CSE 8803 EPI: Data Science for Epidemiology, Fall 2024

Lecturer: B. Aditya Prakash — Sept. 3rd and 5th, 2024
Scribe: Ben Zabriskie, Benjamin Gyimah-Asamoah — Lecture 5 : Other Contagion Models

---

1 Summary : Contagions Everywhere!

This lecture covers extensions to classic SIR epidemic models, such as SIS, SIRS, and SEIR. These, as well as a range of other models, have been developed to capture additional layers of complexity which traditional SIR is unable to or makes simplifying assumptions that lead to unsuitable modeling predictions. Examples of these other models include the Independent Cascade model, SpikeM models, and the classic Lotka-Volterra (Predator-Prey) model. Scenarios in which each of these models may be most suited for are briefly discussed.

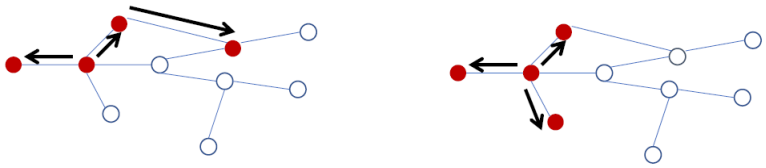


Figure 1: Possible paths of the spread of a contagion in a network

2 Continuation from previous lecture: SIR models and Extensions to SIR

SIR models are simple and intuitive in nature, but often make simplifying assumptions about the model characteristics that do not capture the range of potential factors influencing real world behaviors. To better capture these elements while retaining the intuitiveness of SIR, a substantial number of extended models based on SIR have been developed. Examples of these include SIS (Susceptible-Infected-Susceptible), SEIR (Susceptible-Exposed-Infected-Recovered), and SIRS (Susceptible-Infected-Recovered-Susceptible). The primary differences between these variations are how the compartments are initialized, how the transitions between states occur, and the additional layers of complexity used as input to compute the reproductive number R_0. For example, in the SIRS and SIS models people who have been infected can get better and be infected again. In the SEIR model, a new compartment E is introduced to capture an exposure or incubation period prior to being recognized as infected – potential or probable infections do not immediately transition from the S to I compartments.

2.1 SIR & SIS network models: Discrete Steps

Before discussing further details on these more complex models, it is important to provide more details on how the basic SIR and SIS network models work. The same basic logic can be extended to the more complex models that will be discussed later.

Starting with the SIR network model, we can represent the states as in Figure 2. As for the ODE, there are three possible states: Susceptible (S), Infected (I) or Recovered (R). An

important charateristic of the SIR model is that once an individual becomes infected and recovers, they cannot become infected again.

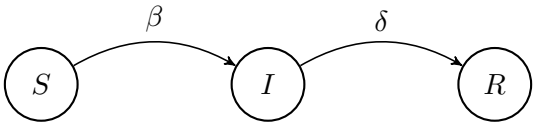
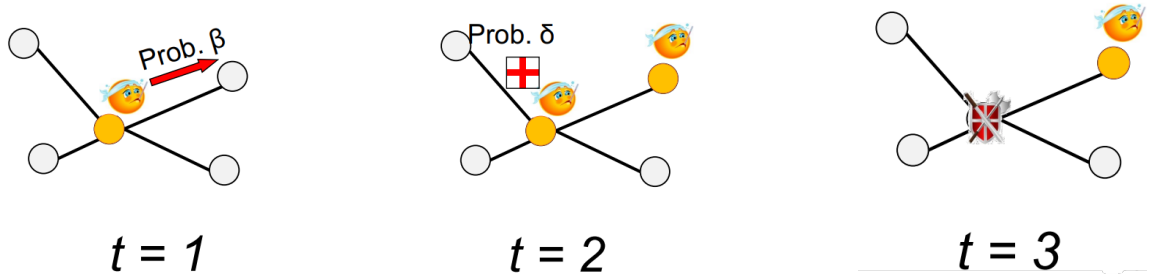


Figure 2: Representation of SIR Network States

We can represent a basic SIR network in Figure 3 from the perspective of an Infected individual in the central node, surrounded by Susceptible individuals. At random, this Infected individual comes into contact with one of its neighbors. This randomly selected Susceptible neighbor is then infected with probability β. The original Infected individual then becomes Recovered with probability δ. This process is shown in Figure 3 across several timesteps. When implementing these network models, transitions between states are often controlled by comparing a random number draw to the transition probability between the two states being considered. This makes the network model stochastic in nature.



(a) (b) (c)

Figure 3: SIR Network Model. Images from Lecture 4, Slide 49

We can also represent diseases where individuals do not become permanently recovered and can become susceptible again. In this case, an SIS model can be used to model "flu-like" diseases. The steps of an SIS network model are very similar as the ones previously described for an SIR model. If we consider the perspective of an infected node, a random neighbor will be selected and then infected with some probability β. The previously infected individual will then transition back to susceptible with another probability α. Because flu-like diseases have no immunity, there is no Recovered state in this case.

2.2 Calibrating SIR network models

Obtaining a suitably appropriate network is the primary importance for network-based models, since incorrect networks will always produce incorrect SIR curves. In the case of differential equation-based (ODE) models, only the hyperparameters β and γ need to be inferred, whereas networks require both β and γ hyperparameters plus the correct construction of the network itself.

There are many considerations when calibrating a model. A key one is which data to use to do the calibration and how reliable that data is. Often, there are many data problems to navigate. These include missing data, biases, reporting lags, etc. For some diseases, COVID-19 being a key example, data on infected individuals is likely to be flawed. Data on

mortality and hospitalizations are likely to be more accurate and thus more suitable to use for calibration. Therefore, we might calibrate the hyperparameters of an SIR ODE model using death data ($R_{observed}$):

$$\{\beta^*, \alpha^*\} = \text{argmin}(R(t) - R_{\text{observed}}(t))^2$$

Based on domain knowledge and expertise, bounds can be placed on these parameters to ensure the optimized values are realistic for the biology.

One example of network construction might be the boroughs of London where the geographical proximity of neighboring boroughs is used to model an SIR network. ODE-based SIR/SIS models make the simplifying assumption of constant homogeneous population mixing, which can be approximated in network models as "cliques". Cliques are complete graphs in which all nodes share an edge with all other nodes in the graph. More complex model heterogeneity can be achieved as the connectivity, betweenness, and centrality of the network changes.

For instance, the London example does not approximate the network as a clique. Instead this network is created by representing each borough as an individual homogeneous population mixing within itself but the heterogeneity of this network comes in by treating each borough node independently. Each node, say node A, is only connected to another node, say node B, if the boroughs representing A and B are neighbouring boroughs. This takes into account a slightly more realistic mixing assumption which would make the model more complex and most likely more accurate.

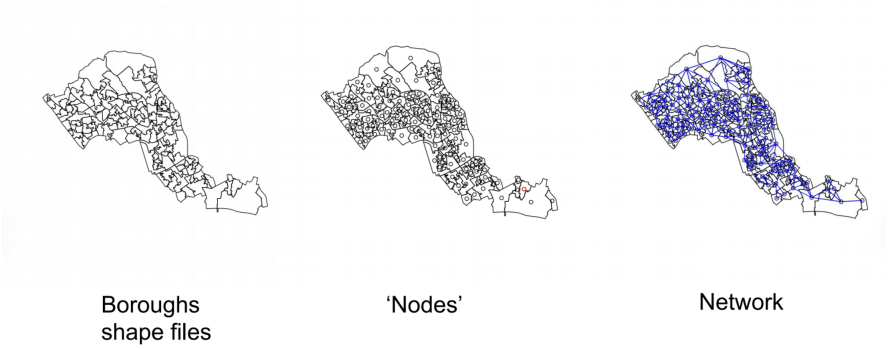


Figure 4: Network diagram for the London borough example. Image from Slide 57 of Lecture 4

Once we have collected reliable time-series data containing infections, recoveries and other relevant parameters we can use statistical methods to estimate β and γ and calibrate our model. Common methods to achieve this include Maximum Likelihood Estimate (MLE), Bayesian Inference, and Least Squares Fitting. MLE finds the parameter values that make the observed data most probable by maximizing a likelihood function that expresses the probability of observing the data given the model parameters.

Bayesian methods incorporate prior knowledge about parameters and update this knowledge based on observed data. A common Bayesian technique is Markov Chain Monte Carlo (MCMC). This technique involves exploring the parameter space to find a distribution of

plausible parameter values. This method is also capable of quantifying the uncertainty in parameter estimates.

Lastly, Least Squares Fitting involves minimizing the sum of squared differences between observed data points and the model's predictions. This is useful for fitting epidemic curves to data, and finding the parameter values that produce the closest match between the model's output and the actual observed case numbers over time.

As for optimization, there are many methods that can be used to find the best parameter values. This includes but is not limited to gradient descent methods for finding optimal parameter values, genetic algorithms for exploring a wide parameter space, and particle swarm optimization for complex, multi-dimensional problems. For network-based models, there are some additional steps needed for optimization. The network must be representative of the community dynamics so adjusting things like the clustering coefficient or node degree can help achieve this.

Some common optimizers are summarized below, along with a short description. [1]
Non-linear optimizers:

- Nelder-Mead: Direct-search optimizer, updates set of test points until it converges
- Levenberg Marquardt: non-linear least squares, similar to gradient descent
- Powell's Method: direct-search optimizer, performs bi-directional search in each direction until it converges
- Broyden-Fletcher-Goldfrarb-Shanno algorithm: gradient-descent, preconditions gradient with curvature information

Bayesian Optimizers:

- Markov chain Monte Carlo: approximates parameter distribution using a Markov chain's stationary distribution
- Maximum Likelihood: Iterated optimizer that adds a perturbation to the parameters to add variability
- Approx. Bayesian computation: Uses simulations to approximate likelihood function
- Probe matching: Aims to minimize difference between simulated and actual data

3 Other Types of Epidemic Models

3.1 Metapopulation Models

All models have trade-offs as the complexity (i.e. the number of model paramters) increases, particularly in relation to the model's transparency and accuracy. Epidemiological models like SIR et al. have found many uses in other realms such as social collaboration initiatives, information diffusion, viral marketing strategies, cybersecurity, gaming, and in the life sciences such as public health, personalized medicine, and ecology.

[1] Information on optimizers paraphrased from Lecture 4, Slides 14-15.

3.2 Independent Cascade (IC) Models

A special case of the network-based SIR model that is particularly useful for modeling "viral" spread (in the context of social media, not *sensu stricto* public health). Independent Cascade models a social network with directed, weighted edges which reflect, for example, individual pairwise probabilities of 'infection' spread from node u to node v. This model is reminiscent of the Twitter "following" system where connections are directed, i.e. Person A may "follow" Person B while Person B may not follow Person A. The IC model aims to capture the "spread of infection", here referring to the amount of interactions, precipitated when a node becomes active and fires an event. The active node tries to "infect" each of its out-neighbors, thus cascading the infection forward in time. The odds that this peer-to-peer activation is successful is based upon the probability of the edge weight from (u, v). As each event is a single discrete action, the activated node tries to activate its nodes only once before reverting to an inactive state. In relation to ODE-based SIR models, network-based IC models extend the β rate of transmission from a single universal rate to a pairwise rate along the edge (u, v). Similarly, the recovery hyperparameter δ is understood to be 1.0.

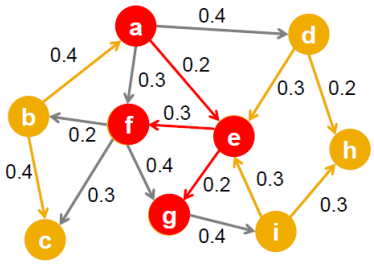


Figure 5: Example IC model diagram with the individual edge weights. Image from Slide 4 of Lecture 5

One key benefit of the IC model is that it is very intuitive for information flow, however this also brings with it some drawbacks: the computational complexity increases exponentially as the number of edges in model grows, and the significant number of parameters (e.g. pairwise β_{ij} , to estimate. These can make it difficult to calibrate IC models accurately.
One feature of the Independent Cascade (IC) model is its inherent tendency for popularity or the rate of spread to decline exponentially over time. This characteristic is deeply rooted in its foundation—the SIR model—which also demonstrates a similar exponential decay in terms of the number of susceptible, infected, and removed individuals over time. However, this model may not be universally applicable for capturing every real-world scenario. For instance, when examining the sharing behavior of blog posts or the virality of online trends, the decline in popularity frequently deviates from an exponential curve. Real-life datasets often indicate that this drop-off more closely follows a power-law distribution, specifically with an exponent of -1.6. This nuance is important because a power-law decay accommodates the possibility of rare but significantly impactful events, often referred to as "long-tail" events. These rare occurrences can cause unexpected rebounds in popularity or spreading rates—phenomena not accounted for when relying solely on the exponential decay assumptions of the SIR and IC models.
Let's look at a real world scenario where this model would be beneficial. Consider a trending hashtag, #SaveThePlanet, initiated by a few influential environmentalists on a social media

platform such as Twitter (or X). In this IC model, each person who tweets this hashtag has a one-time chance to influence each of their followers to retweet it. If the follower decides to retweet, they, in turn, get a one-time chance to influence their own followers, and so on. However, the IC model would predict that the hashtag's popularity would peak quickly and then drop off exponentially as people either choose to retweet or not, effectively moving from the "susceptible" to the "infected" and then the "removed" categories in SIR model terms.
While this is a simplified example, it encapsulates the basic dynamics of the Independent Cascade model, demonstrating how a single event can trigger a cascade of interactions.

3.3 SpikeM Models

A SpikeM Model is a variation of the classic SIR model that uses the Power Law to calculate β, making it not exponential in the same way as standard SIR. This model works by having a series of 'uninformed' nodes, followed by a random "disturbance" such as -for example- Operation London Bridge, the surprise news report covering the death of Queen Elizabeth II. Following the perturbation event, "infection" can spread in the form of word-of-mouth transmission. This model allows for a resurgent or periodic spread after the initial disturbance event and initial spread of infection subsides. This is because Power Law distributions that have mathematical properties making the subsequent "unlikely" disturbance events in a standard SIR model more likely to occur.

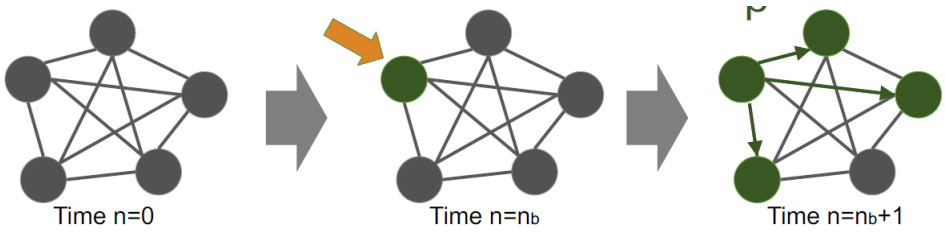


Figure 6: Example SpikeM model diagram. Image from Slide 12 of Lecture 5

3.4 Complex contagions

The idea behind a complex contagion models is that the neighbors of a node expose that node to a transmissible 'contagion'. In a traditional SIR model, the probability of a node becoming infected is $p = 1-(1-\beta)^k$, where k is the number of infected adjacent neighbors. This results in a curve that resembles a Logarithmic or Negative Binomial growth curve and reflects the probability of becoming infected based on the number of infectious neighbors. In effect, this suggests an effect of diminishing returns from the amount of infected neighbors (i.e. the Power Law decay parameter α is less than 0).

Complex contagions may model different behaviors. For example, a sharp rise in the probability of becoming infected once a certain minimum threshold of infectious neighbors is reached, or conversely, a steep drop representing an over-exposure effect such as an individual being "over-advertised" to, which reduces the likelihood of being interested in a product. This type of model may not be suitable for application to many epidemiological scenarios but may be useful for related fields.

3.5 Other types of complex (single contagion) models

Other types of structural models exist which take a different approach to modeling single contagion infection spread based not on individual nodes, but on different levels of granularity, such as a person's social circle, local community, or familial relationships.

Briefly, examples of these models include Decision models which have nodes that each have an activation threshold before the node will activate. An example of this might be at the start of a riot, each person in a crowd may have an activation threshold (e.g. number of disgruntled neighbors beginning to riot) before they join in. Thus, the centrality of nodes in this type of network is a critical factor in infection spread.

Other examples of complex contagion models are voter models where nodes pick an 'ideology' based on their neighbors and it shows how beliefs are picked, or hybrid models where not every node is modeling the same thing. For example one model could show both how things are pick and also how they spread all at once.

3.6 Multiple contagion models

These are extensibly models like the Lotka-Volterra Model, sometimes known as the Predator-Prey model. This model consists of two populations: one of a predator and the other being a prey species. Each of these populations experience fluctiations in population size of time, influenced both by birth/death parameters of their own population (part of each species' carrying capacity) and by the population of the other species. The expected population sizes can be modeled using ODEs and are extensible to any number of competitive species.

The classic Lotka-Volterra model contains 2 species, X and Y, where X represents a total number of prey species and Y represents a total number of predator species that preys on X. An example here would be if X represented of deer and Y represented of lions hunting on that deer. The deer and lion populations(X and Y respectively) interact with one another using the following system of equations:

$$\frac{dx}{dt} = \alpha x - \beta xy \tag{1}$$

$$\frac{dx}{dt} = \delta xy - \gamma y \tag{2}$$

where αx and δxy represent the birth rated of Species X and Species Y respectively, and the βxy and γy terms represent the mortality rate of each population. There are two stable equilibria points: one is trivial – the extinction point, where $x = y = 0$. The second one is the co-occurrence point at $x = \frac{\gamma}{\delta}$ and $y = \frac{\alpha}{\beta}$.

3.6.1 Extensions for multiple contagions.

There can also be two competing contagions fighting for the same population in one model. This could be for real world diseases like influenza and SARS-CoV-2 or this could be used for non-epidemiological fields like modeling the spread of Apple vs Android devices. This requires the model to have two β and two γ parameters to model the traits of the two different contagions.

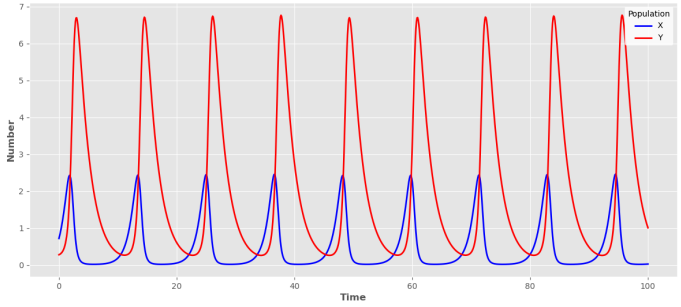


Figure 7: Example Lotka-Volterra model.

There can also be an ϵ (epsilon) value to model the interaction. This is because sometimes one contagion can block off people from getting another, like getting Cowpox preventing people from getting Smallpox. In such a case, $\epsilon = 0$. Sometimes they only grant partial immunity so getting one makes people less likely to get the other but it is still possible. This is like getting one variant of COVID making it less likely someone gets a different variant but it is still possible to get. This is represented by $0 <= \epsilon <= 1$. There could also be two diseases that are unrelated and do not limit people from getting the other if they already have one. This scenario occurs when $\epsilon = 1$.

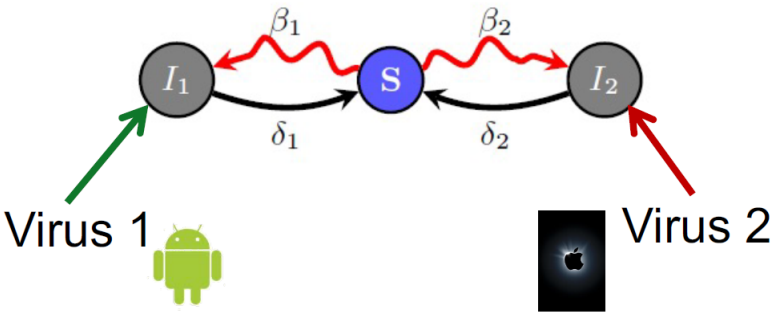


Figure 8: A two-contagion model diagram. Image from Slide 38 of Lecture 5

3.7 Branching Processes in Epidemic Modeling

In a branching process model, we consider the number of new infections generated by each infected individual. The idea of fixed points in these models is essential for understanding the long-term behaviour of an epidemic.

Branching processes are stochastic models that can be used to describe the spread of epidemics, especially in the early stages. Showing how each infected individual can give rise to a random number of new infections, creating a "branch" that can produce further "branches" over time. This continues through generations, where each generation consists of individuals infected by those in the previous one. The branching process helps in understanding the probability of an epidemic dying out after a certain number of generations.

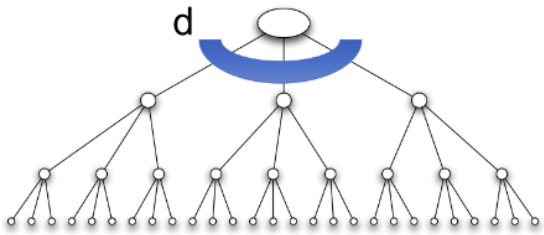


Figure 9: Example of Branching. Image from slide 42 of lecture 4

Fixed points refer to the steady-state solutions where the number of susceptible, infected, and recovered individuals remains constant over time. They are crucial for understanding the long-term behaviour of an epidemic. Using the SIR model as an example fixed points are found by setting the derivatives of the model's equations to zero, indicating no change in the number of individuals in each compartment. A common fixed point is when there are no infected individuals left, signifying the end of the epidemic. The stability of these fixed points determines whether the system will return to the fixed point after a small perturbation or move away from it.

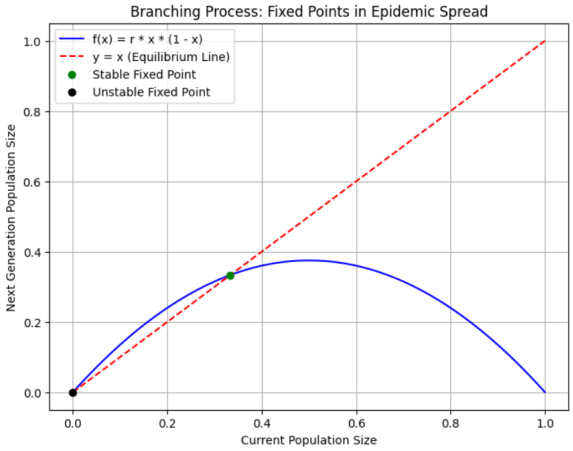


Figure 10: Example of Fixed Points in a Branching Process

Figure 10 depicts the effects of two different fixed points in a branching process model. The x-axis represents the current state or population size, y-axis represents the next state or population size after one generation, curve f(x) shows how the population changes from one generation to the next, and the fixed points occur where f(x) = x, i.e., where the curve intersects the y = x line.

The "Stable Fixed Point" (green) represents when there is a stable equilibrium, if the infection rate is below this point, the epidemic will naturally die out, leading to the Disease-Free Equilibrium. The "Unstable Fixed Point" (black) represents when there is an unstable equilibrium, if the infection rate exceeds this point, the epidemic will grow, leading to the Endemic Equilibrium.

References

- Zarebski, A., & Tonkin-Hill, G. (2018, October 3). Stochastic sir network model on City of London Boroughs. epirecipes. http://epirecip.es/epicookbook/chapters/london-boroughs-network/r
- David Kempe, Jon Kleinberg, and Eva Tardos. Maximizing the spread of influence through a social network. SIGKDD 2003.
- Richardson, M., Kuhn, T., & Juels, A. (2002). RFID privacy: A technical primer and review. In IEEE Pervasive Computing (Vol. 2, No. 1, pp. 34-43)
- Leskovec, J., Adamic, L. A., & Huberman, B. A. (2007). The dynamics of viral marketing. In Proceedings of the 7th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 568-574)
- J. G. Oliveira A.-L. Barabási. "Human Dynamics: The Correspondence Patterns of Darwin and Einstein." Nature, 437, 1251.
- Matsubara, Y., Sakurai, Y., Prakash, B. A., Li, L., & Faloutsos, C. (2012). Rise and fall patterns of information diffusion: model and implications. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining (SIGKDD '12), pp. 6-14
- Nickbakhsh, S., Mair, C., Matthews, L., Reeve, R., Johnson, P. C., Thorburn, F., ... & Jansen, V. A. (2019). Virus–virus interactions impact the population dynamics of influenza and the common cold. Proceedings of the National Academy of Sciences, 116(52), 27142-27150.