

SI Epidemics Model

Jordan Hasler

September 30, 2013

In this project, I looked at modelling the spread of infectious disease on a network. I stochastically simulated how long it would take for all individuals to become infected on a network using a SI model of disease. The SI model is the susceptible infected model. Before stochastically simulating the results, I obtained a system of differential equations and derived graph theoretical properties.

The REGS project was completed under the supervision of Professor Lee DeVille. The authors acknowledge support from National Science Foundation grant DMS 08-38434 “EMSW21-MCTP: Research Experience for Graduate Students”.

Epidemic Theory Origins The study of epidemics [1] is important to learn about how disease spreads and how to control the spread of disease. Networks are of fundamental importance in studying the spread of disease. One can study the spread of disease deterministically and stochastically. The original deterministic model goes back to Kermack and McKendrick. In this model there were three conditions S, I, R which stand for the number of susceptible individuals, the number of infected individuals, and the number of recovered individuals.

If there is a fixed population, $N = S + I + R$, they modeled the dynamics based on the following three equations,

$$\begin{aligned}\frac{dS}{dt} &= -\beta SI \\ \frac{dI}{dt} &= \beta SI - \gamma I \\ \frac{dR}{dt} &= \gamma I\end{aligned}$$

where β is the contact rate at which susceptible's become infected and γ is the recovery rate.

One may also look at the scenario when there is no recovery. In this case, the system is represented as

$$\begin{aligned}\frac{dS}{dt} &= -\beta SI \\ \frac{dI}{dt} &= \beta SI\end{aligned}$$

To see how these equations may be derived, assume that the number of susceptible particles is represented by $[S]$ and that the susceptible S will become infected I and stay infected with no chance of recovery. Further make the assumption that S become I at rate ρ .

Theorem: If we have $[S]$ susceptibles and $[I]$ infecteds with a rate of infection ρ , then we have the following equations

$$\frac{dS}{dt} = -\rho[S I]$$

$$\frac{dI}{dt} = \rho[SI]$$

where $[SI]$ is the total number of susceptible infected links.

Proof: We have an event that susceptible becomes infected which is represented by a decrease in susceptible and an increase in infecteds.

$$[S] \rightarrow [S] - 1$$

$$[I] \rightarrow [I] + 1$$

During a small time interval δt , the probability of infection is

$$\rho[SI]\delta t$$

We now compute the expected value for the susceptibles at $t + \delta t$:

$$\begin{aligned}\mathbb{E}[S_{t+\delta t}] &= \mathbb{E}[S_t] - 1\Pr(\text{infection}) \\ &= \mathbb{E}[S_t] - \rho[SI]\delta t\end{aligned}$$

We then have that

$$\frac{[S_{t+\delta t}] - [S_t]}{\delta t} = -\rho[SI]$$

Taking the limit as $\delta t \rightarrow 0$, we have that

$$\frac{d}{dt}[S_t] = -\rho[S_t I_t]$$

We thus find that

$$\begin{aligned}\frac{d[S]}{dt} &= -\rho[SI] \\ \frac{d[I]}{dt} &= \rho[SI]\end{aligned}$$

This is the deterministic method of studying an epidemic model. One may also study a system stochastically by using random graph models. In particular one may use the Gillespie algorithm [3] to compute the expected time for all individuals in a population to become infected. We assume that we have a fixed population size N . Let the probability of infection $\gamma = 1$. We assume that each individual (a vertex on the graph) becomes infected at a rate kl where k is the number of infected neighbor. For a vertex, we can write this rate as

$$\rho(AI)_i = \rho \sum_{j \neq i} A_{ij} I_j$$

This counts the number of infected individuals. If we compute AI , and the j th entry gives us the number of infected neighbors of individual j . To simulate stochastically the time for an infection to occur, we let T_j be $E(q_j)$ we have T_j be an exponential process of rate $q_j = (AI)_j$. The time until the next infection can be represented by $\delta t_m = \frac{-1}{m} \log(\text{Rand})$. Compute $\min_j T_j$ which is an exponential of rate $\sum_{j \neq i} q_j$ [2]. This minimum occurs at one of the vertices. The time is updated as $T = T + \delta t_m$. The time is updated until all vertices are infected.

Expected Infection

In studying epidemics on networks, one question that we are interested in is how much time will it take for everyone to become infected. We can study this by computing the expected time until everyone is infected. Define a jump chain $S_k = T_k - T_{k-1}$ where T_k is the time up until the k th event. If we model the infection as kI_t where k is a natural number and I_t is the number infected at time t , then we can define expected value for each step in the jump chain S_k . [2]

For each S_k , we have an exponential process with rate q . $\mathbb{E}(kq)$ where k is the number of infections and q where is the exponential process rate. We have that $P(S_k > t) = e^{-kqt}$. [2]

$\mathbb{E}[S_k] = \mathbb{E}[E(qk)] = \frac{1}{kq}$. The expected time for all infections is then the sum over all those infected:

$$\mathbb{E}[S_k] = \sum_i \frac{1}{qi}$$

This gives us a measure of the expected time for all people to become infected. We can furthermore calculate the variance

$$\mathbb{E}[S_k^2] = \frac{2}{k^2q^2} - \frac{1}{k^2q^2}$$

$$V = \mathbb{E}[x^2] - \mathbb{E}[x]^2$$

Giving us that the variance is

$$V = \sum_{i_o} \frac{1}{k^2q^2}$$

Graph Properties

Since the study of epidemics is very much network oriented, graph properties are useful. [1]

Let us consider a network with nodes of type t . In the case of an SI model, the nodes can be of two types susceptible and infected.

Let $\phi(i, t) = 1$ if node i is type t . Otherwise $\phi(i, t) = 0$

Since $\phi(i, t) = 0$ for all fixed except when node i is of type t , we have

$\sum_t \phi(i, t) = 1$ for each i

Let G be the matrix $G_{ij} = 1$ if i, j are connected. Otherwise, $G_{ij} = 0$

The number of singles is defined as $\sum_i \phi(i, t)$.

The number of doubles $[t_1, t_2]$ is $\sum_{i,j} \phi(i, t_1)\phi(j, t_2)G_{ij}$.

The number of triples $[t_1, t_2, t_3]$ is $\sum_{i,j,k} \phi(i, t_1)\phi(j, t_2)\phi(k, t_3)G_{ij}G_{jk}$.

$$\begin{aligned} & \sum_{t_1} \phi(i, t) \\ &= \sum_{t_1} \sum_i \phi(i, t) \\ &= \sum_i \sum_{t_1} \phi(i, t) = N \end{aligned}$$

This says that the number of infected is equal to the sum of the types. In our case, this is the infected and susceptible is equal to the total.

One can also obtain similar formulas for doubles. Namely, we have the formula:

$$= \sum_{t_2} [t_1, t_2] = n[t_1]$$

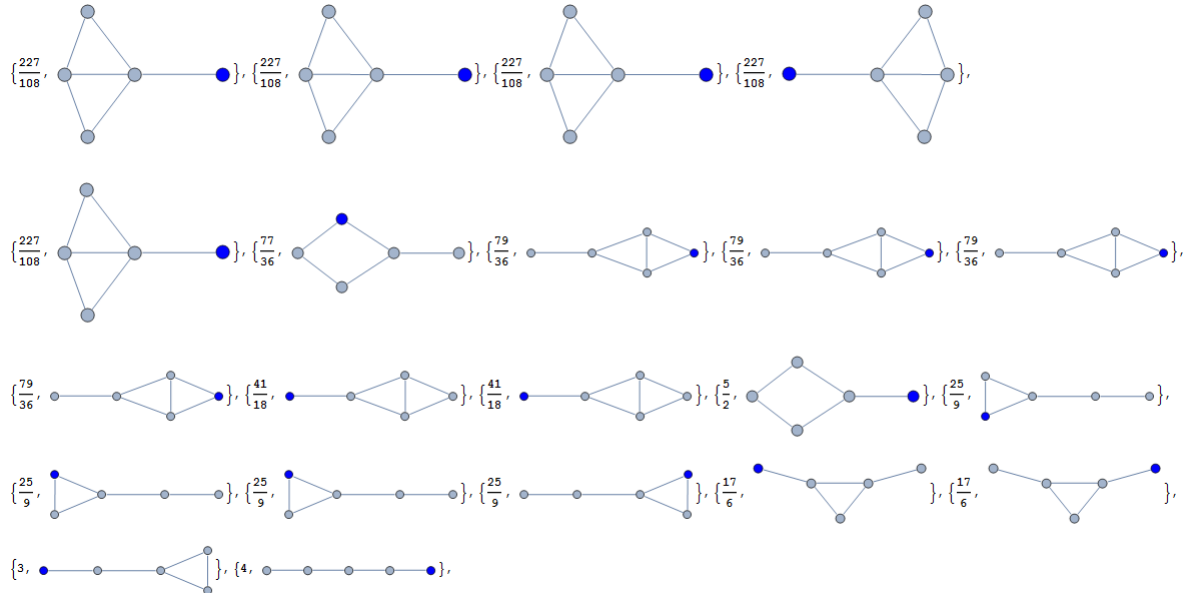
This says that for a homogenous network with on average each node having n neighbors that the sum of the edges starting from a node of type t_1 to a node of a different type is $n[t_1]$.

In our case this means that $[SS] + [SI] = n[SS]$. If each node had on average 2 neighbors, we would have $[SS] + [SI] = 2[SS]$.

I made a computer program to compute the expected time until all individuals are infected. This revealed that the connectivity of the graph is important for describing whether all individuals are infected. Quarantines are important to prevent the spread of an infection. Also revealed was that the same graph may have a different time for all to be infected. This was a result of a different placement of the nodes with respect to each other. Two isomorphic graphs can have different times depending upon whether they start at the same node or not. Two isomorphic graphs which have the same starting node will take the same time to infect since each node will be infected at an exponential rate and the number of neighbors of each vertex will be the same. The following data was obtained for a graph with 5 vertices with the dark blue color indicating the starting node. Each graph has starting node highlighted and the time for all individuals to become infected is included. Notice that the density of the graph determines the time it takes for all to become infected. The more dense the network, the faster the infection spreads.

Future Work

I will study more of the properties that are directly related to the network. For instance, I will study the time it takes for adding multiple starting nodes. Afterwards, I will look at more advanced models such as the SIR model. The comparison of the stochastic model with the dynamical system will also be studied.



References

- [1] Matt Keeling, & Eames. . 2005. *Networks and epidemic models*. J.R. Soc. Interface. 2, 295-307.
- [2] Norris, J.R. *Markov Chains*. CUP. 1997.
- [3] Matt Keeling & Rohani. Modeling Infectious Diseases in Humans and Animals. Princeton University Press. 2008.