In today’s exercise, we will simulate social dynamics in the networks. In the morning, we will study the dynamics of simple contagion in social networks. Specifically, we will build an independent cascade (IC) model and study the role of weak ties in simple contagion.

In the afternoon, we will focus on complex contagion and the threshold model. We will explore how to find the best ‘seeds’ in a social network.

The network we use today is based upon the Twitter network data. A small-world like network is developed based upon the community structure from the Twitter network data. Please download the “day5\_socialdynamics\_Networkdata.csv” for the network data. Codes are provide in R, “day5\_socialdynamics\_Rcodes.r”.

For each question, discuss the answer within your group and write down the answers for Plenary Discussion. For some questions, save the number or figure if they support your answer.

Computer lab, 10:00-11:30:

Follow the instructions in the r file to build Network 1. Visualize Network 1 (by calling “VisNetwork1\_layout” in R). Can you find out 3 weak ties and 3 strong ties in this network? Record the node IDs constituting these 6 edges. And explain the criteria that you have used to identify weak ties and strong ties. **[Q1]**

Next, we will test “the strength of weak ties in simple contagion”. First, build a simple independent cascade (IC) model with the following characteristics:

* Each node in the network has two statuses: infected (value=*1*) or healthy (value =*0*);
* At Day 0, all the nodes in the network are healthy;
* At Day 1, an infected node (*N0*, node ID= 5) is introduced to the network;
* At the following days, all the nodes connecting to an infected node will have a chance of 0.35 (*p*=0.35) being infected.
* Once infected, the node will remain contagious and infected until the end of simulation.

After you build the model, apply it to the Network 1, and record the number of newly infected people by day (i.e., newly confirmed cases by day) for further analysis. **[Q2]**

Based upon Network 1, delete the 3 WEAK ties that you have identified to generate a new network (Network 2). Apply the IC model to the Network 2, and record the number of newly infected people by day for further analysis. **[Q3]**

Based upon Network 1, delete the 3 STRONG ties that you have identified to generate a new network (Network 3). Apply the IC model to the Network 3, and record the number of newly infected people by day for further analysis. **[Q4]**

Generate a plot (Plot\_p0) to compare the daily infected percentages of Network 1, Network 2 and Network 3. What do you observe? Considering the fact that the amount of edges we deleted in Network 2 and 3 account for around 2% of the total edges in Network 1, what is the implication of your observation? **[Q5]**

Try to explain the your observation in Q5 in more details as a network scientist-- by looking at the change of network structure. To do so, we suggest you to check the following network attributes of Network 1, Network 2 and Network 3: 1) Clustering coefficient, 2) Diameter, and 3) Average path length (average of the shortest paths between all pairs of nodes). **[Q6]**

In the above exercises, we assume the probability *p* in the IC model is 0.35, i.e., there is a 35% chance for a healthy person to get infected after contact with an infected person. Do you think your observation in Q5 holds regardless of the contagiousness of the virus? To find out, play around the probability *p* in the IC model. Change the value of *p* to high and low ends, run the IC model again on Network 1, Network 2 and Network 3, and see if you will observe different things. **[Q7]**

Plenary Discussion, 12:00-12:30:

Inspired by the above exercises, can you relate some of the modeling results to the COVID measures implemented in the real world?

Computer lab, 15:00-16:30:

In the lecture, we did a small survey on your ‘threshold’ to restrict your beef consumption to once a week or lower. Here, let’s build a threshold model to simulate the spread of “once-a-week-beef” campaign in Network 1. Below is the model description:

* Each node in the network has two status: adopt the behaviour (value=*1*) or refuse to adopt (value =*0*);
* At Day 0, *no one* in the network adopts the behaviour;
* At Day 1, a group of enthusiastic people (seed nodes, *Ns*, Node ID={1, 2, 5, 11}) in the network decided to take action and adopt their own behaviours (change the status value to 1);
* At the following days, for nodes who haven’t adopted ((value=*0*), they will check the status of their neighbours to decide to adopt or not:
* For example, for node *i*, it has a predefined threshold of *Ɵi*, as defined in the *r* file. And node *i* has *Ni* number of neighbours in the whole network.
* Among all the neighbours of Node *i*, if more than *Ni\* Ɵi* of them have adopted, node *i* will also adopt and change the status value to *1*.
* Once adopted, the status of this node will remain as *1* till the end.

Codes for the threshold model is provided in the r file. Apply the threshold model to Network 1. By activating 4 nodes (id=1, 2, 5, 11), how many people in the network can be activated? **[Q8]**

Recall the diffusion of virus we simulated in the morning: A network-wide contagion has been observed with only one node (id=5) being activated at Day 0, for various values of *p*. Use the theory of “width of bridge” in complex contagion, explain why a network-wide contagion is not happening here. **[Q9]**

Imagine now the project team of the “once-a-week-beef” campaign plans to increase the membership by spending part of the budget to hire four “ambassadors”. The team leader suggests that the ambassadors should be the four people having the most number of friends in this network. Recall the Influence Maximization problem we discussed in the lecture and its solutions, can you relate this idea to one of the solutions we mentioned? Guided by this idea, which nodes should be selected as ambassadors? What is the expected outcome suggested by the threshold model? (Note: All nodes are Inactive before the ambassadors are hired, i.e., node 1, 2, 5 and 11 are not the seeds anymore) **[Q10]**

The project team now seeks advice from your group on the “best” ambassadors they should choose. If the cost to hire any people in this network as the ambassador is the same, whom will you suggest?

1. To answer this question, run the greedy algorithm and see if the solution is different from the seeds suggested by the team leader. (Codes for the greedy algorithm is provided in the r file.)
2. What are the node attributes of the seeds provided by the greedy algorithm? Are they high-degree or high-betweenness nodes?
3. If you have time, try to explore other heuristics and see if any one of them can provide better outcomes than the greedy algorithm. **[Q11]**

Plenary Discussion, 16:30-17:00:

Reflect on the common marketing or vaccination strategies in the real world, after today’s program, do you think you can propose a better strategy?