Modeling Worm Propagation in Large Networks

CS556 Computer Security Assignment Project 1

Chandrima Ghatak

Department of Systems Engineering
Colorado State University

Contents:

1. Project Introduction

For this Assignment project, we are simulating the propagation of worms in large networks. Three different network topologies are being used for this purpose:

1. Erdos-Renyi
2. Barabasi-Albert
3. Watts-Strogatz

We are using Python for programming, and the *networkx* package.

Initially, the networks are generated and stored in CSV files as pairs of nodes (representing an edge). These networks comprise three topologies, with 1000, 1500 and 2000 nodes for each network type.

For the three topologies, we ran the experiment 10 times in each network, and we recorded the average time for completing one round of simulation (1 round = 1-time unit).

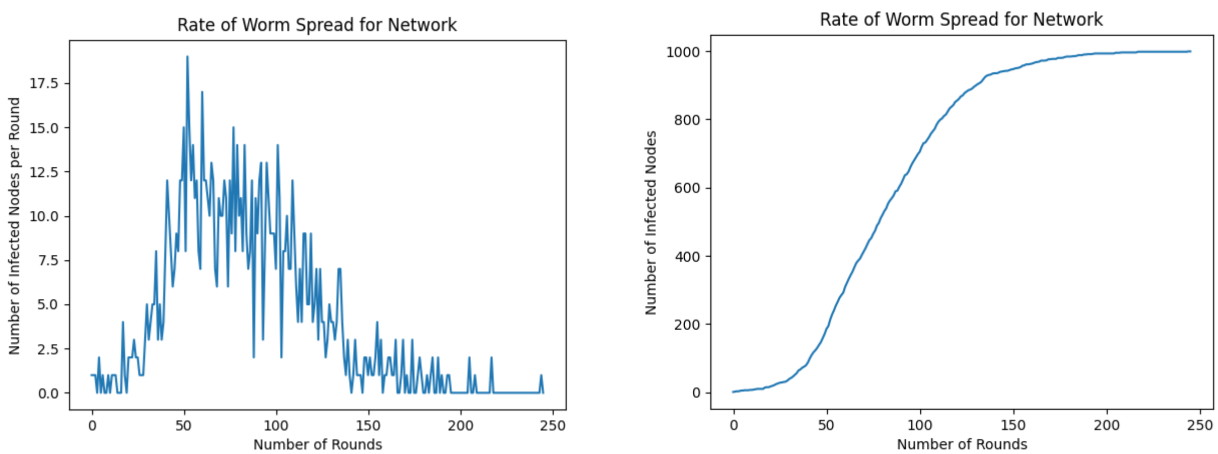
We have two types of observation for different networks and their topologies:

i) **Propagation without a cure:** A graph is created with the number of rounds vs the number of infected nodes at the end of infection of all nodes.
ii) **Propagation with a cure:** A graph is created with the number of rounds vs the number of infected nodes at the end of curing all nodes.

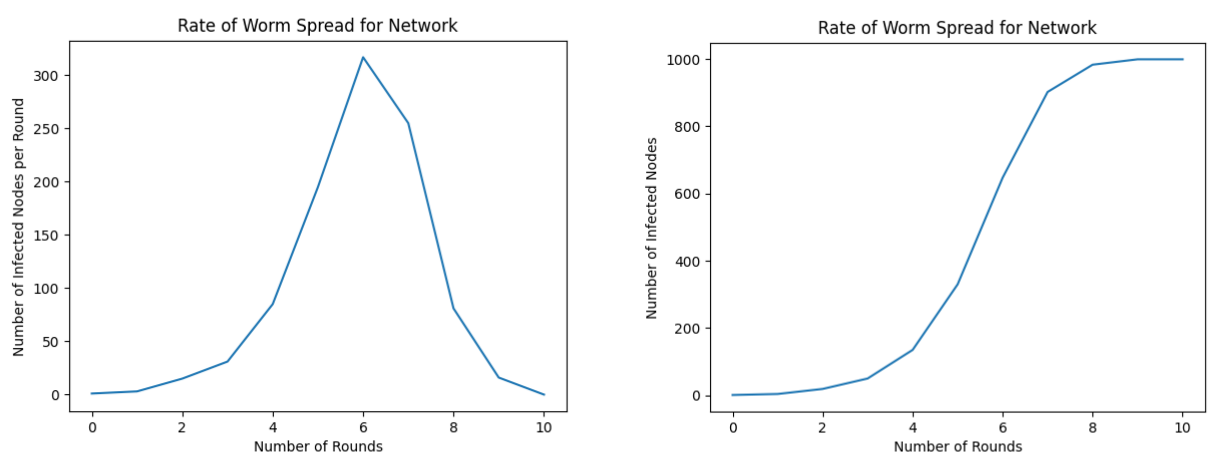
2. *Observations based on the propagation graphs in different networks without a cure*

Our first task for the project was to simulate the propagation of worms within a network. As discussed, we have three networks with 1000, 1500, and 2000 nodes each. We took the initial point of infection = 1, and probability of infection = 0.01 for each case, we then run the simulation a total of 10 times. We record the average number of rounds taken to finish the simulation in each case. Additionally, we assigned a probability of 0.01 to the worm, which indicates the previous node selects each uninfected node adjacent to an infected node, in order to infect it, with a probability of 0.01.

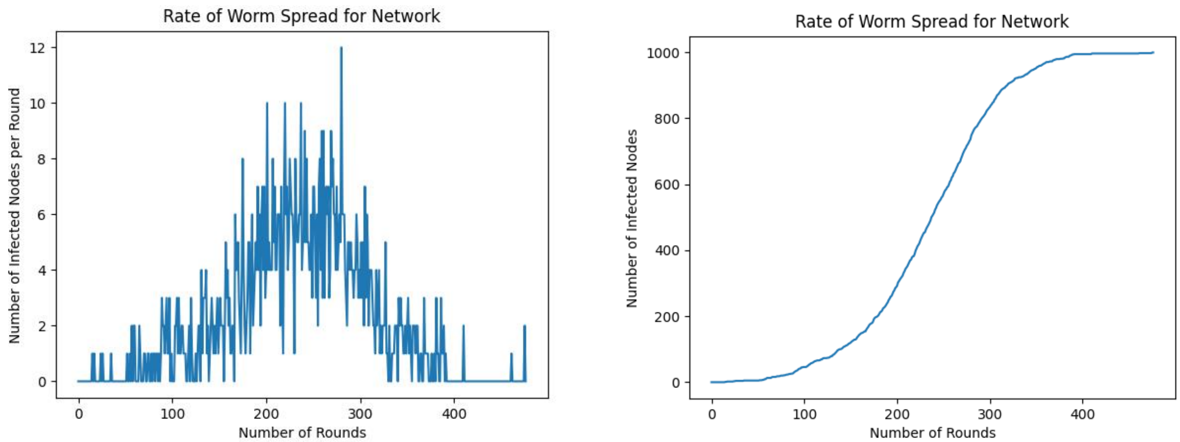
For 1000 nodes:



Barabasi-Albert: Average Number of rounds = 246.1

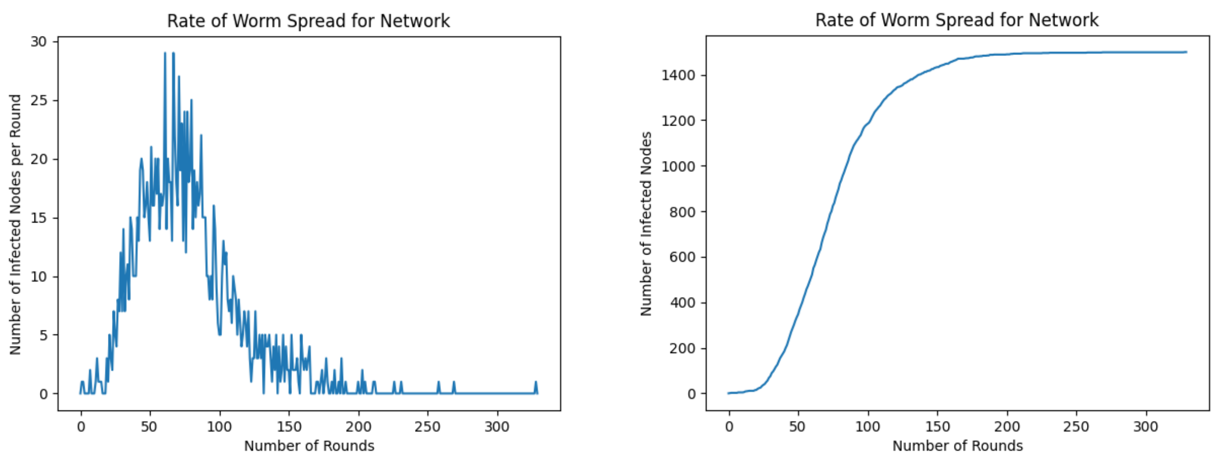


Erdos-Renyi: Average Number of rounds = 11.0

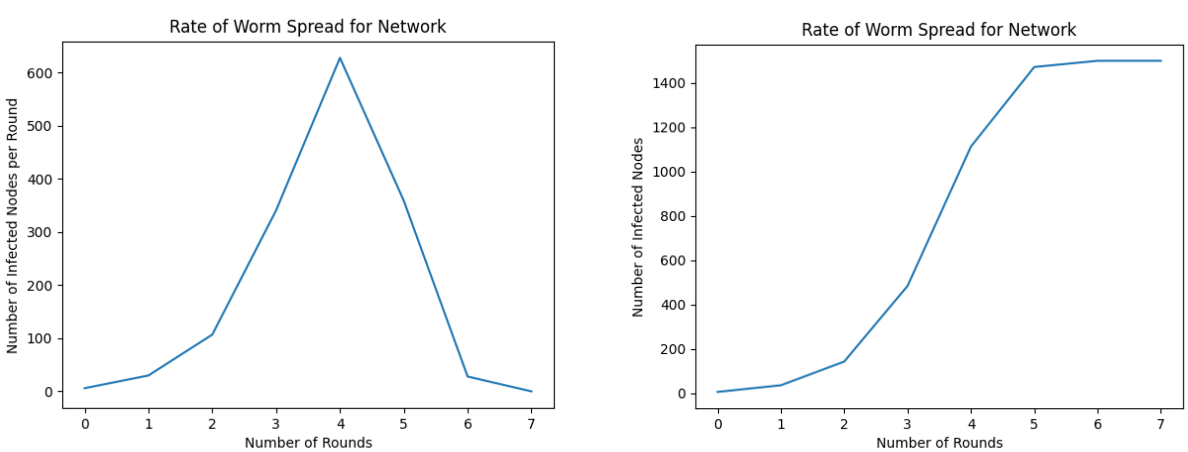


Watts-Strogatz: Average Number of rounds = 477.6

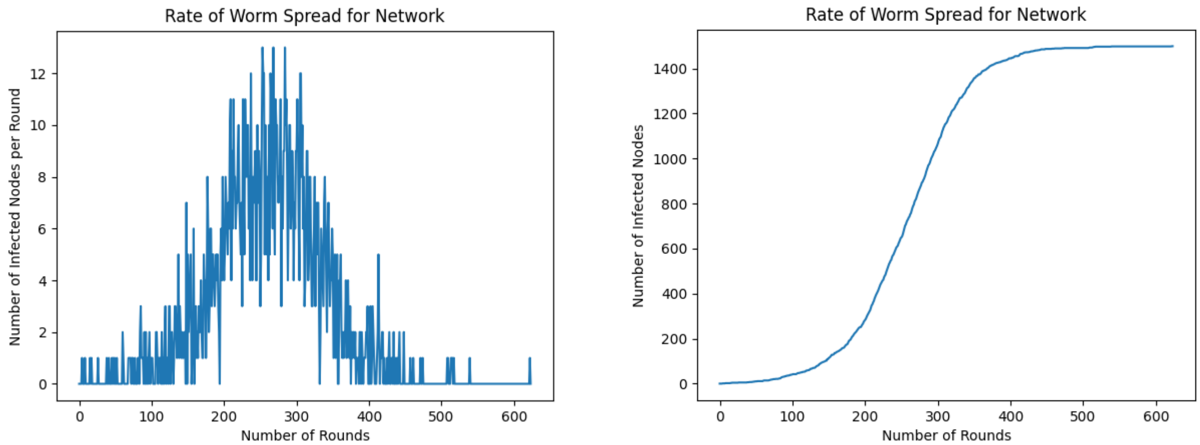
For 1500 nodes:



Barabasi-Albert: Average Number of rounds = 330.5

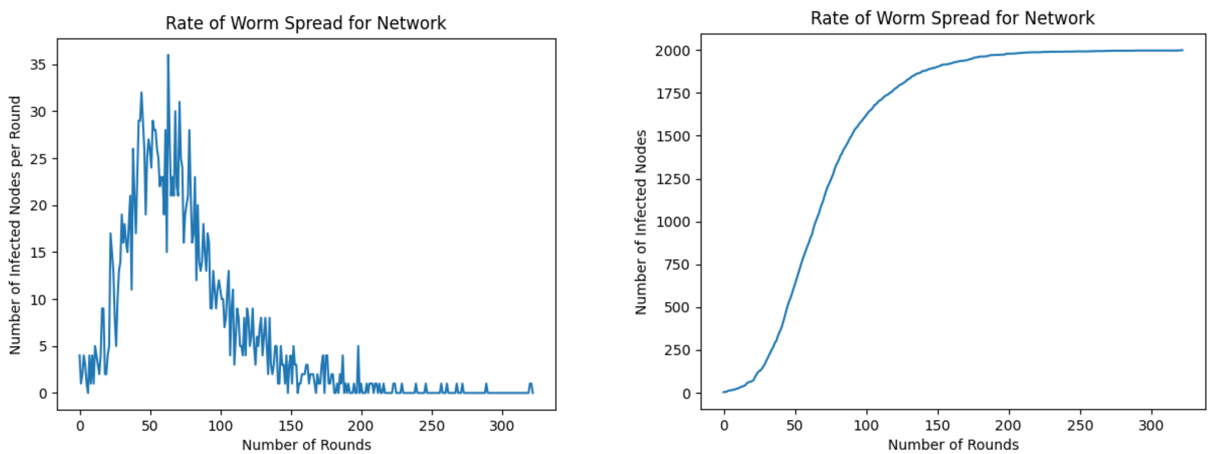


Erdos-Renyi: Average Number of rounds = 8.1

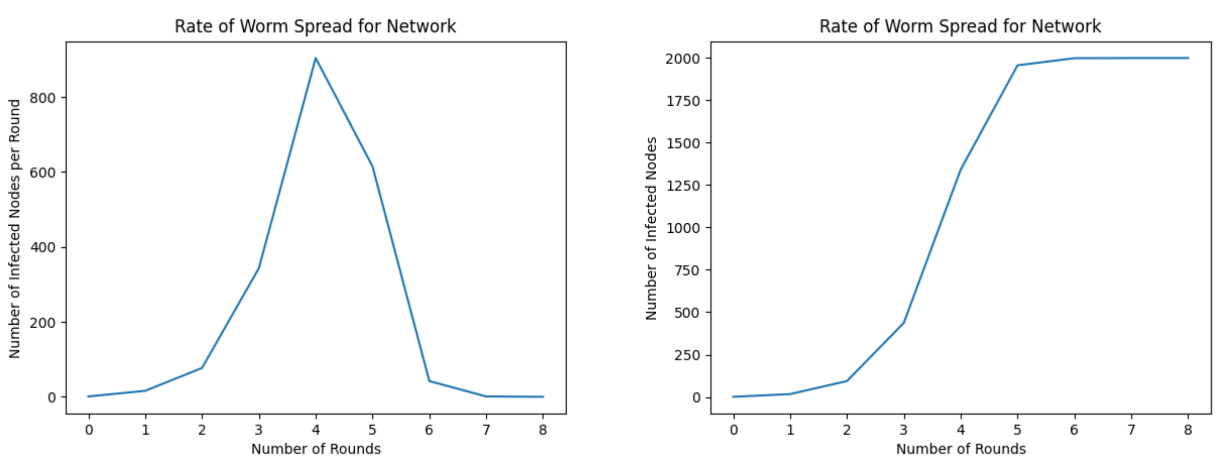


Watts-Strogatz: Average Number of rounds = 625.1

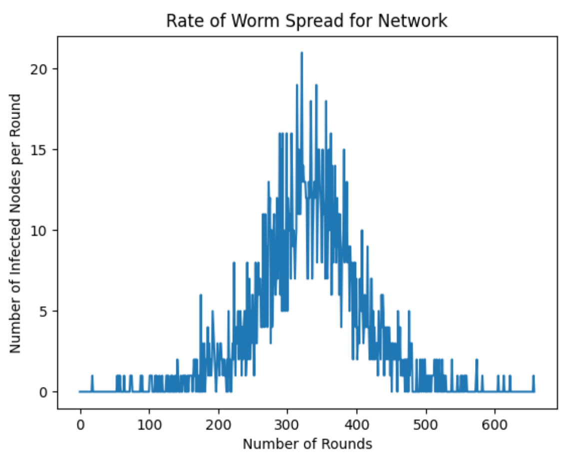
For 2000 nodes:

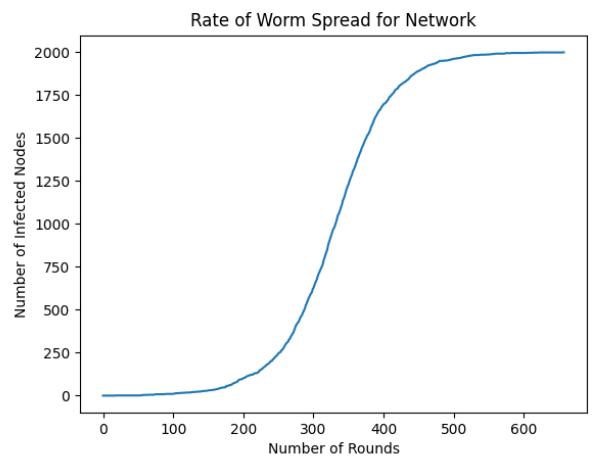


Barabasi-Albert: Average Number of rounds = 323.9



Erdos-Renyi: Average Number of rounds = 9.2





Watts-Strogatz: Average Number of rounds = 658.0

2.a) Shape of the graphs

The above graphs validate our first question, whether the shape of the graphs is S-shaped when a cumulation of nodes is considered, and bell-shaped when infected nodes at a given time is considered. Our graphs are different versions of bell-shapes and S-shapes.

For the first case, the graphs are bell-shaped because, when the worm starts propagating, gradually the number of nodes infected at a given time increases, until a certain point, when number of uninfected nodes starts reducing, as a result of which the number of nodes infected per time reduces as well. For example, in the case of Erdos-Renyi, for 1000 nodes, the worm propagation escalates at time 4, reaches its peak by 6, and gradually decreases till 9, after which we barely see any propagation.

For the second set of graphs, we are considering a collective of the number of nodes infected at a given time, it produces an S-shape, because at first, the propagation is slow, then we see a rapid rise in the total number of nodes infected, but soon the curve flattens, and with a smaller number of uninfected nodes remaining the propagation slows down. For example, in the case of Watts-Strogatz, for 2000 nodes, we see a slow propagation, until 250, when the total number of infected nodes increases rapidly, till time 400, when the curve flattens, creating an S-shaped graph.

2.b) Rate of propagation for different networks without cure

From the above graphs, we can spot significant differences for different network types, for the rate of propagation of the worm. For 1000 nodes, Erdos-Renyi is totally infected by an average of 8 rounds, whereas Barabasi-Albert is taken over in an average of 200 rounds. Watts-Strogatz is infected after 400 average rounds. This means that Watts-Strogatz shows the highest resilience, Barabasi-Albert is mediocre, and Erdos-Renyi shows the least resilience. For 1500 and 2000 nodes, similar characteristics are observed. Hence, we can conclude that the order of resilience for each network type against a worm is: Watts-Strogatz > Barabasi-Albert > Erdos-Renyi

2.c) Difference in the graphs

The S-shaped graphs look almost similar for each of the networks and the number of nodes, with one very significant difference due to the fact stated in the above paragraph. The propagation is much slower for Watts-Strogatz, while Erdos-Renyi has the fastest propagation. Barabasi-Albert network shows a middle rate of propagation compared to the other two. Hence the main difference in the graph lies in the number of rounds, or time.

3. Analogy with observations during the current pandemic.

A very similar situation was encountered by all human beings during the current pandemic. The network in this case was socializing with other human beings, the actual networking (before the internet), the worm was the virus SARS-CoV-2, and the disease that spread due to this was COVID-19. The virus propagated from one person to another via respiratory droplets. The virus was first found in Wuhan, China in December 2019, and patient 0 was recently identified there, who first got infected with the virus. This person was our first node in the propagation of the worm in the network. This person infected all with whom they socialized. These infected people propagated the virus, behaving very much like the nodes in our network.

The propagation of this virus happened at such a high rate that Wuhan first went under lockdown to prevent the spread, and gradually by end of March 2020, almost every country in the world went under lockdown. The symptoms of this disease were very severe, with dry cough, breathing issues, loss of smell and taste, even acute pulmonary problems in serious cases. Many people become victims of this disease and suffered severely during the first wave. To this day, we have a record of more than 435,000,000 falling victims of this disease, and almost 6,000,000 people losing their lives to this disease.

The prevention of this disease was not as easy as well. Social distancing, always wearing a mask, avoiding touch, sanitizing hand (human being have a tendency of touching their face with a hand probably 16 times an hour), etc., which if followed by common people, can reduce the spread.

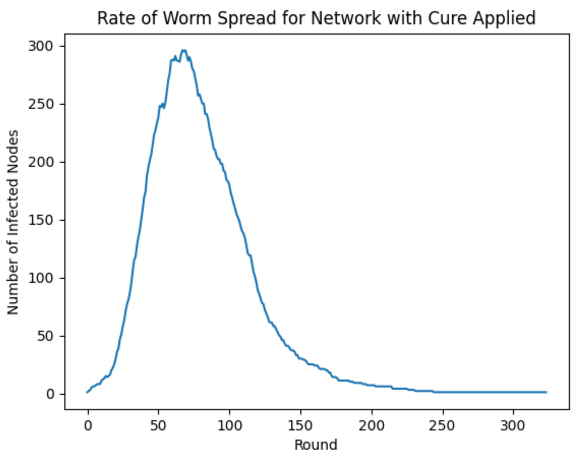
If one becomes infected, medical help (either cure at home, or at a hospital) is required. Much like the cure that propagates along with the worm in the next segment of our experimentation. This is like the inoculation that the cure performs in the network. Then comes vaccination, which immunizes (if not immunize, vaccination reduces the symptoms majorly), which is very much how the cure immunizes each node in our network.

4. Effect on worm propagation when cure is applied.

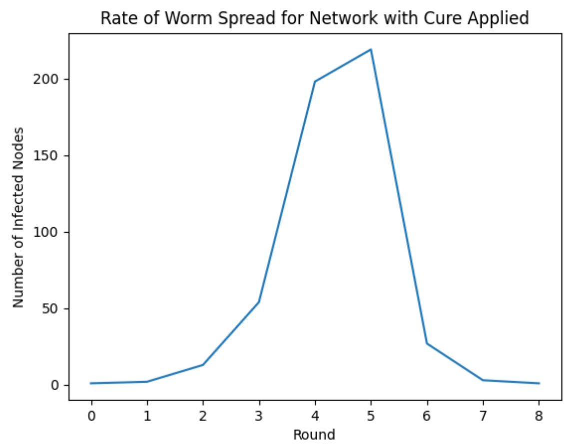
Now, we have tried to observe the result of propagation of worm, but when a cure is present. Like the behavior of the worm, we have a cure which propagates from one node to another, either inoculating them (if they are uninfected), or first inoculating and later immunizing them. The propagation of both the worm and the cure takes place simultaneously in the simulation, and we observe the results.

The plots below show the number of infected nodes with respect to time (rounds), when a cure is applied, for 2000 nodes, in all three network types. We set the probability value to 0.01 for the worm propagation, and 0.01 for the cure propagation. We observe the number of infected nodes increases at first, for all the network types. Eventually the cure catches up, thereby decreasing the number of infected nodes in the system, till the network is completely inoculated and cured.

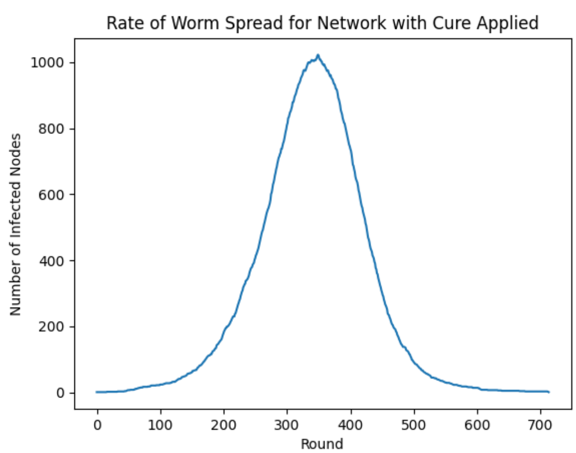
Let us consider the graphs for 2000 nodes:



Barabasi-Albert



Erdos-Renyi



Watts-Strogatz

In all the above cases for all three networks, we notice a similar bell-shaped graph. The difference from the previous case, where no cure is present is very clear. When a cure is applied, there are distinct changes. In all the networks, the cure seems to prevent the entire network from being taken over. For all the networks, the number of infected nodes reaches the highest and then decreases, ultimately inoculating and immunizing all nodes.

4.a) Rate of propagation for different networks with cure

For the cure to work, and to fully inoculate and immune all the nodes, each different network type shows a different rate of propagation for both the worm and the cure. As we see in the graphs, a similar type of average data is noticed for our 10 different observations. But for the worm to propagate to a maximum number of nodes for each network it is somewhat like this: To reach the maximum infected state, it took an average of 82.9 for the Barabasi-Albert, 5.5 for Erdos-Renyi, and 375.6 for Watts-Strogatz.

From the above observations, we can infer that, the resiliency among the three networks can be put as: Erdos-Renyi < Barabasi-Albert < Watts-Strogatz

On the other hand, to inoculate and immune the network it takes around an average of 7.9 rounds for Erdos-Renyi, 322.3 for Barabasi-Albert, and 709.5 for Watts-Strogatz. We cannot really comment on the rate of cure to take over the worm in each network from the available data, except for the fact that the rate of cure is somewhat comparable to the rate of infection for each individual network.

5. Resilience Enhancement Strategies

From the above experiments, we can easily conclude that the type of network plays a very significant role in propagation of worms. In both the experiments we observe this, first when the worm propagates in the networks without a cure (2.b) and then with a cure (4.a). We noticed in each case that the worm propagates much faster in Erdos-Renyi, the rate is slightly lower for Barabasi-Albert, but Watts-Strogatz shows maximum resilience. On studying the networks, we see Erdos-Renyi has a lower distance between paths and the number of nodes adjacent to another node is lower, whereas it is the complete opposite in the case of Watts-Strogatz. Watts-Strogatz network has a very high number of nodes adjacent to one another, and the distance between the nodes is also higher.

Resiliency in a network increase when the cure has a better chance of spreading, additionally, it takes more time for a worm to traverse in a network where the nodes are placed at a higher distance. Furthermore, we also noticed something about the application of the cure. If the cure is applied to a node on the periphery in the network, it takes more time for the cure to spread, whereas, when the cure is applied to a hub node (with many adjacent nodes), the cure spreads faster.

As per the above observations, two strategies can be suggested:

- A network must comprise of higher number of adjacent nodes
- A network must comprise of nodes placed further apart
- The cure should be introduced to a hub node with higher connections and adjacent nodes.

6. *References*

https://networkx.org/documentation/networkx-1.10/reference/generated/networkx.classes.function.get_node_attributes.html

https://networkx.org/documentation/stable/reference/generated/networkx.classes.function.set_node_attributes.html

https://networkx.org/documentation/stable/reference/generated/networkx.generators.random_graphs.erdos_renyi_graph.html

https://networkx.org/documentation/networkx-1.9.1/reference/generated/networkx.generators.random_graphs.barabasi_albert_graph.html

https://networkx.org/documentation/networkx-1.9/reference/generated/networkx.generators.random_graphs.watts_strogatz_graph.html