Implementing an SIR Model on Various Network Topologies

May 26, 2021

Kenny Chen

1 Introduction

As the COVID-19 pandemic continues, we are reminded and encouraged to stay six feet apart from others as to not contract the disease from infected persons. This motivated us to explore how contact network modeling can begin to give us an idea at how a disease spreads throughout a population.

1.1 Creating A Network Graph

We can model social networks through a mathematical framework of a network where we have a graph $G = (V, E)$ where G consists of V, a set of vertices (also called nodes) and E, a set of edges (also called links). In our case, nodes represent different individuals and edges can be thought of as the interaction or connection between two individuals.

1.2 Network Models

There are several ways to construct a graph with various characteristics and shapes and for our purposes, we investigate the spread of diseases on three network models to simulate the various networks we might encounter in real life.

- **Erdos Renyi Random Graph**: The classical random graph model popularized by Erdos and Renyi (1959) starts with a collection of graphs with the same given number of nodes and edges and gives each graph equal probability. In mathematical terms, we can define a collection of graphs as $\mathscr{G}_{N_v,N_e}$ of all $G = (V, E)$ with the same given number of nodes and edges. Then we assign probability $\mathbb{P}(G) = \binom{N}{N_e}^{-1}$ for each $G \in \mathscr{G}_{N_v,N_e}$ where $N = \binom{N_v}{2}$ is the total number of distinct pairs (Kolaczyk 2014).
- **Watts-Strogatz Small-World** : The small world network structure identified by Wattz and Strogatz (1998) aims to create networks with high clustering and low average path lengths. They were motivated by their observations of these types of networks in the real world that had high clustering but low distance between nodes. To create a small world, we begin by starting with the graph in a lattice structure made of n vertices and connect them to r of its neighbors. Then we "rewire" the edges by moving one end of each edge to another node with probability p (Kolaczyk 2014).
- **Preferential Attachment Model**: Another type of model we want to explore is the preferential attachment model created by Barabasi-Albert (1999) which embodies the idea of "the rich get richer". This type of model works like this: we start with an initial graph $G^{(0)}$ of $N_v^{(0)}$

vertices and $N_e^{(0)}$ edges. Then at each step $t = 1, 2, ...$, the current graph is changed to create a new graph by adding a new vertex with some degree m and the m edges are attached to m different vertices of the current graph. The probability of the new vertex being connected to an old vertex v is given by $\frac{d_v}{\sum_{v' \in V} d_{v'}}$. We can see that there is a preference towards those with higher degrees and after t steps, we will start to see vertices with very high degrees emerging (Kolaczyk 2014).

- **Zach's Karate Club**: While this is mostly a simulation study, we also wanted to explore an actual network with real individuals and relationships between individuals. This network depicts the interactions between members of karate clubs observed by Wayne W. Zachary (1977).

1.3 SIR Model

With these various network topologies, we are curious if the spread of diseases changes depending on the structure of the network and to model disease spread, we use a basic SIR model which is a compartmental model that places individuals into one of three compartments: susceptible, infected, or recovered. However, we won't be looking at the mathematical version but relying on the algorithm implemented in the python package (EoN) for the textbook "Mathematics of Epidemics on Networks" by Kiss, Miller, and Simon (2017).

This algorithm simulates continuous time SIR epidemics in static networks with transmission rates τ and recovery rates γ. This implementation uses a priority queue Q where events are stored and the first is executed. If the event is a transmission and the individual is susceptible, then they become infected. Their recovery is inserted into Q and transmission to incidental nodes may be added. If the event is recovery, the node then recovers. A picture of the algorithm from the book is shown below.

In our SIR model, we chose our tau = 0.056 and gamma = 0.0455 based on the paper "Estimation of SIR model's parameters of COVID-19 in Algeria" by Dilip and Lounis (2020). Moreover, we begin each infection by targeting the node with the highest degree centrality to see how it spreads when we might have a "super spreader".

[2]:

```
Input: Network G, per-edge transmission rate τ, recovery rate γ, set of index node(s)
  initial_infecteds, and maximum time t_max.
Output: Lists times, S, I, and R giving number in each state at each time.

  function fast_SIR(G,τ, γ, initial_infecteds, t_max)
      times, S, I, R ← [0], [|G|], [0], [0]
      Q ← empty priority queue
      for u in G.nodes do
          u.status ← susceptible
          u.pred_inf_time ← ∞
      for u in initial_infecteds do
          Event ← {node: u, time: 0, action: transmit}
          u.pred_inf_time ← 0
          add Event to Q                                   ▷ ordered by time
      while Q is not empty do
          Event ← earliest remaining event in Q
          if Event.action is transmit then
              if Event.node.status is susceptible then
                  process_trans_SIR(G, Event.node, Event.time, τ, γ, times, S, I, R, Q, t_max)
          else
              process_rec_SIR(Event.node, Event.time, times, S, I, R)
      return times, S, I, R
```

1.4 Network Characteristics

For our three networks, they are randomly generated with a pre-determined seed. Each network consists of 500 nodes and the choice for number of edges was chosen arbitrary to the extent that we are putting in enough edges such that the overall density of the network is approximately 10%.

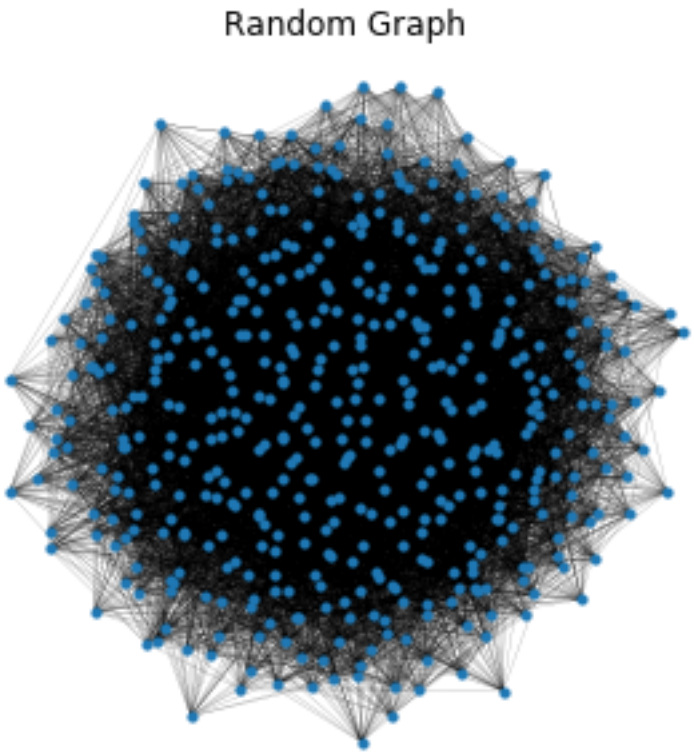
In the figure below, we see each network shares very similar network characteristics except for the clustering coefficient (frequency of connected triples to close and form triangles).

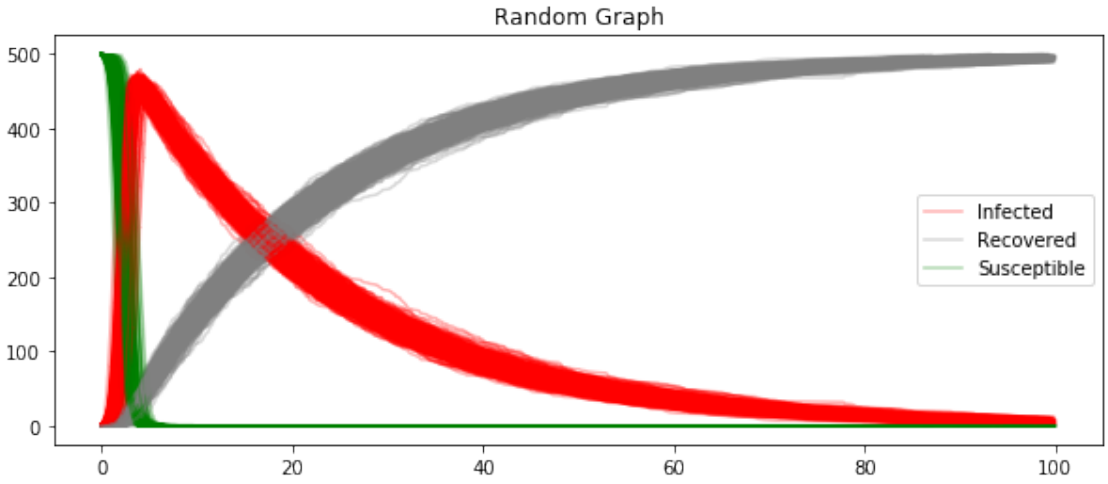
	Attribute	Erdos-Renyi Random	Watts-Strogatz Small-World	Barabasi-Albert Preferential
0	Number of Nodes	500	500	500
1	Number of Edges	12515	12500	12771
2	Density	0.100321	0.1002	0.102373
3	Average Degree	25.03	25	25.542
4	Average Path Length	1.90549	1.91284	1.91346
5	Transitivity	0.0997073	0.162317	0.181909
6	Clustering Coefficient	0.0999317	0.163431	0.186897
7	Is fully connected	True	True	True

	Attribute	Karate Club
0	Number of Nodes	34
1	Number of Edges	78
2	Density	0.139037
3	Average Degree	2.29412
4	Average Path Length	2.4082
5	Transitivity	0.255682
6	Clustering Coefficient	0.570638
7	Is fully connected	True

2 Results

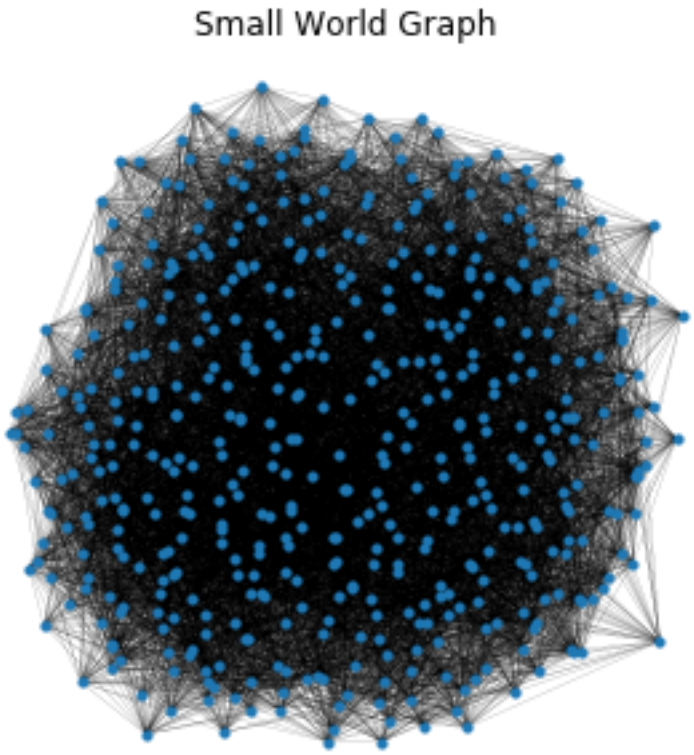
2.1 Random Graph

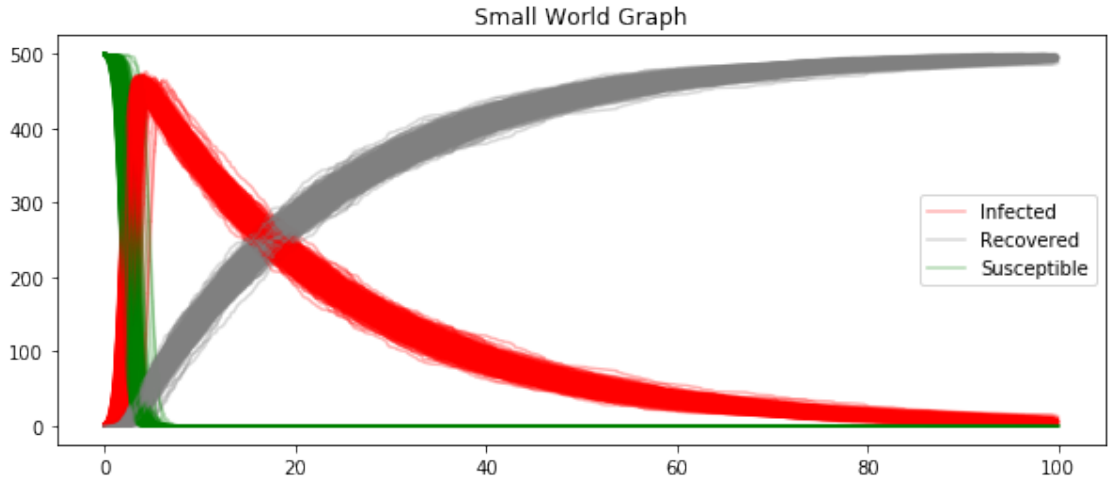




Running 500 simulations, we see that nearly all the nodes get infected at some point given our parameters in the SIR model. Nearly all nodes get infected by around 5-7 days marks. We also notice that the results are also pretty consistent throughout the 500 simulations as indicated by the thick bands.

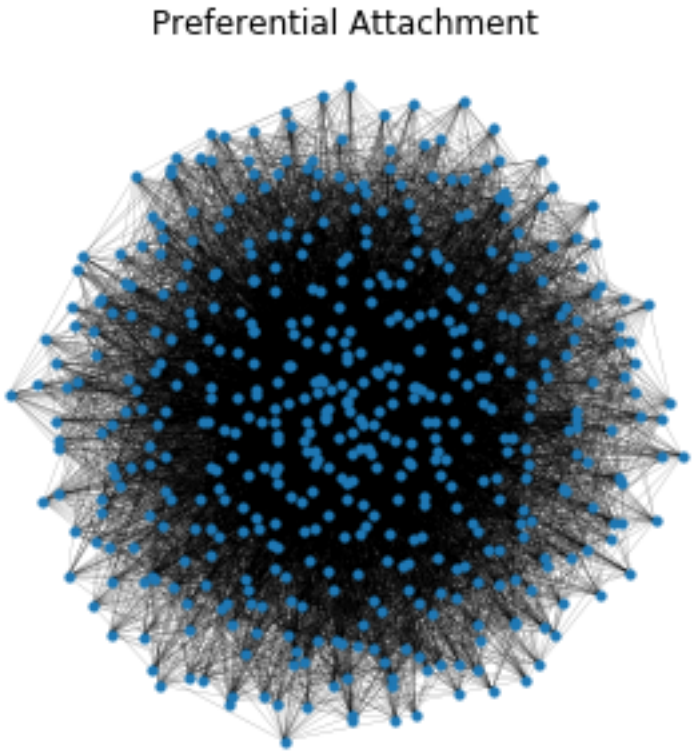
2.2 Small World

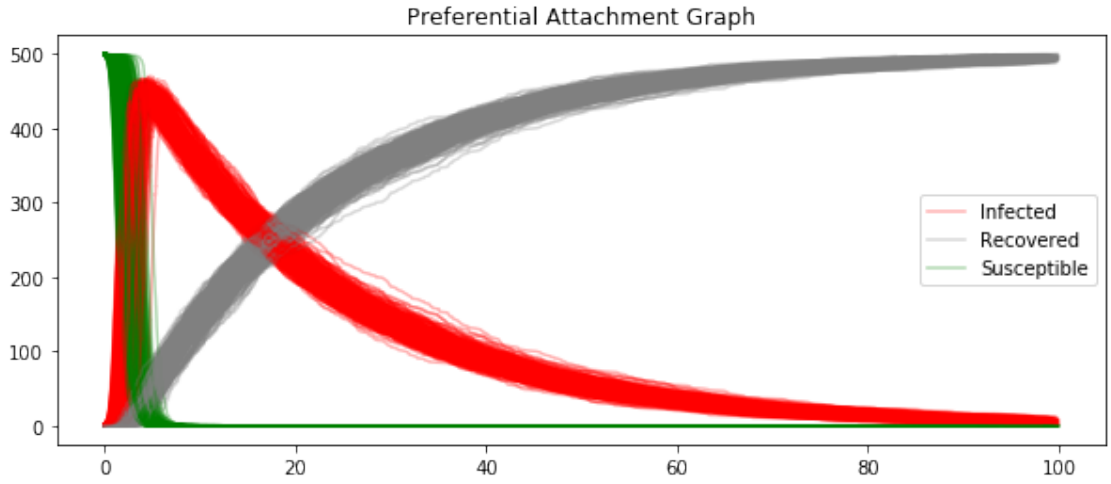




In the small-world graph, we get really similar results to the random graph where nearly all the nodes get infected at some point given our parameters in the SIR model and it is around the 5-7 days mark. We also notice that the results are also pretty consistent throughout the 500 simulations as indicated by the thick bands.

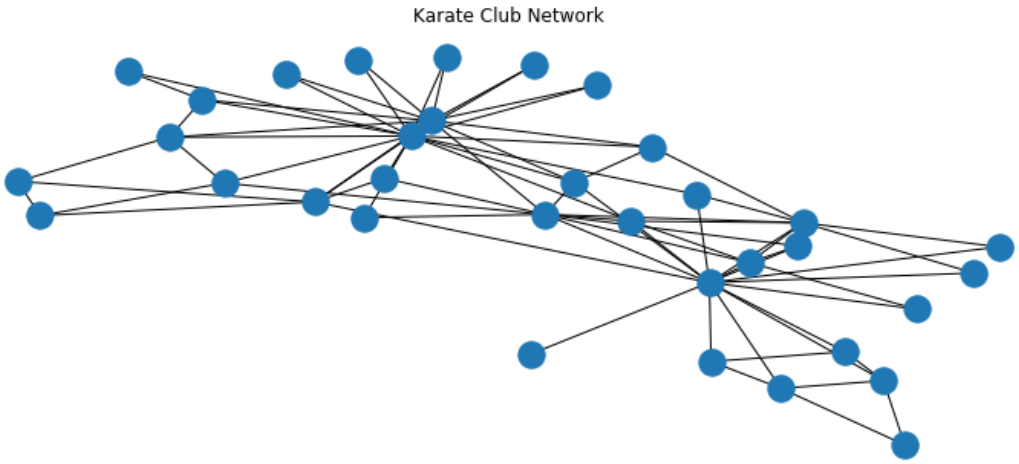
2.3 Preferential Attachment Graph

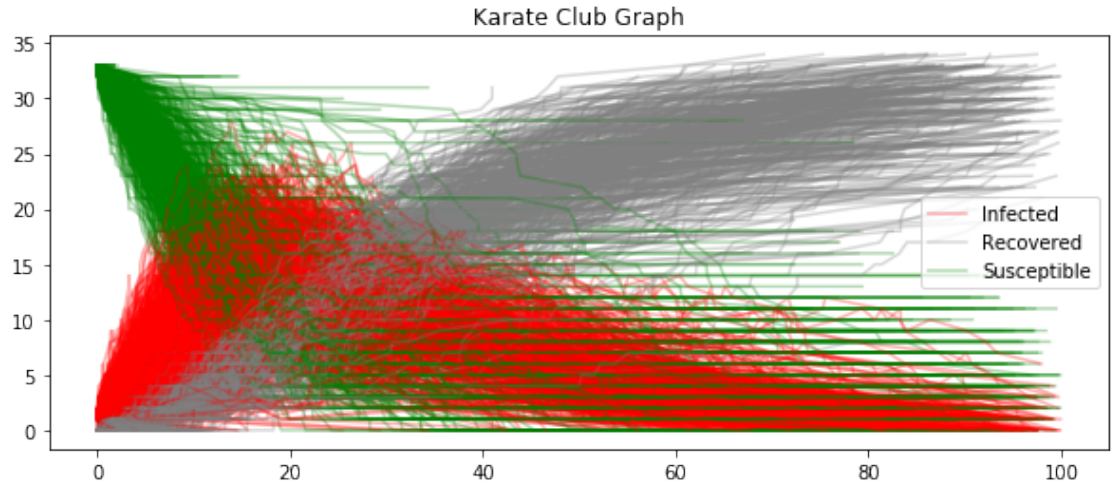




In the preferential attachment graph, the results are also similar to both the random and the small-world. We see this exponential rise in cases where nearly every node gets infected at some point. Given these networks have very similar characteristics, tuning some of the network parameters slightly might see more changes.

2.4 Karate Club





In the karate club, due to there being less nodes, the variation between simulations seems way larger. We still notice a really quick rise then slow dip in infections as seen in the previous one but the band is not as tight as the others.

3 Discussion

In the three simulated networks, we see very similar trends in the rise of infected nodes. They all exhibit exponential rises, peaking around days 5-7 and nearly infecting the entire network population of 500 nodes. Moreover, the graphs also display a level of consistency when running through the 500 simulations as seen by the thick bands. Whereas in the karate club network which is a much smaller network, we see more variability in the simulations. These results might be explained by the fact that the networks as shown above had really similar characteristics besides the clustering coefficient.

3.1 Limitations

There were several limitations to this simulation study, from network formation to model implementation. The three networks we looked at are generally not representative of real social networks although small-world and preferential attachment may try to replicate certain aspects of real world networks. Furthermore, we also only considered static networks while social networks are dynamic and relationships/contact often form or dissolve over time.

In addition, the SIR model is a rudimentary model that only considers three possibilities and places unrealistic assumptions such as assuming infections happen instantaneously and recovery rates are independent of time. We also didn't consider stochastic effects as the SIR model is a deterministic one. Given more time, exploring dynamic networks and stochastic models would be the next step in modeling infectious diseases.

3.2 Conclusion

We did not see noticeable differences in SIR models on the three simulated networks which raises the question of what network parameters matter the most in the spread of diseases. In a future simulation study, we would like to see how tuning various parameters affect the spread in

combination with various network topologies and using more complex models that considers vaccination rates or the ability for people to get reinfected again. While that remains a problem for another day, our study shows the usefulness of networks in modeling diseases and can allow us to have a better understanding of infection diseases and could serve as a useful tool to inform policy and public health decisions.

4 Citations

A. Barabasi, R. Albert, Emergence of scaling in random networks. Science 286(5439), 509–512 (1999)

D. Watts, S. Strogatz, Collective dynamics of 'small-world' networks. Nature 393(6684), 440–442 (1998)

Kiss, Miller, and Simon, "Mathematics of Epidemics on Networks" (Springer, 2017).

Kolaczyk, Eric D., Csardi Gabor, Statistical analysis of network data with R. New York: Springer, 2014.

Lounis, Mohamed, and Dilip Kumar Bagal. "Estimation of SIR model's parameters of COVID-19 in Algeria." Bulletin of the National Research Centre vol. 44,1 (2020): 180. doi:10.1186/s42269-020-00434-5

Miller et al., (2019). EoN (Epidemics on Networks): a fast, flexible Python package for simulation, analytic approximation, and analysis of epidemics on networks. Journal of Open Source Software, 4(44), 1731, https://doi.org/10.21105/joss.01731

P. Erdős, A. Rényi, On random graphs. Publ. Math. Debrecen 6(290), 290–297 (1959)

Zachary W. (1977). An information flow model for conflict and fission in small groups. Journal of Anthropological Research, 33, 452-473.