Network Dynamics I

Erik Waldemarson

1 Centrality in Input-Output Network of Goods

In this section, I will examine the most central financial sectors of the countries Sweden and Indonesia, with respect to three different centrality measures.

1.1 In-degree and out-degree centrality

The in-degree and out-degree centralities are simply the nodes with the highest (weighted) number of ingoing and outgoing neighbours, respectively. More formally for a network with weight matrix W, the sectors with highest centralities are defined as

$$\text{Sector max in-degree centrality} = \operatorname{argmax} w^{-} = \operatorname{argmax} W'\mathbb{1},$$
$$\text{Sector max out-degree centrality} = \operatorname{argmax} w = \operatorname{argmax} W\mathbb{1},$$

where w^{-} is the in-degree vector and w is the out-degree vector.

The top three most central sectors for each country have been gathered in Table 1.

Table 1: Three most central sectors (in- and out-degree centrality) for Sweden and Indonesia.

Number	SWE		IDN	
	in-deg cent.	out-deg cent.	in-deg cent.	out-deg cent.
1	19 Radio	43 Other Business Activities	4 Food products	31 Wholesale & retail traide; repairs
2	21 Motor vehicles	39 Real estate activities	30 Construction	1 Agriculture
3	43 Other Business Activities	31 Wholesale & retail trade; repairs	31 Wholesale & retail trade; repairs	2 Mining and quarrying (energy)

1.2 Eigenvector centrality of largest connected component

The eigenvector centrality z is found by solving the equation

$$z = \frac{1}{\lambda_W} W' z,$$

where λ_W is the dominant eigenvalue of W, i.e. the eigenvalue with largest magnitude value. This is an eigenvalue problem with a guranteed $z \neq \mathbf{0}$ solution since W is a nonnegative square matrix (according to Perron-Frobenius theorem).

I begin by finding the most connected component in the graph. I then create a new weight matrix W_{Comp} for that component. I then find the three most central sectors for each country by calculating each z for W_{Comp} and then finding the top three $\operatorname{argmax} z$. The results have been collected in Table 2.

Table 2: Three most central sectors (eigenvector centrality) for the largest components in Sweden and Indonesia.

Number	SWE	IDN
1	15 Fabricated metal products	17 Office
2	28 Steam and hot water supply	20 Medical
3	5 Textiles	23 Aircraft & spacecraft

1.3 Katz centrality

Finding the Katz centrality z means solving the equation

$$z = \frac{1-\beta}{\lambda_W} W' z + \beta \mu,$$

for some choice of scalar $1 > \beta > 0$ and vector $\mu \neq \mathbf{0}$. Solving for z yields the formula

$$z = (I - \frac{1-\beta}{\lambda_W} W')^{-1} \beta \mu.$$

Does a $z \neq \mathbf{0}$ exist? The answer is yes. The matrix $(I - \frac{1-\beta}{\lambda_W} W')$ is invertible iff it has only non-zero eigenvalues. The spectral radius of $-(1-\beta)/\lambda_W W'$ is simply $\rho = (1-\beta)$ according to the Perron-Frobenius theorem. Adding I will shift every eigenvalue one-step right along the real line (+1 to the real part). This means that as long as $(1-\beta) < 1$, no eigenvalue can ever be equal to 0 since ρ is not large enough (all eigenvalues lie within a circle in the complex plane with center at (1,0) and radius ρ). Thus the matrix is always invertible and there exists a solution $z \neq \mathbf{0}$.

I will solve for the Katz centrality using $\beta = 0.15$ and for two different μ. The first $\mu^{(1)} = \mathbb{1}$, and the second $\mu^{(2)}$ is defined as $\mu^{(2)}_{31} = 1$ for sector "31 Wholesale & retail trade: repairs" and $\mu^{(2)}_i = 0$ for all others.

The three most central sectors using Katz centrality for each μ and country have been gathered in Table 3.

Table 3: Three most central sectors (Katz centrality) for Sweden and Indonesia. Here $\beta = 0.15$ and $\mu^{(1)} = \mathbb{1}$ and $\mu^{(2)}_i = 1$ for $i = 31$ and 0 for all others.

Number	SWE		IDN	
	$\mu^{(1)}$	$\mu^{(2)}$	$\mu^{(1)}$	$\mu^{(2)}$
1	21 Motor vehicles	31 Wholesale & retail trade; repairs	4 Food products	4 Food products
2	19 Radio	21 Motor vehicles	32 Hotels & restaurants	31 Wholesale & retail trade; repairs
3	43 Other Business Activities	19 Radio	1 Agriculture	32 Hotels & restaurants

1.4 Discussion

It's clear from the results from the previous sections that the choice of centrality and choices of intrinsic centrality can have a large impact on the most central sectors and should thus be chosen carefully.

2 Influence on Twitter

In this part I take a look at a subgraph of the social network X (formerly Twitter).

2.1 Iterative PageRank

The aim is to calculate the PageRank centrality of a graph, which is equivalent to solving the equation

$$z = (1 - \beta)P'z + \beta\mu,$$

where $P = \text{diag}(w)^{-1}W$ is the stochastic matrix of the graph. Note that if any $w_i = 0$, i.e. a node has no outgoing neighbours, then diag(w) is not invertible. However this problem can be solved by simply adding a self-loop to node i, e.g. $W_{ii} = 1$ without loss of generality.

Similarly as before I solve for z and get the formula

$$z = (I - (1 - \beta)P)^{-1}\beta\mu.$$

By arguments similar to the one in section 1.3 the matrix is invertible since P is square nonnegative and the dominant eigenvalue $\lambda_P = 1$. However, instead of solving for the inverse I will solve this iteratively. Note first that for any matrix Q you can write the power series

$$(I - Q)^{-1} = \sum_{k=0}^{\infty} Q^k,$$

assuming that $(I - Q)$ is invertible. This is just the power series for $1/(1 + x)$ applied to matrices. Thus I finally get the equation

$$z = \sum_{k=0}^{\infty}(1 - \beta)^k P^k \beta\mu.$$

For the pratical implementation I use sparse matrices to save memory. I noted that W was not initially a square matrix. This is because of how the data was obtained, which was by crawling through the network by one user's followers, then the followers' followers and so on. The result is that when stop at the last nodes, we have no idea how many followers those nodes have. This means that there are coloumns missing in the matrix (representing ingoing neighbours). I solved this by making the simplyfing assumption that all the last nodes were sources, and filled the rest of W with zero columns until it was square. I also added self-loops to W where necessary (which happened to be only node 1).

I then went ahead and solved for the PageRank centrality using $\beta = 0.15$ and $\mu = \mathbb{1}$. I iterate through the sum until I reach some k for where $||(1-\beta)^k P^k \beta\mu||_2 < 10^{-6}$. The five most central nodes were (in descending order) 1, 2, 112, 9 and 26.

2.2 Simulation of stubborn nodes

In this part I simulate an the graph as an opinion dynamics network with some stubborn nodes u. A stubborn node is a node that has a constant value over time. These can be interpreted as users that never change their opinion, i.e. 'stubborn' (very common on Twitter/X).

If for some time t we have the values of each node in a vector $x(t)$, we can separate this vector into one part with non-stubborn and stubborn nodes $x(t) = [\underline{x}(t) \quad u]'$. In a similar manner we can divide P into a matrix for non-stubborn nodes Q and one for stubborn nodes E.

The following equality holds

$$\underline{x}(t+1) = Q\underline{x}(t) + Eu.$$

I simulate this equation setting the stubborn nodes to be $[80 \quad 1]'$ with values $u = [0 \quad 1]$ for 1000 iterations. I used $\beta = 0.15$ and $\underline{x}(1) = 0.5 \cdot \mathbb{1}$. The result is shown for some selected nodes in Figure 1.

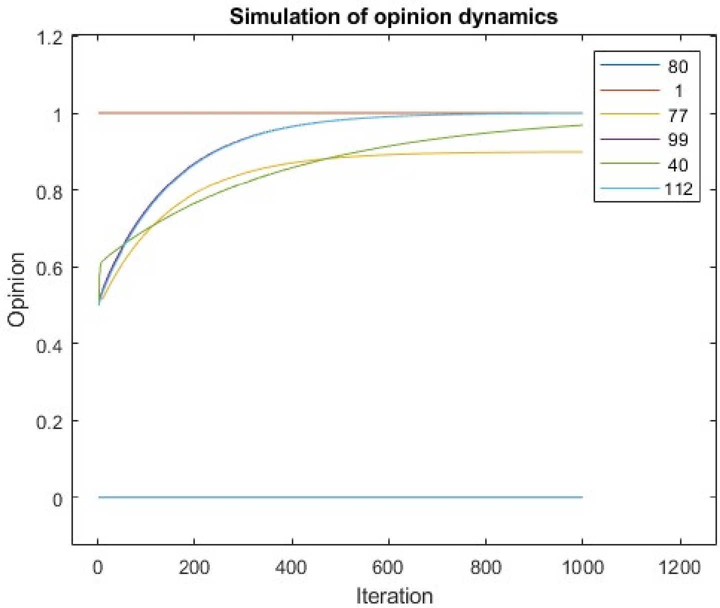

Figure 1: Simulation of node 80, 1, 77, 99, 40 and 112 of Twitter network with stubborn nodes $[80 \quad 1]'$ with values $u = [0 \quad 1]$ for 1000 iterations with $\beta = 0.15$ and $\underline{x}(1) = 0.5 \cdot \mathbb{1}$.

As one can see, node 1 always has value 1, while node 80 has value 0, and the remaining ones starting out with an average in-between, around 0.4, but increases towards a value of 1. This is because I purposefully selected a node with high PageRank centrality (node 1) and the other one was just random (80). It seems that central nodes with high centrality will start to dominate the network, however I will explore this in more detail in the next section.

2.3 Stationary opinion distribution

In this part I continue with the same opinion dynamics simulation as in the previous section, however now I am interested in seeing how the distribution of values changes depending on the choice of stubborn nodes. In every experiment I always had the most central (PageRank) node 1. I experimented to see it affected a random node 80, and what happened if I changed their values. I then compared the most central node to the second and third most central nodes, 2 and 112.

The resulting probability distributions of each value was then collected been collected in Figures 2-6.

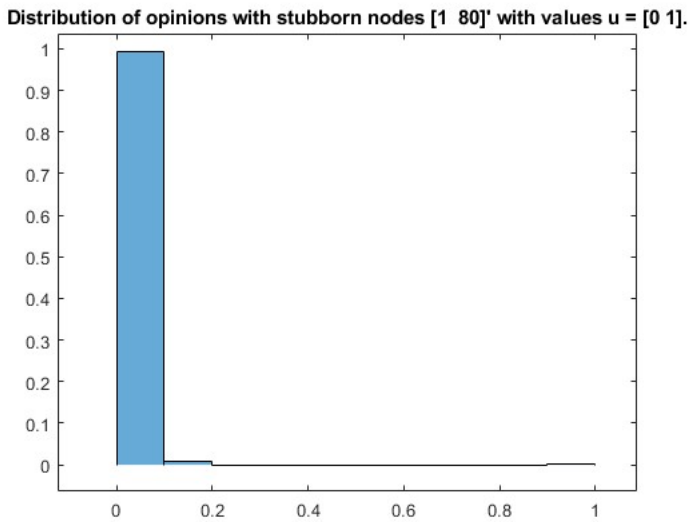

Figure 2: Probability distribution of value (opinion) of nodes (users) in Twitter network. Stubborn nodes are $[1 \quad 80]'$ with values $u = [0 \quad 1]$ using 1000 iterations with $\beta = 0.15$ and $\underline{x}(1) = 0.5 \cdot \mathbb{1}$.

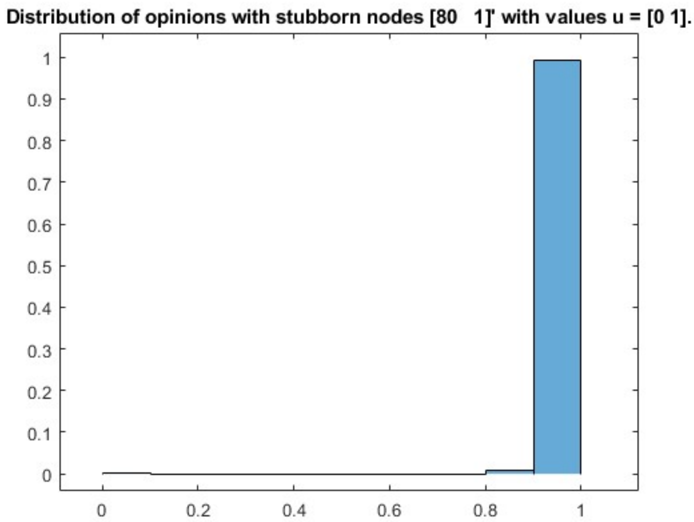

Figure 3: Probability distribution of value (opinion) of nodes (users) in Twitter network. Stubborn nodes are $[80 \quad 1]'$ with values $u = [0 \quad 1]$ using 1000 iterations with $\beta = 0.15$ and $\underline{x}(1) = 0.5 \cdot \mathbb{1}$.

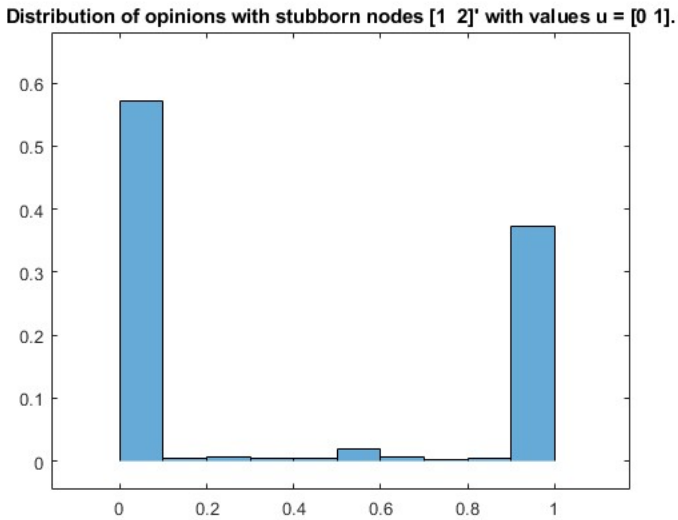

Figure 4: Probability distribution of value (opinion) of nodes (users) in Twitter network. Stubborn nodes are $[1 \quad 2]'$ with values $u = [0 \quad 1]$ using 1000 iterations with $\beta = 0.15$ and $\underline{x}(1) = 0.5 \cdot \mathbb{1}$.

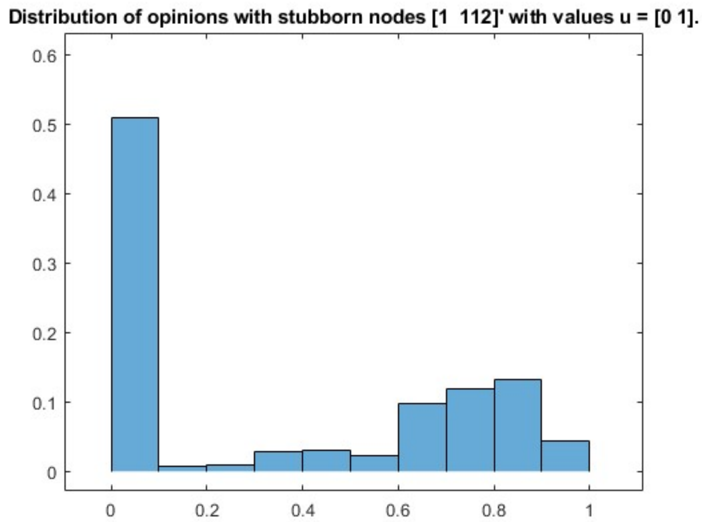

Figure 5: Probability distribution of value (opinion) of nodes (users) in Twitter network. Stubborn nodes are $[1 \quad 112]'$ with values $u = [0 \quad 1]$ using 1000 iterations with $\beta = 0.15$ and $\underline{x}(1) = 0.5 \cdot \mathbb{1}$.

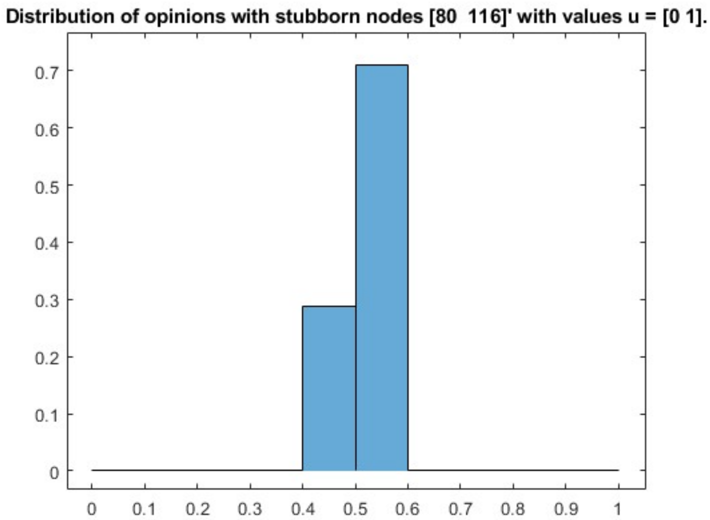

Figure 6: Probability distribution of value (opinion) of nodes (users) in Twitter network. Stubborn nodes are $[80 \quad 116]'$ with values $u = [0 \quad 1]$ using 1000 iterations with $\beta = 0.15$ and $\underline{x}(1) = 0.5 \cdot \mathbb{1}$.

It's clear from these results that nodes with high centrality will (given enough time) dominate the opinion dynamics. If two stubborn nodes have opposite opinions, then the opinion network will divide into two camps of opposing opinions. Also the degree of centrality matters a lot. Just comparing Figures 4 and 5 shows a big difference with just a slight difference in centrality. When the stubborn nodes were just random (and likely not very central) the opinions were usually around the intial value.

The interpretation is that stubborn users with a large number of followers can easily have a big impact on the opinion dynamics and potentially dominate the entire network.