Problem Set n.3

Clinton Leung, 200870
Dante Gustav Jarhos, 200821
Gian-Piero Lovicu, 200752
David Ampudia Vicente, 82929

Optimal Seeding

Question 1. Consider the model we saw in class where a household with λ informed connections will choose to become informed if

$$\frac{\alpha^\lambda}{\alpha^\lambda + (1-\alpha)^\lambda} > \frac{\tilde{c}_j}{\tilde{\pi}}$$

for some $\alpha > 0.5$ and $\frac{\bar{c}}{\pi}$ the cost to benefit ratio of the new technology. The above equation essentially means that each farmer has an adoption threshold τ_i which determines the minimum λ such that she decides to adopt.

(a) For $\lambda = 1$ and $\lambda = 2$, conduct technology diffusion simulations for your village, where each individual draws an adoption threshold τ from a $\mathbb{N}(\lambda, 0.5)$, but truncated to be strictly positive.

We choose a network of Mali Terrorists taken from here.

(b) Provide the identity of the optimal pair of seeds under $\lambda = 1$ and $\lambda = 2$

Optimal seeding chooses nodes 14 and 25 for $\lambda = 1$. For $\lambda = 2$, the optimal seeds are 8 and 18.

(c) Now run the same Monte Carlo simulation as in part (a) but using randomly chosen seeds. Compare the rates of information diffusion with the optimal seeds in part (a) and discuss the results.

When $\lambda = 1$ this corresponds to a simple diffusion model, with the threshold for adoption distributed normally around 1 adopting neighbour for each node. On the other hand, when $\lambda = 2$ this simulates a complex diffusion process, with the threshold for adoption increasing to be normally distributed around 2 adopting neighbours. The simulations show that for simple diffusion ($\lambda = 1$), random seeding underperforms optimal seeding by around 18 percent. In contrast, under complex diffusion random seeding is 50 per cent worse.

For $\lambda = 1$ with optimal seeding around 28 nodes had adopted after three periods. In contrast, with random seeding the average level of diffusion was lower, but not substantially at around 23 nodes. This supports the idea of random seeding from the Akbarpour paper - with only a few more seeds, random seeding could have achieved almost the same amount of diffusion as the optimal seeds.

For $\lambda = 2$, optimal seeding resulted in adoption by 8 nodes compared to around 4 nodes for random seeding. In general, the complex diffusion process results in a much lower rate of adoption than for simple diffusion. This makes sense because the threshold for adoption is twice as high. With such a low rate of diffusion a few more seeds placed randomly would have achieved a similar result to optimal seeding, though the small number of periods makes it difficult to know whether more iterations would have led to a higher rate of diffusion under optimal seeding.

The optimal seeds correspond to larger eigenvector centralities (Figure 1) . The optimal seeds do not necessarily represent the highest eigenvector centralities in the graph and this means that strategically placed nodes (with not necessarily highest eigenvector centrality) may have been more effective at inducing adoption. Figure 2 shows the betweeneess centrality of the optimal seeds. The optimal seeds are not necessarily the highest in the graph and does not give a strong indicator that the betweeness centrality is that important for seeding aside from the fact that you don't want to be an end node that has zero or very low betweeness centrality.

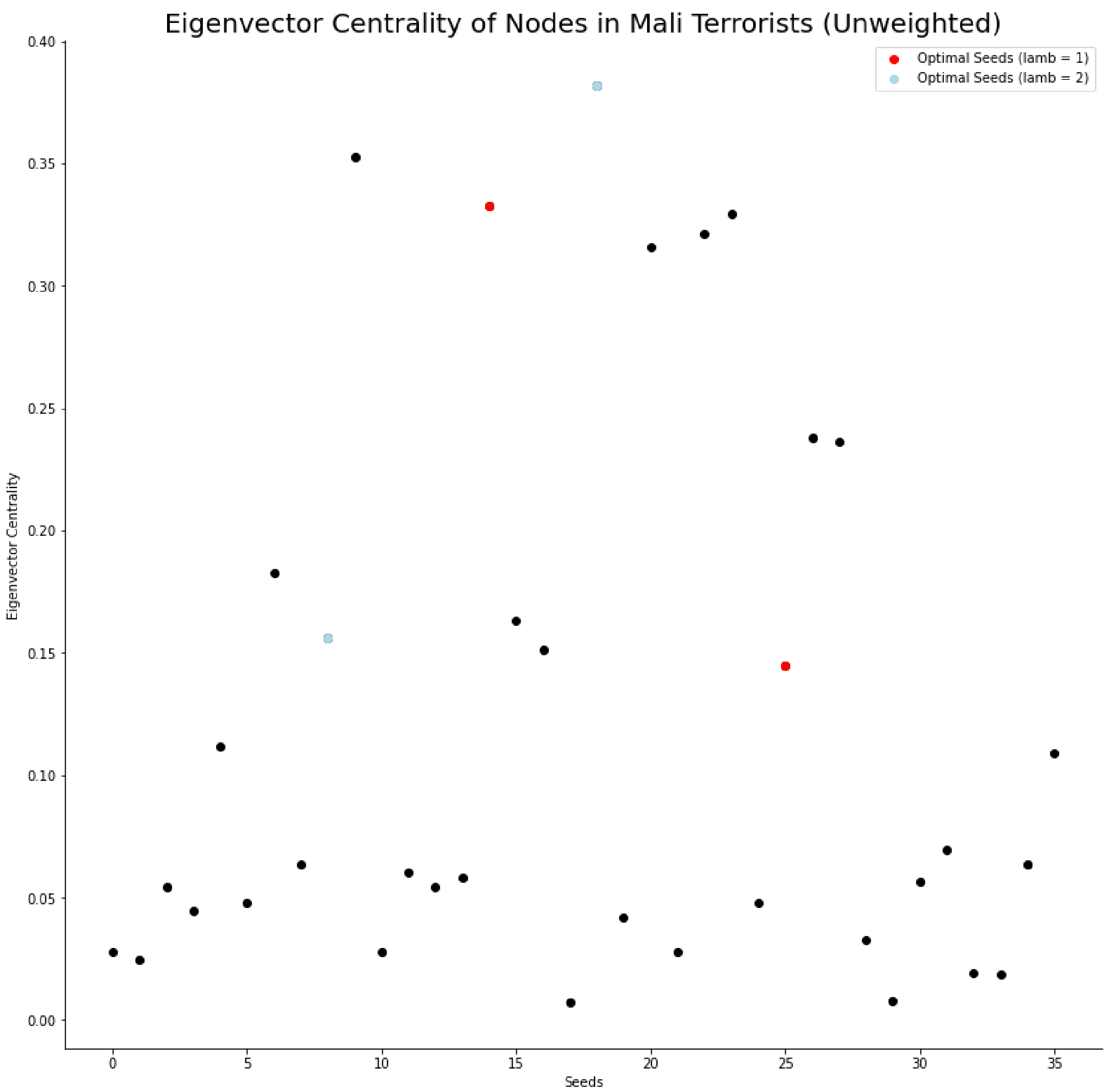


Figure 1: Eigenvector centralities of optimal seeds in Mali terrorist network

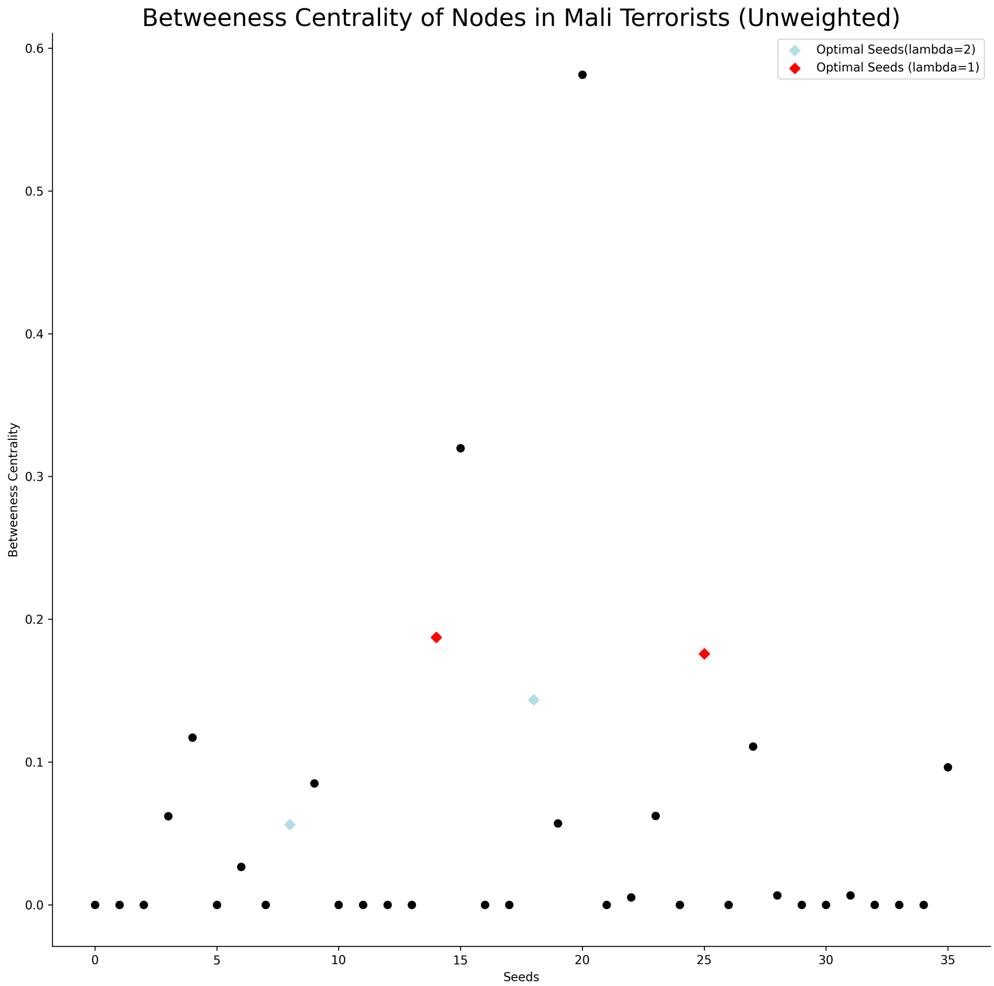


Figure 2: Betweeness centralities of optimal seeds in Mali terrorist network

Figure 3 shows the positions of the optimal seeds in the graph for $\lambda = 1$ in red and $\lambda = 2$ in blue.

For $\lambda = 1$ other nodes with high eigenvector centrality (18, 20) are located near the optimal seeds (14 and 25). However, it more efficient to seed from 14 and 25 than from 18 and 20, because these seeds have a shorter geodesic path **within** (important!) each of the two main communities (split at node 20) than nodes 20 or 18. For example, the geodesic distance from each of the optimal seeds to all other nodes on their side of the graph is 3 steps. For nodes 18 and 20 it is 4 steps (assuming node 20 is in both communities and node 18 is in the LHS community). Figure 4 shows why this is important in the simple diffusion process - the seeds are located to try to foster adoption among more of the fringe nodes. Altogether, 32 nodes have adopted out of a possible 36 with the remaining nodes at the fringe of the graph.

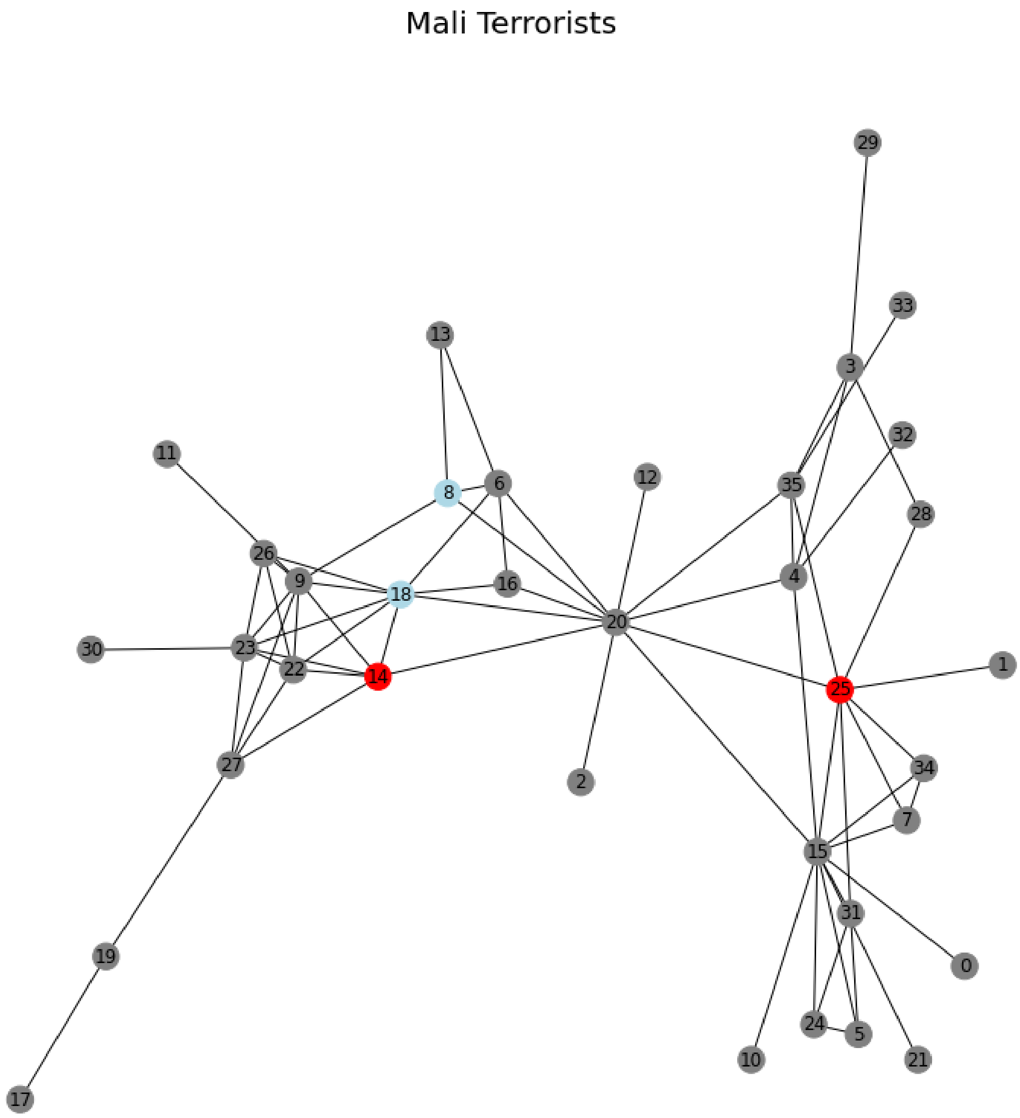


Figure 3: Location of the initial optimal seeds

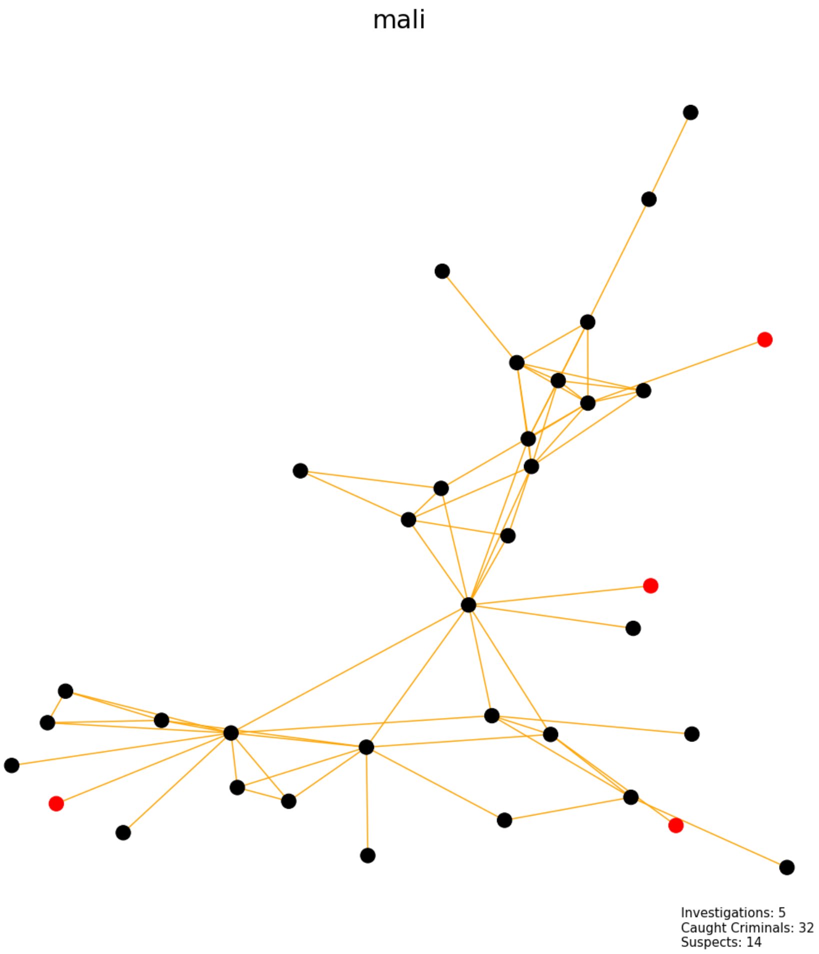


Figure 4: Final graph for $\lambda = 1$ (red corresponds to nodes that did not adopt, all other nodes successfully adopted.

For $\lambda = 2$, the optimal seeds are located in one cluster of the graph and have common neighbours. It makes sense that the optimal seeds are located close together, since a higher threshold of adoption is needed to induce adoption among common nodes. It also explains why overall diffusion is so much lower than the simple diffusion process - over 3 periods the the diffusion process does not have the ability to make the jump from one side of the graph to the other (say through node 20).

Social Learning

2. Consider the following variation of DeGroot's opinion updating model. At time 0 , the opinions of the agents are given by a column vector p_0 (assume that to begin with they all have different opinions). Let G be a standard DeGroot updating matrix and assume that the associated network is aperiodic and strongly connected. Now suppose that the updating process

(in our modified model) is as follows:

$$p_{ti} = \lambda_i \left(Gp_{t-1}\right)_i + (1-\lambda_i)\, p_{0i}, \text{ where } \lambda_i \in (0,1) \text{ for all } i$$

An interpretation is the following: First the agent takes the usual weighted average of last period's opinions, but then he always goes home and reads up on his original opinion. So that ultimately he places a weight of λ_i on the "current trend" and $(1-\lambda_i)$ on his initial opinion. What happens to the opinions of the agents in the long-run? Does each agent's opinion converge to something? In general, do they reach a consensus? Be as precise as possible in your answer.

Let p_T be the column vector of opinions at some finite terminal time T, and define λ as the column vector of the individual λ_i. Then, the updating process reads

$$p_T = \lambda G p_{T-1} + (1-\lambda)\, p_0$$

which is a recursive formula over the individual's previous beliefs. Hence, we can rewrite it as

$$\begin{aligned} p_T &= \lambda G \left[\lambda G p_{T-2} + (1-\lambda)\, p_0\right] + (1-\lambda)\, p_0 \\ &= G^2\lambda^2 p_{T-2} + G\lambda\,(1-\lambda)\, p_0 + (1-\lambda)\, p_0 \\ &= G^2\lambda^2 \left[\lambda G p_{T-3} + (1-\lambda)\, p_0\right] + G\lambda\,(1-\lambda)\, p_0 + (1-\lambda)\, p_0 \\ &= G^3\lambda^3 p_{T-3} + G^2\lambda^2\,(1-\lambda)\, p_0 + G\lambda\,(1-\lambda)\, p_0 + (1-\lambda)\, p_0 \end{aligned}$$

This process can be repeated until the left-most vector of beliefs also becomes the original vector, at which point the expression reads

$$\begin{aligned} p_T &= G^T\lambda^T p_0 + (1-\lambda)\sum_{j=0}^{T-1} G^j\lambda^j p_0 \\ &= \left[G^T\lambda^T + (1-\lambda)\sum_{j=0}^{T-1} G^j\lambda^j\right] p_0 \end{aligned}$$

As $T \to \infty$, ie. over the long run, the expression above becomes

$$\lim_{T\to\infty} p_T = \left[G^\infty\lambda^\infty + (1-\lambda)\sum_{j=0}^{\infty} G^j\lambda^j\right] p_0$$

where the right-hand side geometric series has a closed form solution

$$\lim_{T\to\infty} p_T = \left[G^\infty\lambda^\infty + \frac{1-\lambda}{1-G\lambda}\right] p_0$$

Importantly, three possible scenarios suggest different there may be multiple long-term equilibria.

First scenario: individuals never read their original opinions, ie. $\lambda = [1,\ 1,\ \dots\ 1]$. This implies

$$\lim_{T\to\infty} p_{T|\lambda=1} = G^\infty p_0$$

This result is identical to the unmodified DeGroot opinion updating model. Since the associated network is aperiodic and strongly connected, it follows that agents' opinions will converge to a weighted average of the original opinion vector. Hence, they will reach consensus in a finite time frame.

Second scenario: for at least some individuals, $0 < \lambda_i < 1$. These agents listen to themselves in each update iteration, and their long-run probability is as follows[1]

$$\lim_{T\to\infty} p_{T|0<\lambda<1} = \frac{1-\lambda}{1-G\lambda} p_0$$

In this alternative scenario, long-run beliefs will be stable at some weighted average of the original beliefs. However, these beliefs will not become homogeneous across members of society, so no consensus will materialize. Instead, a smaller λ_i will increasingly bias long-run beliefs towards individual's own original opinions.

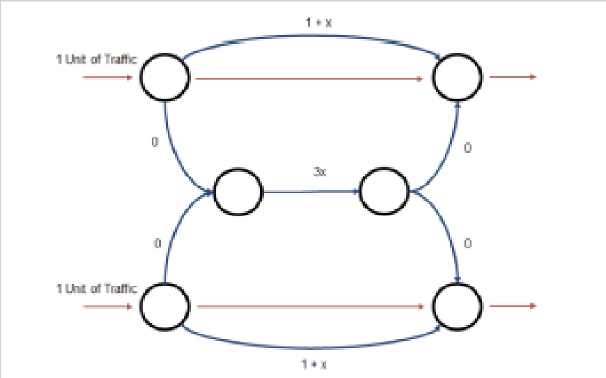
Third scenario: one where $\lambda_i = 0$ for all $i \in \{1, 2, \cdots, n\}$. Here, individuals exclusively listen to themselves, so trivially

$$\lim_{T\to\infty} p_{T|\lambda=0} = p_0$$

All in all, because the associated network is aperiodic and strongly connected, which is true even if individuals listen to themselves since these are cycles of length 1, opinions are stable over the long-run. However, only in scenarios where $\lambda_i = 1$ for all individuals does convergence across society occur. Instead, smaller values of lambda create an incentive to stick to one's own original opinions, and in the degenerate case where $\lambda = 0$ society's beliefs are always their own original thoughts.

Routing on roads

3. Consider the traffic flow game pictured in the figure below. There are two origin-destination pairs. A unit of traffic needs to flow from the upper left to the upper right, and another unit needs to flow from the lower left to the lower right. The cost functions are given in the figure.



(a) What is the social optimal routing, and what is its total cost?

See the code attached that solves the constrained optimisation problem presented by this network.

The socially optimal outcome is for 0.786 to travel along the top and bottom edges and for 0.429 to travel along the middle edge. The overall cost of this route is 3.357.

[1] We assume without loss of generality that all λ_i are less than 1, in order to keep consistent with the vector notation.

(b) What is the equilibrium routing? What is the welfare loss relative to the optimum?

The equilibrium routing is where there is 0.714 flow going through top and bottom edge each and a flow of 0.571 going through the middle edge. This has an overall cost of 3.429 which is higher than the social optimum cost. This was once again solved using the code provided to model the constrained optimization problem after adding in constraints that set the costs of all available paths to be equal for each agent.

We can get the welfare loss by subtracting the social optimum cost from the general equilibrium cost and we can get the price of anarchy by the ratio for the two. This implies a welfare loss of 0.072 and a price of anarchy of 1.021.

(c) Suppose you can impose constant tolls on some edges and constant subsidies on others. Design a system of tolls and subsidies to implement the social optimum as an equilibrium.

To calculate the tolls we need to solve a system of equations that ensures the agents actions are in equilibrium and that the balance of payments is even. This amounts to three equations:

$$
\begin{aligned}
c_1x_1 + c_2x_2 + c_3x_3 = 0 \\
1 + x_1 + c_1 = 3x_2 + c_2 \\
1 + x_3 + c_3 = 3x_2 + c_2
\end{aligned}
$$

Here c, indicates the toll cost, and x indicates the flow of each edge (subscripts are in order from top edge to bottom edge). The first equation keeps the balance of payments even. The second and third equations keep the agents' actions in equilibrium once you add the tolls. We can consider $x_3 = x_1$ and $c_1 = c_3$ due to the symmetry of the problem to get a solution which turns out to be:

$$
\begin{aligned}
c_1 &= \frac{x_2(3x_2 - 1 - x_1)}{2x_1 + x_2} \\
c_2 &= \frac{-2c_1x_1}{x_2}
\end{aligned}
$$

Since we are using these tolls and subsidies to achieve the flows in the social optimum we can substitute those values of $x_1 = 0.786$ and $x_2 = 0.429$ and we get that for both the top edge and the bottom edge there is a subsidy of -0.107 and for the middle edge there is a toll of 0.393. We can indeed verify that, up to rounding, this has a balance of payments that is even. The flow of the top and bottom edges are 0.786 and the flow in the middle edge is 0.429. We verify the results once again in Gurobi.

We used Gurobi optimizer for this homework because we are already familiar with its use from our optimization class and this meant we didn't have to compute this by hand or risk making small math errors. Unfortunately, you might not have it installed to check our script. It is possible to get a free academic license (what we have) here. I have also included the .lp files that show the optimization model we wrote in case you don't want to spend the time to install.