

Assignment #4 – Epidemics

Chris J. Riederer, Zhikun Ma (TAs)

A. Chaintreau (instructor)

Why there are three parts in this assignment: Each part fulfills one of the objectives of the class:

- **Manipulate concepts:** Getting Familiar with the technical concepts used in class, by reproducing similar arguments. Being proficient by manipulating the object to answer some small-size problem.
You are expected to answer this question rigorously, the answer can be quite short as long as it contains all the required argument to justify your answer.
- **Experience the concepts :** Being able to reproduce these concepts in real or synthetic data. Study their properties in real examples.
- **Connect the concepts to real-life:** Interpret a problem you find in light of the concepts or principles you have learned. Develop a critical eye to determine how the concepts introduced are useful in practice.

How to read this assignment : Exercise levels are indicated as follows

(\rightarrow) “elementary”: the answer is not strictly speaking obvious, but it fits in a single sentence, and it is an immediate application of results covered in the lectures.

Use them as a checkpoint: it is strongly advised to go back to your notes if the answer to one of these questions does not come to you in a few minutes.

(\curvearrowright) “intermediary”: The answer to this question is not an immediate translation of results covered in class, it can be deduced from them with a reasonable effort.

Use them as practice: how far are you from the answer? Do you still feel uncomfortable with some of the concepts and definitions? which part could you complete quickly?

(\curvearrowleft) “tortuous”: this question either requires an advanced concept, a proof that is long or inventive, or it is still open.

Use them as an inspiration: can you answer any of them? does it bring you to another problem that you can answer or study further? It is recommended to work on this question only AFTER you are done with the rest!

PART A — MANIPULATING THE CONCEPTS

Exercise 1: Continuous epidemics in the $S \leftrightarrow I$ case(5.5pt)

Motivation In this exercise, you will show that the final outcome in a epidemics with recovery to susceptible state depends on the parameter.

We consider a continuous (or fluid flow) model of an epidemic in a population, in the special case where nodes infect susceptible nodes with rate β and recover (*i.e.*, come back to the susceptible state) with rate γ . We denote by $y(t)$ the fraction of the population of nodes that are infected. Note that the fraction $x(t)$ of susceptible nodes satisfies $x(t) = 1 - y(t)$. According to our fluid flow assumption, both are real valued variables which increase and decrease continuously with time t .

Starting at time $t = 0$ with a fraction y_0 , the evolution of this epidemic may be modeled as the solution of the following differential equation:

$$\frac{dy}{dt} = \beta x \cdot y - \gamma y = \beta(1 - y) \cdot y - \gamma y .$$

- 1 (\curvearrowright) Assume that $\beta \neq \gamma$, find the solution of this differential equation for the initial condition $y(t) = y_0$ at time $t = 0$.
- 2 (\curvearrowright) What is the limit behavior of the epidemics after a long time (*i.e.*, characterized by $\lim_{t \rightarrow \infty} y(t)$)? What is the role of y_0 ?
- 3 (\curvearrowright) Assume now that $\beta = \gamma$, is the limit different from one of the two cases above? Can you compare how fast this limit is approached?

PART B — EXPERIENCING THE CONCEPTS

Coding Assignment Submission All programming should be written in one file. The name of this file should be `{UNI}_homework4.py`, with `{UNI}` replaced with your UNI. For example, if my UNI is `cjr2149`, I would name my assignment `cjr2149_homework4.py`. This file should be uploaded into your Drop Box on Courseworks before the deadline. Please name the file correctly, paying attention to the extension, and do not compress your file before uploading. Points may be subtracted if you do not follow these procedures.

Exercise 1: Coding Problem: Infections (19 + 0.5pt) The CDC believes that an outbreak of the ONOES virus is imminent. They've called on the students of COMS W4995-1 to help stop an epidemic. In this coding assignment, we will simulate an outbreak of the ONOES virus and try to develop effective vaccination strategies. The ONOES virus spread in the following way:

- The population is represented by a graph.
- Some nodes start infected in the graph.
- At each timestep, a node has a chance of infecting a neighbor with some probability p .
- Nodes with multiple infected neighbors have multiple chances to be infected.
- Some nodes can be vaccinated.
- Vaccinated nodes can never be infected and cannot pass on infection.

All code will be evaluated on the graph “infection_graph.txt”, available on the course web page. This graph is in adjacency form[`at`], so please read it using the `read_adjlist()` function.

1. (\curvearrowright) Create a function `next_infected = infection_timestep(G, p, infected)`. Let G be an undirected NetworkX graph, p be the infection chance, and `infected` be a list of nodes that are infected. The function should iterate through all neighbors of the infected nodes and infect them with probability p . If a node has n infected neighbors, you should try to infect that node n times. The function should then return all infected nodes.
2. (\curvearrowright) Create a function `next_infected = infection_timestep_vaccine(G, p, infected, vaccinated)`. Let G be an undirected NetworkX graph, p be the infection chance, `infected` be a list of nodes that are infected, and `vaccinated` be a list of vaccinated nodes. The function should create a new list of infected nodes and return it. However, for this function, the vaccinated nodes can never get sick and cannot pass on infection.

Now we will investigate the effects of different vaccination strategies.

3. (\rightarrow) Create a function `number_infected = SimulateNoVaccine(G)`.

This function should pick 10 nodes randomly from the graph, and set these to be the initial infected group. It should run `infection_timestep` for 50 timesteps. Set $p = 0.01$.

The function should return the number of infected nodes at the end of the simulation.

4. (\rightarrow) Create a function `number_infected = SimulateSimpleVaccine(G)`.

This function should pick 10 nodes randomly from the graph, and set these to be the initial infected group. Pick 7000 nodes randomly to be your vaccinated set. It should run `infection_timestep_vaccine` for 50 timesteps. Set $p = 0.01$.

The function should return the number of infected nodes at the end of the simulation.

5. (\rightarrow) Create a function `number_infected = SimulateFriendVaccine(G)`.

Members of the CDC, after taking COMS 4995-1, have decided to start a “bring your friend to the flu shot” program. Instead of picking 7000 nodes randomly, pick 3500 nodes. Have each of these 3500 randomly select a friend. Vaccinate the 3500 nodes and their randomly selected friend. Your seed set should be unique, but the “friends” do not have to be unique. That is, multiple people may decide to bring the same friend to get a vaccine, in which case, less than 7000 total people will be vaccinated.

Now, pick 10 nodes randomly from the graph and set these to be the initial infected group.

Run `infection_timestep_vaccine` for 50 timesteps. Set $p = 0.01$.

The function should return the number of infected nodes at the end of the simulation.

6. (\leftrightarrow) Create a function `number_infected = SimulateMyVaccine1(G)`.

Now it's your chance to develop a vaccination strategy! Pick any set of 1000 nodes to be your vaccinated group. You may assume you have full control over the graph; you may use whatever global properties of the graph that you like. Like the previous questions, run the infection model for 50 timesteps with $p = 0.01$.

The function should return the number of infected nodes at the end of the simulation.

Please include a write-up of your strategy for this function with the written portion of the assignment.

This question is on a half point, and you will receive it if you are among the 2 best vaccination strategies proposed. Otherwise, you'll have fun with competition.

7. (\leftrightarrow) Create a function `number_infected = SimulateMyVaccine2(G)`.

Now it's your chance to develop a vaccination strategy! Pick any set of 1000 nodes to be your vaccinated group. Use a different strategy than the one you used for `SimulateMyVaccine1`. You may assume you have full control over the graph; you may use whatever global properties of the graph that you like. Like the previous questions, run the infection model for 50 timesteps with $p = 0.01$.

The function should return the number of infected nodes at the end of the simulation.

Please include a write-up of your strategy for this function with the written portion of the assignment.

The student with the best strategy will receive extra credit!

PART C — CONCEPTS AT LARGE

Exercise 2: Blog post (15 pt) For the part C of this assignment, you will have to write your own blog post.

The blog will be readable by anyone, and will receive post from our lab and you. We will set up account for each of you (using the email address you indicated in piazza) to be part of the tumblr blog that is already existing.

Format: The blog should contain a link to an existing resource (probably online, but any reference that is accessible in anyway can be accepted, we can potentially make document available ourself), and at least two paragraphs (you are welcome to write more, just keep in mind that web's readers prefer quality over quantity).

You should in a first paragraph describe the reference to give a good account of its content (whether it's an article in the general press, a scientific , and in a second paragraph you should draw some connection with elements of the class or open questions. The idea is for you to develop a perspective as a commentator, ideally so someone would read the reference differently after your comments.

You should at anytime keep a positive tone (keep in mind that the people you write about like the authors of an article may read you as well). You are welcome to draw criticism of course, they should simply remain constructive and also recognize that limitations of a work are generally easy to find once the work is made.

When is it due? The blog will be due the week before and during Thanksgiving break, so you have any chance to reflect and perhaps discuss it with your friends or family during this time. So the due date will be on Monday 26th at 4pm. Keep in mind that you ought to find a topic you care about, this is probably the best way to maximize your grade and have fun.

Prior to your post, you will have to indicate to us on Monday 19th the resource you have chosen to discuss about.

How will it be graded? We will take into account: (1) the relevance of the topics and resource to the topic of the class (how social networks shape science and our society), (2) its originality, (3) the depth of the analysis you provide, (4) the precision of the connection you made with scientific elements (which could be things seen in class, but could also be something else, as long as you are rigourous), (5) the quality of the language and presentation. Note that these are roughly sorted from the most important criterion to the last, which means we don't expect you to write Shakespeare quality. As long as you are clear and say something interesting, we won't be focusing on the quality of the text.

Why don't I simply wait for someone to post and copy it? First of all, plagiarism will be taken very seriously. It's also probably a bad idea to wait for the first student to post and copy (although you are of course encouraged to read others' post and potentially get some inspiration), because like all of us, he or she will make it in a rush before it's due. Second, we will look at post in chronological order and if you happen to choose the same reference or the exact same topic, you will have to account for this post as being known before yours (and should provide an alternative view or a counterpoint). That would not bring you originality point, and it would overall be pretty boring, and not much easier.

Note also that you need to let us know the choice of the resource you want to present a week in advance. This way, if by any chance two people want to do the same thing, we can check that the conflict is properly handled (probably we'll ask you to send it to us separately by email and we'll post it).

I am lost, it's too complicated, what exactly do you want? I have an idea but I am not sure if it fits. Please tell us so (either me directly or any TA). We are here to help you. We will be especially available next week during office hours (since you have less maths, you can use them for that), and in class, and of course on piazza or email. We won't share your ideas with others.

On the other hand, any last minute request (say on Sunday 18th to find a resource) will be ignored. We are here to help you, but you need to ask us before the week end.

OK, I got it, I want to get some places I can look at for inspiration. You are really welcome to be creative. And there is a flurry of article in the NYtime, Fast company, The economist, etc. that deals with social media everyday. The difficulty is to find some with actual interesting content, and it's important that it's not just superficial to have something interesting to discuss. Here are some resources we recommend:

- The blog of our class: Right now the archive contains about a year and a half of articles collected on the various aspects of social media in relation with societal issues:
<http://social-network-and-computing.tumblr.com>
Browse archive at: <http://social-network-and-computing.tumblr.com/archive/>
- The blog from the class *Networks* organized in Cornell
<http://blogs.cornell.edu/info2040/2012/>
- Websites of important conferences: ACM WWW, ACM WSDM, Workshop WIDS, WIN.
- I have a list of people who contacted me (engineers in start-up, artists using social media) who kindly accepted to be volunteer for "short interview". This could be done on the phone, but I would encourage to go there and meet with them if you can. This is optional. You are welcome to contact me if you are interested to conduct one of those and write a short paragraph describing their work and products. We would of course take into account that you make extra effort to get original content in your post.