CSE 8803 EPI: Data Science for Epidemiology, Fall 2023

Lecturer: B. Aditya Prakash
Scribe: Tejas Pradeep, Saahir Dhanani

September 13, 2022
Lecture 6: Dynamics of modeling

---

1 Overview of Dynamics of Models

In this lecture we broadly covered the Dynamics of Models: looking at how models behave long-term and what factors influence their long-term behavior. In particular, we are focusing on Thresholds and Stability Points with respect to single-virus models and competing contagion models.

Thresholds are boundary points where a model's long-term behavior goes from one phase too another phase (e.g. dying out vs epidemic). Stability Points are positions a model tends to return to or oscillate around over the long term regardless of perturbations. In addition to Threshold and Stable points there are also Unstable points. Perturbations to these points tend to lead to large changes in model behavior. In order to better understand this behavior this lecture introduces us to a couple of concepts to make it easier to analyze such system in the context of a more general framework.

The first concept is the generalized S*I*V* model which can be seen as a more general model that encompasses SIR, SIS, and SIRS models. The second concept introduced is the idea of Effective Strength. The third concept introduced is the idea of modelling epidemics as networks being modified by an Non-Linear Dynamic System (NLDS)

1.1 A Fundamental Question

How do we determine a condition where a virus will go extinct quickly regardless of the initial infection conditions? We can determine this condition mathematically by finding a threshold at which a given virus will either go extinct or become invasive.

Establishing this threshold for a given virus is important for multiple reason. For example, when a virus is below a threshold, we might not have to 'worry' so much and take less precautionary steps, such as social distancing and lock-downs, to reduce its spread. Or when it is above the threshold, we can consider stronger measures like vaccination efforts. It can also help enhance forecasting of 'what-if' scenarios.

An example of such threshold behavior can be seen in figure 1.

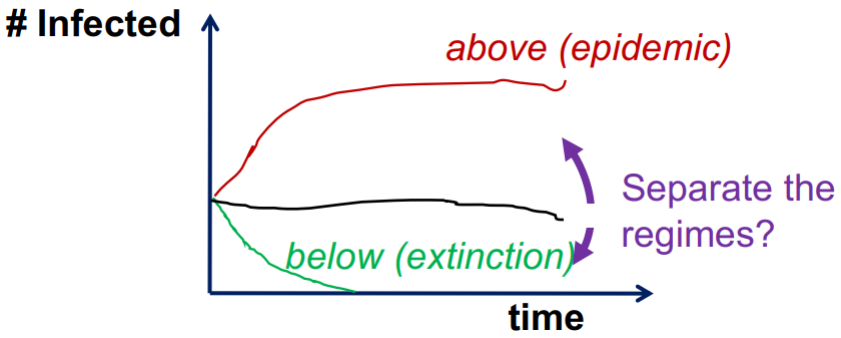


Figure 1: Black line indicated threshold tipping point where below the line a virus would go towards extinction and above the line would move towards and epidemic. [3]

In addition to threshold points, we can also look at stable, unstable, and neutral points. Stable points are characterized by the fact that small perturbations to the system will not lead to large changes in long term behavior. Unstable points are characterized by the fact that large perturbations to the system will lead to large changes in long term behavior. Finally neutral points are characterized as neither stable nor unstable. They can be thought of as a ball on a flat plane or a phase plot full of dots (zero magnitude trajectory arrows).

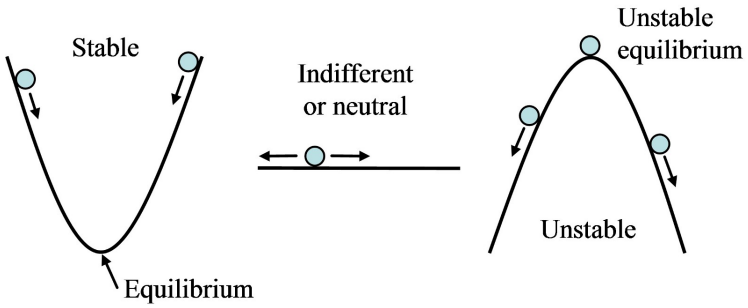


Figure 2: Stable, unstable, and neutral points [6]

2 Model Generalization

All of the SIR, SIS, SIRS,... models that we have looked at so far can be seen as specialized forms of a more general S*I*V* model where they can be an arbitrary number of "suscep-

tible" S states, one or more "infections" I states, and zero or more "vigilant" V states. An outline of this general state transition graph can be seen in figure 3.

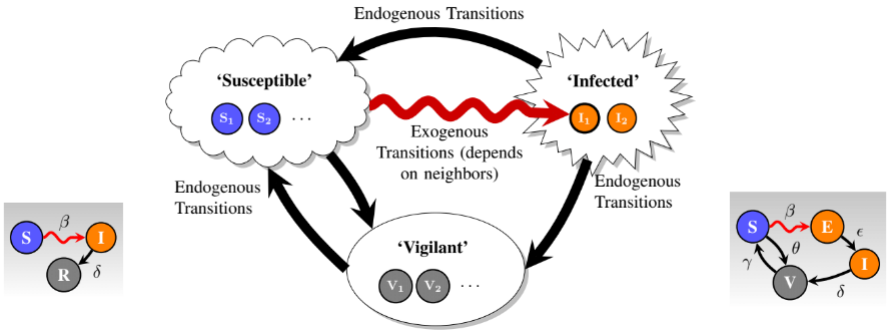


Figure 3: Diagram outlining the S*I*V* transition diagram. [5]

As seen in figure 3, "Susceptible" states can transition to "Infected" states via "Exogenous" transitions that depend on the states of neighbors, and "Infected" states can revert back to an "Susceptible" state via an "Endogenous" transition that can happen independently of neighboring states. "Infected" states can transition to "Vigilant" states via Endogenous transitions. Finally, "Vigilant" and "Susceptible" states can endogenously transition between each other.

To form the more specific models such as SIR and SIS, we just restrict the number of states and transition pathways accordingly. For example an SIR model can be viewed as an S*I*V* model with just one susceptible state, one infectious state, and one vigilant state (recovered); where the transition pathways between each state have been restricted to an exogenous infection pathway from the susceptible state and the infected state, and a recovery endogenous pathway between the infected state and the recovered state.

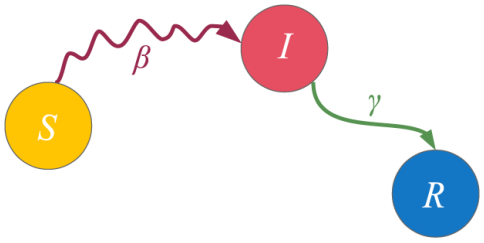


Figure 4: The state transition diagram for an SIR infection model.[7]

3 Thresholds for arbitrary static networks

A long standing question in epidemiology has been determination the epidemic threshold, which has been defined in the literature as, 'the minimum level of virulence to prevent a viral contagion from dying out quickly' [5]. Given that the epidemic threshold can be challenging to determine depending on a contagion's virulence. In this lecture, we were introduced to a generalized framework for analyzing thresholds for static graphs based off of a factor called Effective Strength (s), which can be thought as similar, if not the same, as the reproductive number (R0) which we have learned of in previous lectures. Thus, if s <

1 it is below the threshold for mass spreading event, if s $>$ 1 then it is above the threshold for an epidemic, and if s = 1 then it is at a tipping point [5].

The Effective Strength factor (s) consists of two components, λ (the largest eigenvalue of the network's adjacency matrix) and some constant C_{VPM} that is dependent on the virus propagation model being used.

$$s = \lambda * C_{VPM}$$

The constant C_{VPM} typically consists of constants describing the state transition probabilities between the various epidemic states. For example, if we consider the SIS model, then:

$$s = \lambda * \frac{\beta}{\gamma}$$

where γ is the recovery probability, β is the viral transmission probability, and λ is the largest eigenvalue determined in from the network's adjacency matrix (A).

Other effective strength formulations can be found in the table in figure 5

Table III

Models	Effective Strength (s)	Threshold (tipping point)
SIS, SIR, SIRS, SEIR	$s = \lambda_1 \cdot \left(\frac{\beta}{\delta}\right)$	$s = 1$
SIV, SEIV	$s = \lambda_1 \cdot \left(\frac{\beta\gamma}{\delta(\gamma+\theta)}\right)$	
$SI_1I_2V_1V_2$ ($\sim$ H.I.V.)	$s = \lambda_1 \cdot \left(\frac{\beta_1 v_2 + \beta_2 \epsilon}{v_2(\epsilon+v_1)}\right)$	

Figure 5: Table containing the effective strengths for various Virus Propagation Models. [5]

3.1 A deeper dive into λ

If we consider the fact that an adjacency matrix represents all connections between all nodes in a graph, then the largest λ value of an adjacency matrix represents the strength of connectivity [4] between all nodes in the graph. For example, figure 6 shows how λ is empirically correlated with higher levels of inter-node connectivity within the given graph.

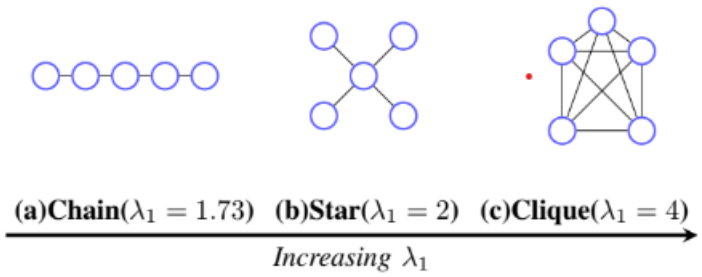


Figure 6: Diagram illustrating how higher λ values are correlated with higher inter-node connectivity in a graph. [5]

When we examine the graphs from Figure 6a and b, each graph has an equal number of nodes and edges, and there respective λ values are relatively close in value. However, Figure 6c has the same number of nodes but has more edges and thus is more connected. From graphs a and b, we determine that what influences the strength of λ is not the number of edges but the way nodes are connected. In the Chain graph, for any susceptible node to be infected at a future time point, an infected node can only infect its neighbor node(s) as their connection is linear. From the Star graph, if the central node is infected then there is a higher probability that in will infect at least one node it is connected with. However if the central node moves to the recovered state before infecting other nodes, the infection will die out and its connected nodes will not become infected. However, when we examine Figure 6c, we can see that the connection between every node increases the likelihood that all nodes will eventually become infected at some future time point. Finding the largest λ of each graphs adjacency matrix shows a strong correlations with the strength of connectivity within a graph. Overall, what λ tells us is that the chance of an epidemic increases when the inter-node connectivity is high, especially when λ is large. Using λ to make forecasting predictions of a virus with increase infectivity could have suggested better ways to take precautionary steps to reduce the transmission of the Omicron variant.

Below are some applications that benefit from finding the threshold of arbitrary static networks.

1. **Public Health Interventions**: Understanding λ can guide public health interventions. For instance, if λ is high, indicating strong inter-node connectivity, targeted vaccination campaigns could focus on "super-spreader" nodes to effectively reduce λ and thereby the potential for an outbreak.

2. **Network Optimization**: In computer networks or social networks, λ can be used to identify key nodes that, if removed or secured, could significantly reduce the risk of network failure or information spread.

3. **Financial Systems**: In financial networks, understanding λ could help in identifying the most interconnected banks or institutions. Targeted regulations could then be applied to these entities to prevent systemic risks.

4. **Emergency Preparedness**: λ can be used to model the spread of wildfires, floods, or other natural disasters across a network of affected regions. Resources could then be optimally allocated to the most connected nodes to mitigate damage.

Understanding λ as a threshold has real-world implication. For example, removing nodes from a given graph with greater connectivity will decrease the size of λ. In a real-world example, we could vaccinate specific individuals (nodes) that have the highest connectivity to help reduce the the size of λ so that the number of future infections drops beneath the effective strength (s) threshold. Additionally, λ has implications for 'what-if' scenarios (briefly mentioned in 1.1 above) as it can be used to forecast future epidemics if a virus becomes more infections/virulent. A very good and relatable real-world example of this is the current COVID epidemic. In October of 2021, the Delta variant was still the dominant strain causing infections in the USA. By mid-January 2022, the Omicron strain swept through the US generating the greatest number of infections within a short period of time at an alarmingly fast rate.

4 Non-Linear Dynamic Systems (NLDS)

Non-Linear Dynamic Systems (NLDS) are mathematical models that describe how the state of a complex system evolves over time in a non-linear manner. In the context of contagion spread through a network, NLDS can be employed to capture the intricate interactions and probabilistic states of each node in the network.

- **Graph**: Represents the network where each node is an entity that can be affected by the contagion. Edges between nodes indicate some form of interaction or connectivity that can facilitate the spread of the contagion.
- **Probability Vectors** $\vec{P}_t$: These are time-stepped vectors that store the probabilistic "state" of each node with respect to the contagion at a given time t. The size of each probability vector is $mN \times 1$, where N is the number of nodes and m is the number of possible states a node can be in (e.g., susceptible, infected, recovered in an SIR model).
- **Function** $G(\vec{P}_t)$: This is the non-linear function that defines how the probability vectors $\vec{P}_t$ change over time. It models the transition between each time-step, effectively capturing the dynamics of how the contagion propagates through the network.

4.1 SIR Example

In the case of an SIR VPM, the dimensions of each $\vec{P}_t$ vector would be $3N \times 1$ because an SIR model has three epidemiological states: Susceptible, Infected, and Recovered. The NLDS $G(\vec{P}_t)$ would then model how these states transition over time, taking into account the non-linear interactions between nodes.

Think of the entire process as a stop-motion animation. Each frame ($\vec{P}_t$) represents the state of the system at a particular time. The transition between frames is orchestrated by the non-linear function G, which dictates how each node's state evolves as the contagion spreads.

By using NLDS in this manner, one can capture the complex, non-linear interactions that occur in real-world systems, making it a powerful tool for understanding and predicting the dynamics of contagion spread through networks.

For example in, Figure 7, you can see that each probability vector has a size of $mN \times 1$ where N is the number of nodes in the graph and m is the number of states a node can be in with respect to the contagion. This allows us to keep track of how likely each node is in a particular contagion state at a particular time.

We can then use a NLDS $G(\vec{P}_t)$ to model how those probabilities change over time as the contagion propagates through the system. This process is kind of like a stop motion animation where each $\vec{P}_t$ is an image and the transition between each image time-step is defined by the non linear function 'G'.

For example if we wanted to simulate a SIR VPM as an aformementioned NLDS, the dimensions of each $\vec{P}_t$ vector would be $3N \times 1$ as an SIR model has 3 epidemiological states.

$$\vec{P}_{t+1} = G(\vec{P}_t)$$

where

$$G : \mathbb{R}^{mN} \to \mathbb{R}^{mN}$$

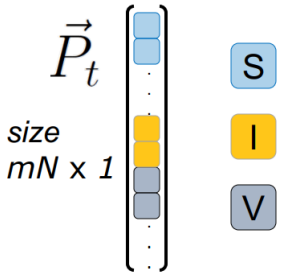


Figure 7: At each point in time in the simulation there is an $mN \times 1$ probability vector $\vec{P}_t$ representing how likely each of the N nodes are in any one of the m (in this case 3) contagion states. [3]

is defined as the following system of equations:

$$P_{S,i,t+1} = P_{S,i,t}\zeta_{i,t}(I)$$
$$P_{I,i,t+1} = P_{S,i,t}(1 - \zeta_{i,t}(I)) + (1 - \delta)P_{I,i,t}$$
$$P_{R,i,t+1} = P_{I,i,t} + P_{R,i,t}$$

Where $\zeta_{i,t}$ is the probability that the node i is not attacked by any of its infections neighbors. [5]

5 Analyzing Dynamics with Multiple Contagions

5.1 Synchronization

Synchronization is a special case where the long term behavior of a graph enters a periodic state. This typically happens when a SIRS disease periodically cascades through a highly connected parts of the system leading to a wave in infections and then slowly burns through more sparsely connected parts of the system to form a trough before burning through the highly connected parts of the system once all the recovered nodes have become susceptible again. Examples of this in the real world include measles cases in the UK as seen in Figure 8. [2]

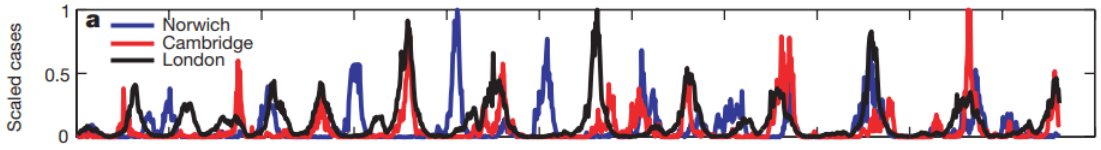


Figure 8: Case counts for measles across multiple cities in the UK. [2]

5.2 Multiple Competing Viruses

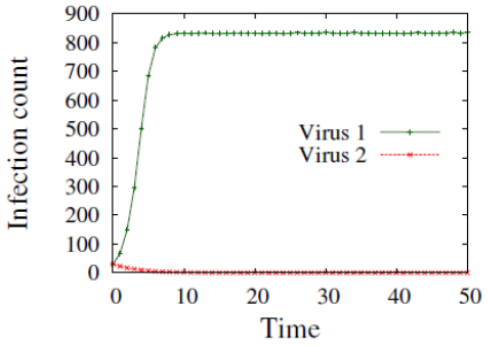
In general, when multiple viruses compete the strongest virus usually dominates. Where strength is measured by communicability and positioning. E.g. a highly transmissible on an isolated island is not going to dominate a less transmissible virus that starts out in a highly connected place like New York City. In addition, viruses that are too weak to propagate in the environment typically die out while mediocre propagators tend to go endemic.

5.2.1 Strong vs Weak

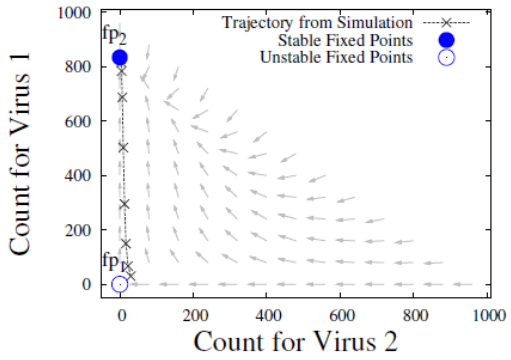
For example in the Strong vs Weak case in Figure 9a we can easily see from the Time-Plot that the Infection count of the stronger virus 1 will usually dominate the weaker virus 2 over time.

Then in the corresponding phase-plot in figure 9b we can see that there are two fixed points fp_1 and fp_2. Since virus 2 is weak, there is no stability fixed point for virus 2 as at any population it's case count will drop to zero. There is a semi unstable fix point threshold at fp_1, where if there is any minor positive perturbation in the case count for virus 1 the graph will inexorably go to fp_2 where the stronger virus 1 has dominated the case count. However as long as the case count for virus 1 remains zero, any amount of the weaker virus two cases will go to zero over time.

In addition the phase plot has a line representing the path followed by the time-plot in figure 9a



(a) Over time, the stronger Virus 1 dominates the weaker Virus 2 in terms of case counts. [4]



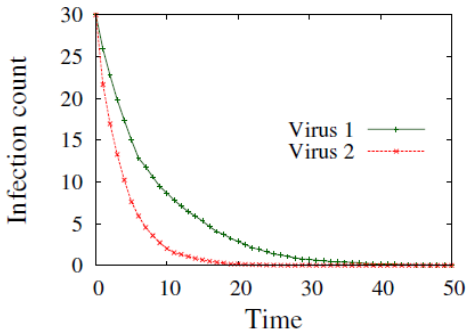
(b) In most cases, the simulation reaches the stable fp_2 where the stronger virus 1 prevails, except for the unstable fixed point fp_1." [4]

Figure 9: Strong vs Weak Virus

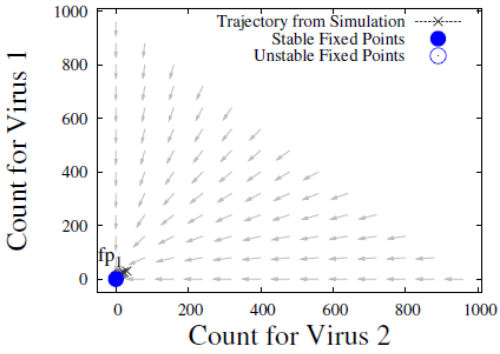
5.2.2 Weak vs Weak

In the case of two weak virus, the case counts for both will drop to zero over time no matter the initial starting point. An example of this can be seen in figure 10a

The fact that this drop to (0, 0) will happen for any initial case count for the weak viruses 1 and 2 is highlighted in figure 10b where all the trajectory arrows point to the stable fixed point fp_1 at (0, 0).



(a) Over time, both weak viruses die out. [4]



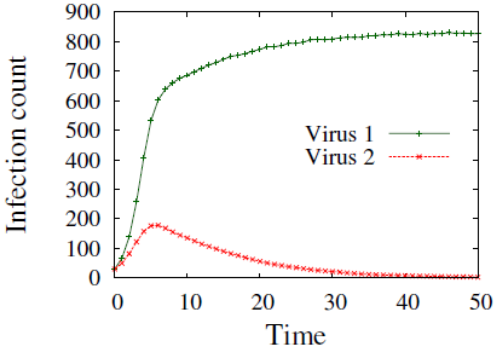
(b) The trajectory arrows for both weak viruses all end up in the fixed point fp_1[4]

Figure 10: Weak vs Weak Virus

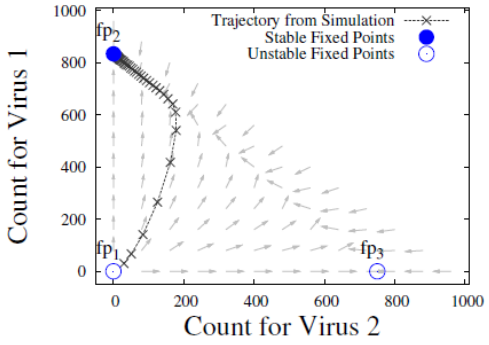
5.2.3 Strong vs Strong

Finally in the Strong vs Strong case, the stronger of the two viruses typically wins long term as seen in figure 11a.

However there are certain unstable fixed points where the less strong virus will win out (e.g. if the less strong virus is the only one introduced to the environment). Otherwise as evidenced by the trajectory arrows in 11b, we tend to end up in the stable fixed point (fp_2) where the stronger virus dominates.



(a) When two strong viruses compete the stronger one or the one in the better position will eventually win out. [4]



(b) The trajectory arrows typically converge to fp_2, where the stronger virus 1 dominates. Exceptions include fp_1 and the semi-stable fixed point fp_3. fp_3 acts as a stable sink when virus 1's case count is zero, and virus 2 has a non-zero positive count. [4]

Figure 11: Strong vs Strong Virus

5.2.4 Cooperation and mutual immunity

In addition to competing, viruses can also coexist in the same host at the same time. When this happens we often see examples of cooperation, partial mutual immunity, and full mutual immunity. For example in figure 12, we are using a modified flu-like SIIIS model to model

browser adoption where a user can install either Chrome, Firefox, both browsers, or neither browsers. As seen in figure figure 12 users have some probability β_1 and β_2 of adopting Chrome or Firefox respectively. Once they have adopted either one of the browsers they have an $\epsilon\beta_2$ and $\epsilon\beta_1$ chance of adopting the other browser as well. [3]

This ϵ value is called the "Interaction Factor" and is used to model how the browser "contagions" affect each other's adoption rate once they have been established. If the adoption of one browser fully excludes the adoption of the other, then ϵ is zero and this is case of "Full Mutual Immunity". However, if the adoption of one browser simply reduces the probability of adoption of the other, then ϵ is less than 1 (but greater than zero) and this is case of "Partial Mutual Immunity". Lastly if the adoption of one browser increases the probability of adopting the other, then ϵ is greater than one and this is case of "Cooperation". [1]

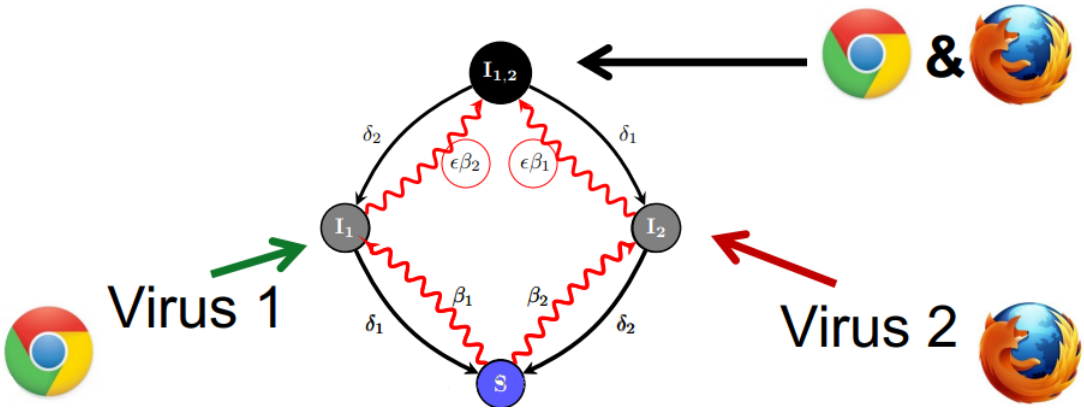


Figure 12: A modified SIIIS model for Web Browser adoption used to illustrate cooperation and mutual immunity. [3]

6 References

References

[1] A. Beutel, B. A. Prakash, R. Rosenfeld, and C. Faloutsos. Interacting viruses in networks: Can both survive? In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 426–434. ACM, 2012.

[2] B. T. Grenfell, O. N. Bjornstad, and J. Kappey. Travelling waves and spatial hierarchies in measles epidemics. *Nature*, 414:716–723, Dec. 2001.

[3] B. A. Prakash. CSE 8803: Epi Data Science for Epidemiology: Lecture 6: Dynamics of Models, Sept. 2022.

[4] B. A. Prakash, A. Beutel, R. Rosenfeld, and C. Faloutsos. Winner Takes All: Competing Viruses or Ideas on fair-play Networks. *WWW*, 2012.

[5] B. A. Prakash, D. Chakrabarti, M. Faloutsos, N. Valler, and C. Faloutsos. Threshold Conditions for Arbitrary Cascade Models on Arbitrary Networks. *ICDM*, 2001.

[6] P. Rai. In search of equilibrium, Sep 2018.

[7] A. Stanoev, D. Trpevski, and L. Kocarev. Modeling the spread of multiple concurrent contagions on networks. *PloS one*, 9:e95669, 06 2014.