeterid.com/`).

a) Compute *iteratively* the PageRank with $\beta = 0.15$ and $\mu = \mathbb{1}$, find the five most central nodes.

b) Simulate the discrete-time consensus algorithm with two stubborn nodes, e.g., one with value 0 and one with value 1. You can initiate the rest of the nodes to the neutral value 0.5. Plot how the opinions change over time for a handful of nodes of your choice.

c) Investigate how the the choice of nodes with respect to their PageRank change the stationary opinion distribution. Plot the opinion distribution for a couple of choices of the stubborn nodes.

Hint: you can use the matlab command `hist()` to plot the histogram.

References

[1] Florian Blöchl, Fabian J Theis, Fernando Vega-Redondo, and Eric O'N Fisher. Vertex centralities in input-output networks reveal the structure of modern economies. *Physical Review E*, 83(4):046127, 2011.