Understanding the Micro and Macro diffusion dynamics of Influenza like viruses using Network Models

Jagdeep S. Gill
Graduate Student, MS Data Science
Indiana University, Bloomington, IN
jaggill@umail.iu.edu

ABSTRACT

The pandemics similar to **Spanish flu of 1918** are some of the high probability worse case scenarios that should concern us all. The global community should be prepared for the outbreak of such pandemics because they have the potential to cause death and destruction on a global scale and disrupt global economies. We live in a global world where human mobility has greatly increased and time to travel long distances has decreased due to Air transport networks. These changes to human travel patterns and increased speed of travel provide rapid long distance links for the transmission of infectious diseases such as influenza. According to a World Bank report, a serious pandemic can result in millions of deaths world wide and according to even most conservative estimates a pandemic can destroy up to 1% of the global GDP.[2] Therefore, understanding the mechanisms of spread of infectious diseases is important to control and contain these pandemics.

This research project explored the micro and macro diffusion dynamics of influenza like viruses using both the traditional epidemiological models such as the compartmental models, **SI, SIS and SIR models** and the Network models, especially the scale-free network models.[3] The compartmental models and the network models can help us understand the diffusion dynamics of influenza like infections and can help us effectively immunize susceptible individuals and quarantine infected individuals in order to effectively control and contain the outbreaks of these infections.

The compartmental model most appropriate for influenza outbreak is the **SIR model**.[3] Since most real world networks are scale-free networks, this project explored the properties of scale-free networks in spreading influenza like viruses. [3] This research study demonstrates the applicability of the concept of robustness of scale-free networks and the effectiveness of strategically immunizing or quarantining main hubs in these networks to control and contain the spread of these epidemics.[3]

In this research study the data published in the article, "Influenza in a Boarding School" in "British Medical Journal", 4 March 1978 [1] was used to model and simulate the micro level diffusion dynamics of influenza infection.[6] Next the network structure of international Airline networks was explored to understand the macro level diffusion dynamics of influenza like viruses. The data for the airport network analysis was obtained from Young Joon Oh's website. [7].

The airport network follows the power-law degree distribution and can be modeled using the scale-free networks with an exponent of **1.8**±0.2. [8]

We demonstrate that immunizing and/or quarantining the highest degree nodes is the most effective way to control and contain the outbreak of influenza like viruses at both the micro and macro levels.

1. INTRODUCTION

Throughout history pandemics have done more damage to human civilization than any war or other natural calamity. The recent outbreaks of **SARS**, **H1N1** ,**Ebola** and the **Zika Virus** have underlined the risks of spread of infectious diseases that have increased significantly with modern air travel. There are several research studies that have tried to model and predict the diffusion of influenza like viruses in the context of network models. Understanding the diffusion dynamics of spread of viruses is important to devise a efficient and effective mitigation strategy against the spread of pandemics. Since it is hard to quickly produce and administer vaccines in these scenario, it is important to focus on key segments of the population that can be vaccinated in order to achieve maximum benefits in terms of controlling the spread of these epidemics.

In this research study the data from the British Medical Journal .[1] was analyzed using the compartmental epidemic models as well as the network models. The epidemic model most suitable for this data is the SIR model. This was based on initial data analysis and the research paper on "Modeling Influenza epidemics and pandemics: insights into the future of swine flu (H1N1)", in BMC Medicine Journal [4] and the research paper, "The role of the airline transportation network in the prediction and predictability of global epidemics". [8]

Most real world networks are scale-free with power law degree distribution. The Airline network resembles the scale-free networks with a few high degree nodes (hubs). [3] Therefore, the global airline network was modeled using the scale-free network. The contact between students in boarding school can also be modeled as a scale-free network. The research paper, "Scale-free networks without growth" [9] suggests that the growth may not be a necessary condition to create scale-free networks such as the friendship networks of school children. The boarding school data was simulated using SIR model and scale-free network models.

1.1 Research questions

The main research question is to determine if the micro and macro diffusion dynamics of influenza like viruses follow SIR epidemic model and scale-free network models. The main motivation behind this research is to understand and apply the properties of SIR epidemic model and scale-free network models to immunize susceptible individuals and quarantine infected individuals in order to control and contain epidemics in most efficient and effective manner.

Keywords

Epidemic, Pandemic, Epidemic Threshold, Compartmentalization, Homogenous Mixing, Susceptible-Infected (SI) Model, Susceptible-Infected-Susceptible (SIS) Model, Susceptible-Infected-Recovered (SIR) model, Scale-free networks

2. METHODS

The methodology used in this research project is to study the diffusion dynamics of influenza like viruses at a micro and macro level using epidemiological and network models. The simulations using these models will help us understand the diffusion dynamics and the strategies to control and contain these epidemics. The micro diffusion dynamics were studied using data for the spread of influenza virus in a boarding school in UK in 1978. There were 763 students in the boarding school. The influenza infection started with one student that had travelled to Hong Kong during school break and returned to boarding school with influenza infection. This data was initially analyzed using a simple visualization as shown in Figure 1 below.
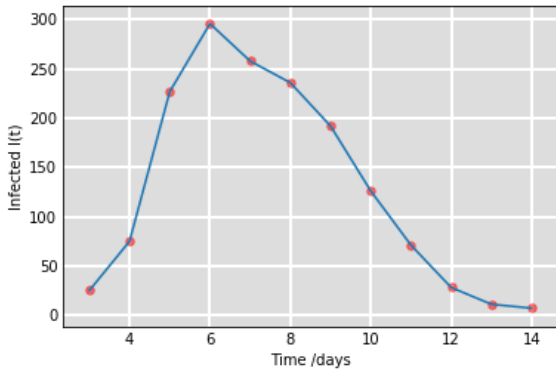

Figure 1: Graph showing the diffusion of Influenza epidemic in boarding school in UK in 1978.

As shown in Figure 1, the spread of influenza among boarding school students followed a curve similar to SIR epidemic model curve. The infection peaks around day 6 with a maximum of 296 students infected with influenza at that point in time. The number of infected students then drops after this peak and only about 7 students were left with infection on day 14.

2.1 Selecting the appropriate Epidemiological Model

The first step is to select the appropriate epidemiological model to understand the diffusion dynamics of influenza in-

fection. There are three basic models to choose from such as, SI, SIS and SIR models. The SIR model is most suitable to model an infection like Influenza. [4] This model compartmentalizes the population into three states or compartments; Susceptible (S), Infectious (I) and Recovered (R). The susceptible individuals are all healthy individuals in the population who have not contracted the infection yet. The infectious individuals are those who have contracted the infection and can infect others. The recovered are the individuals who have been infected before, but have recovered and have developed immunity and are not infectious anymore.[3]

The SIR model was implemented in python with the number of nodes equal to student population in the boarding school and the initial infection rate of 0.05%. The generic SIR model for this data set is shown in **Figure 2** below.
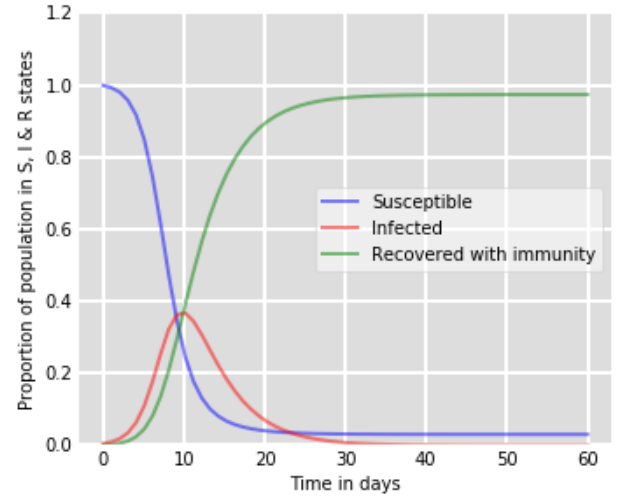

Figure 2: The Susceptible-Infected-Recovered (SIR) Model, implemented in Python.

As shown in **Figure 2** above, initially the entire population is in **Susceptible (S) state**. As the infection starts to spread in the population, some fraction of the population gets infected and move to **Infected (I) state.** The rate of transition from S to I state depends on the effective contact rate, beta (β). Some fraction of infected individuals either die of the disease or are recovered and acquire immunity from infection. This transition from I to R state depends on the mean recovery rate gamma (γ). These recovered/removed individuals are moved to **Recovered (R) state.** For simplicity we assumed that all recovered individuals have 100% immunity to further infections from the same viral strain going forward. Another assumption of this model is that the population size N is static. The differential equations that were used to implement the SIR model are given below.

$$ds/dt = -\beta * S * I/N$$
$$ds/dt = \beta * S * I/N - \gamma * I$$
$$ds/dt = \gamma * I$$

Here β is the effective contact rate of the disease, meaning

an infected individual comes into contact with βN other individuals per unit time. The fraction that is susceptible to contracting the infection is S/N. γ is the mean recovery rate, meaning $1/\gamma$ is the mean period of time during which an infected individual can pass infection to others. [5] In the final regime of SIR model all infected nodes are recovered and therefore the number of infected individuals goes to zero over long period of time. $i(\infty) = 0$. [3]

2.1.1 Choosing appropriate values for β and γ

There are three key parameters required for SIR model simulation. These parameters are β, γ and the number of individuals initially infected (I_0). The β value determines how fast the infection spreads through susceptible population and the γ value determines how fast infected individuals recover from infection. Data visualization tools were used to find the appropriate values of β and γ for the boarding school data set for the simulation. First the γ values were changed while β values were kept constant at 0.3 as shown in Figure 3 below.
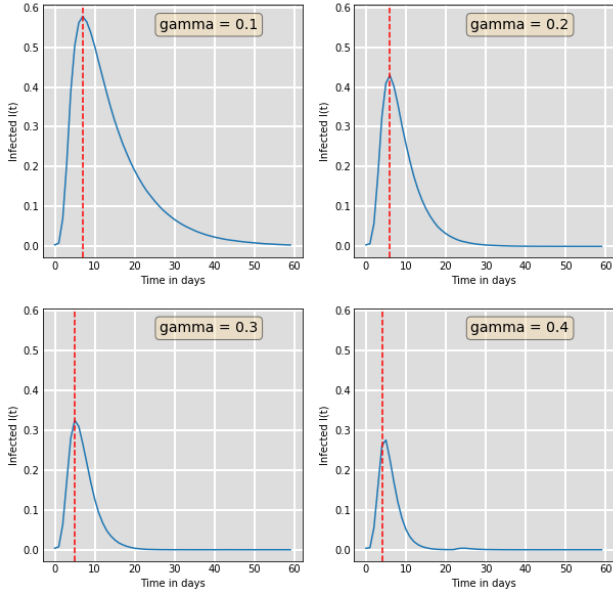

Figure 3: Understanding the effects of different γ values on diffusion dynamics of influenza in boarding school population. The β values were kept constant at 0.3.

As can be seen in Figure 3, the γ value that gives the SIR curve similar to the plot of our data in Figure 1 is approximately $\gamma = 0.25$. This SIR curve has a peak around day 6 of about 39% of the population being in infected state. The majority of students who contracted influenza infection had recovered by day 14 when only 7 students were still in infected state. These results can be simulated by using a $\gamma = 0.25$.

Next the optimal β value for the simulation was determined by varying the β and keeping γ constant at 0.25 as shown in Figure 4. The optimum β value that gives SIR curve similar to boarding school data plot is a $\beta = 0.3$.

The 3rd parameter required for this simulation is the initial infection rate. This value was calculated from the source data, since initially at t_1 only 3 students got the infection, this value was set at 0.005 which is equal to approximately 3 students, **initial** $infection(I_0) = N * 0.005$
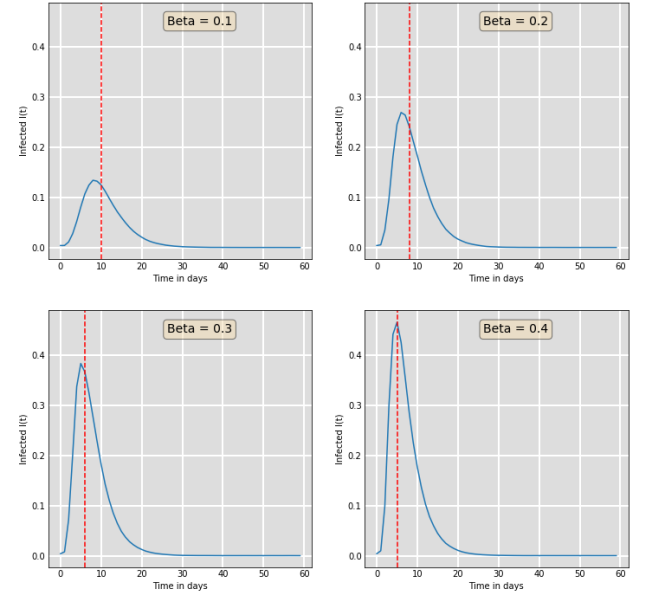

Figure 4: Understanding the effects of different β values on diffusion dynamics of influenza in boarding school population. The γ values were kept constant at 0.25.

2.2 Simulation of diffusion dynamics of influenza using network models

The simulation using the random and scale-free networks was performed using number of nodes (N) equal to population in boarding school data. The first step in the simulation was to create a random network and a scale-free network with nodes (N = 763).
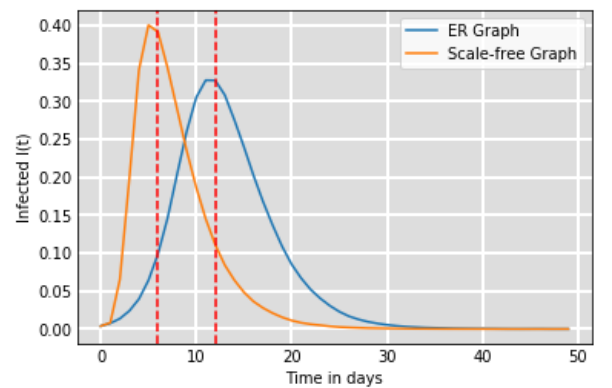

Figure 5: The comparison between ER and scale-free graphs using boarding school data.

Next the SIR curves for random and scale-free networks were plotted as shown in Figure 5. The comparison of these

curves shows that the Scale-free network simulation curve gives a better fit for diffusion of influenza in boarding school compared to the curve for random graph. In scale-free network model, the number of infected peaks around day 6 and is about 39% of the population. This is similar to the graph in Figure 1 for the boarding school influenza epidemic. However, in the random graph, the infection peaks around day 12 and is only 33% of the population. Even though, it is possible to change the parameters β, γ and initial infection rate to fit the data to actual curve, the curve for scale-free network will always be higher and to the left of random network curve. This is because of the unique properties of scale-free networks such as the hubs with power-law degree distribution that can quickly spread the infection to large segment of the population compared to random networks.

2.3 Simulation of immunization strategies using scale-free network models

The effects of random immunizations and selective immunizations of main hubs in scale-free networks was compared. The scale-free network with N=763 and power-law exponent of 2.5 was created using alpha=0.35, beta=0.60, gamma=0.05, delta_in=0.4, delta_out=0.4. The simulation was performed as per below.

1. Randomly remove 5, 10, 15 and 20 nodes from scale-free network.

2. Select highest degree nodes in scale-free network and sort them in descending order in a list. Remove 5, 10, 15 and 20 nodes from the list.
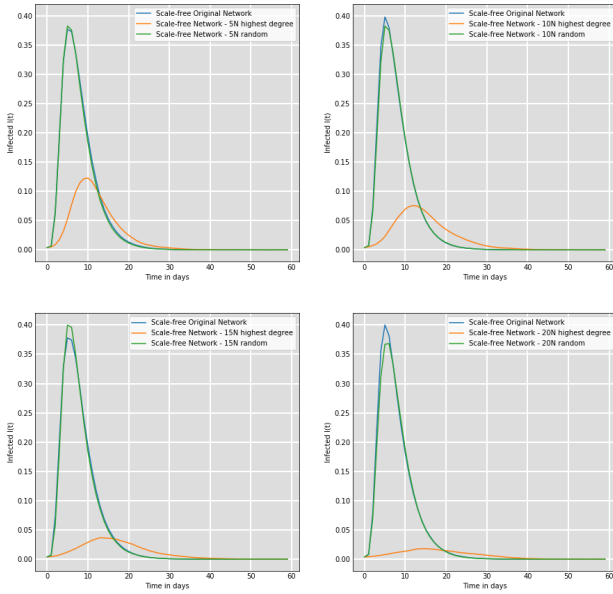

Figure 6: Comparing the effects of randomly and strategically removing 5, 10, 15 and 20 nodes from scale-free network to simulate the effectiveness of random and strategic immunization in scale-free network.

Next the SIR curves for the original scale-free network were plotted along with SIR curves after randomly and strategically (highest-degree nodes or hubs) removing nodes. Four subplots were created to compare the effects of removing 5, 10, 15 and 20 nodes respectively as shown in Figure 6. There seems to be no effect of randomly removing nodes in this scale-free network because these curves almost overlap the original curves in each subplot. However, strategically removing the highest degree nodes greatly impacts the SIR diffusion curves. Therefore, strategically immunizing the hubs in scale-free networks would be most effective strategy to control and contain epidemics.

3. MACRO DIFFUSION DYNAMICS OF INFLUENZA ON AIRLINE TRANSPORT NETWORK

The macro diffusion dynamics of influenza were studied using the airline transport networks. Understanding the network structure of airline networks is important to strategically screen, quarantine and control spread of epidemics in modern globally connected world. According to the research study, "The role of the airline transportation network in the prediction and predictability of global epidemics", air-transport-network properties can help explain and predict the spread of epidemics. According to this study, the worldwide airport network follows a power-law degree distribution with exponent **1.8**±0.2. [8] Therefore, understanding the airline transport network using the properties of scale-free networks can help us understand the diffusion dynamics of influenza like viruses at a macro level.

3.1 Airline-network data analysis

The Airline-network data was downloaded from http://openflights. org/. The nodes and edges data was then loaded into Gephi. The network statistics were calculated in Gephi as shown in Figure 7.

The Airline-network is a directed graph with 5,623 nodes and 37,596 edges. The average degree is 6.686 and the average weighted degree is 12.877.

As shown in Figure 8, the Airline-transport network follows the power-law degree distribution with exponent **1.8**±0.2. Therefore, we can apply the properties of scale-free networks such as average degree, clustering coefficient, centrality, modularity, assortativity and inhomogeneity to understand the diffusion dynamics of influenza like viruses.

4. RESULTS

The results of modeling and simulation experiments using boarding school data show that influenza like viruses spread much quicker on scale-free networks compared to random networks. Therefore, further simulations were done using the scale-free networks to understand the effects of immunizing susceptible nodes or quarantining the infected nodes. The simulation was performed by removing nodes randomly and preferentially in order of highest degree. The results of this simulation show that removing the nodes randomly has no effect on diffusion dynamics of influenza in boarding school network. However, when the same number of highest degree nodes were selectively removed it dramatically reduced the infection rate as shown in Figure 6. The effect

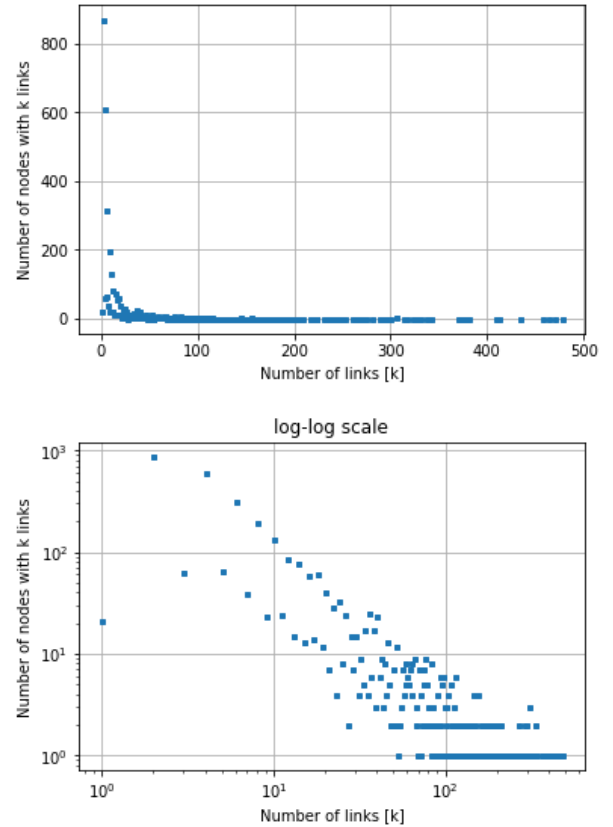**Figure 7: Airline network statistics calculated using Gephi.**

Figure 8: Airline transport network degree distribution in linear and log-log scale demonstrating power-law degree distribution of scale-free networks.

on infection rate seems to be linearly related to number of highest degree nodes that were removed.

The results of data analysis of airline network shows power-law degree distribution as shown in Figure 8. Therefore, the properties of the scale-free networks can be used to model and simulate the diffusion dynamics of influenza infection on airline transport networks.

5. DISCUSSION

The main purpose of this investigation was to understand the micro and macro diffusion dynamics of influenza infection in context of underlying network structure. The SIR epidemic model seems to capture the diffusion dynamics of influenza in boarding school. The boarding school population N is heavy-tailed distributed as per the results of Zipf's law. [8] Therefore, using scale-free network structure to understand the diffusion dynamics of influenza in boarding school was the right choice. It was demonstrated using the simulation methods that as predicted based on the literature research, strategically removing the high degree nodes or hubs is the most effective way to disrupt the diffusion processes on scale-free networks. However, the random immunizations are totally ineffective in controlling and containing the spread of influenza like viruses. Since the airline network is also scale-free, the same principles can be

applied to screen, immunize and quarantine travelling population during an outbreak of serious influenza like viruses.

6. CONCLUSION

This research study confirms that the SIR model is the most appropriate model to simulate the spread of influenza like viruses in a fixed size population. The findings of this research study are consistent with the literature research indicating that the human networks and airline networks follow scale-free network models. Therefore, based on the SIR model and simulations of disrupting hubs in scale-free networks by strategically removing nodes, we can devise a strategy to control and contain the spread of influenza like epidemics. The most effective approach for screening, immunizing and quarantining travelling populations would be to find the hubs in the underlying networks. However, finding the hubs can be easy in airline networks but finding the hubs in human networks can be challenging. There are additional concepts such as the "Friendship paradox" that can be applied to identify the hubs in a social human networks for effective screening, immunization and quarantine. The idea behind this concept is that by following the edges from some random nodes to their friend's friends, we can reach the hubs in human networks. This concept was not the focus of this research study but can be explored further in light of findings of this research study.

7. REFERENCES

[1] Influenza in a boarding school. *BRITISH MEDICAL JOURNAL*, 1978.

[2] website, July 2017. Retrieved from, "http://www.worldbank.org/en/topic/pandemics", on Feb 12, 2018.

[3] A.-L. Barabasi. *Network Science*. Cambridge University Press, 2016.

[4] B. G. W. Brian J Coburn and S. Blower. Modeling influenza epidemics and pandemics: insights into the future of swine flu (h1n1). *BMC Medicine*, June 2009.

[5] C. Hill. *Learning Scientific Programming With Python*. Cambridge University Press, 2015.

[6] M. Martcheva. *An Introduction to Mathematical Epidemiology*. Springer, 2015.

[7] Y. J. Oh. Visualizing airline routes network. online, 2017.

[8] M. B. Vittoria Colizza, Alain Barrat and A. Vespignani. The role of the airline transportation network in the prediction and predictability of global epidemics. *PNAS*, 103(7), Feb. 2006.

[9] B.-H. W. Yan-Bo Xie, Tao Zhou. Scale-free networks without growth. *ScienceDirect*, July 2007.

The code for this project is available on public github address below.

https://github.com/Jaggill2/NetworkSciFinalProjectGillJag.git