Epidemic Control: Simulation of Delaying Epidemic through flight cancellation.

November 28, 2016

Pratyush Singh, Unmesh Kanchan, Ashwini Aras

Abstract:

Study of Spreading of disease through travel network has recently become one of the important research topic. Airplane network is most crucial among them as they can quickly propagate the disease around the world. In this simulation we have tested various strategies to slow down the propagation of epidemic. We have used open flight airplane network dataset to test the different strategies to slow down the propagation around the network. One of the main and effective strategy to do this is to shut down the big airport, these airports are giant HUB in the network but at the same place this strategy is very expensive for mankind. The other strategy would be edge removal, cancellation of critical flight. We have used the SEIR model to simulate the infection across the network of 6977 airports and 67663 flights. We have tested different flight cancellation strategies to limit the spread of infection. The number of infected airports after applying these cancellation strategies is used as a parameter to rank the strategies.

Introduction:

Airline network is crucial for humanity; it connects the people around the globe and opens the corridor for business and communication. Recently these network have been studied through network science in order to shed some light on the global problems like Epidemic. Previous studies have discovered that the structure of these network has played an important role to give us new insight about the spreading of epidemic. We in this simulation project have also used the model of Airline network and tested few algorithms to limit and delay the spreading of epidemic.

Most of these spreading and propagation models works on two phenomena i.e. modularity and scale free real-world network. Modularity means many connections within the module and very few connections outside of module. Scale free network have highly connected nodes called HUB. The Airline network which we have used are both modular and Scale free.

We have referred a paper "Critical paths in a metapopulation model of H1N1: Efficiently delaying influenza spreading through flight cancellation by Jose Marcelino, Marcus Kaiser "and we have implemented **edge based betweenness centrality, Jaccard similarity, Clustering coefficient, High degree (Hub Removal) and random edge removal strategies.**

Method:

We Have used Python, Networkx Libraries to analyze the network, other python packages are also used like Matplot Lib to analyze the Result. We have used the data set from openflight.org.

We have used following few centralities measure to find the appropriate routes for flights cancelation.

Edge Based Betweenness Centrality:

Recent studies on airline network has shown that cohesiveness of the airline network did not arise from high degree node but due to a lesser connected airport had a more central role. This can be identified by high edge betweenness centrality- which is a ratio of all pairs shortest path crossing each node. We tested the impact of topological changes caused by incremental simulation of edge removal on the airline network and found it effective as removal of such edges will quite likely increase the average number of steps needed for spreading.

Jaccard Coefficient:

Neighborhood similarity structure of two nodes is also an important measure to test because dissimilar nodes who do not share same set of neighbors are more likely to be a link between two different network structure that might represent a shortcut between remote regions. This can be identified by jaccad coefficient, lower the value of jaccard coefficient more likely it is the connection between different network.

Clustering Coefficient:

Finding the cluster in the graph is also a good measure it means the degree to which nodes in a graph tend to cluster together. The nodes with low clustering coefficient means they are loosely coupled within the network and have weak ties. Removal of such nodes from the cluster could reduce the propagation of infection. These nodes can be identified by computing clustering coefficient.

We have also implemented random edge removal and Hub removal strategies for comparison of these centralities and to test the efficacy of our algorithm.

Algorithm:

- **Load Network:**
  - First complete network through the networkx libraries is loaded. Then isolated nodes are remove from the network.

- **Compute edges for betweenness centrality:**
  - To compute the candidate edges based for edge betweenness centrality, first each edge is assigned a weight which is equal to normalized degree of connected nodes (edges connecting High degree nodes should have high weight because if infection reach at high degree node it can propagate the infection to many nodes). Then the edge based betweenness centrality is computed for each edge.

- **Compute edges connecting low clustering coefficient:**
  - To compute the candidate edges connecting low clustering coefficient nodes, clustering coefficient of each connected pair of nodes is computed and clustering weight is assigned to the edge which is a sum of clustering coefficient of the connected nodes.

- **Compute edges for jaccard coefficient:**
  - To compute the candidate edges for jaccard coefficient all the edges jaccard coefficient is calculated and stored in ascending order.

- **Compute HUB for Hub removal:**
  - To compute the hub nodes are stored in the descending order (higher degree to lower degree).

- **Find the starting nodes:**
  - Randomly ten high degree nodes are chosen per simulation as the starting node to spread the disease.

- **Main Loop: [for each strategy]**
  - Initially all nodes status is set as susceptible.
  - Starting nodes are set as infected.
  - One to hundred incremented by interval of five percent of edges are selected from edge removal pool to remove.
  - Iteration for 100 days.
    - Randomly status of successor of infected nodes are set as exposed based on the number of flight between the nodes.
    - Based on the probability(BETA) exposed are set to infected.
    - Based on the probability(GAMA) infected are set to recovered.
    - Vaccination is done after 10 days of infection.
    - On the day(iteration) of flight cancellation candidate edges are removed from the network.
    - Once the pair of nodes are recovered the connecting edge is restored.

- **Analyzer:**
  - The analyzer will read all the files and draw a combined line graph for each simulation.

- GraphML file is generated after each main loop.

- All the result will be stored in a folder in current directory named with current time.

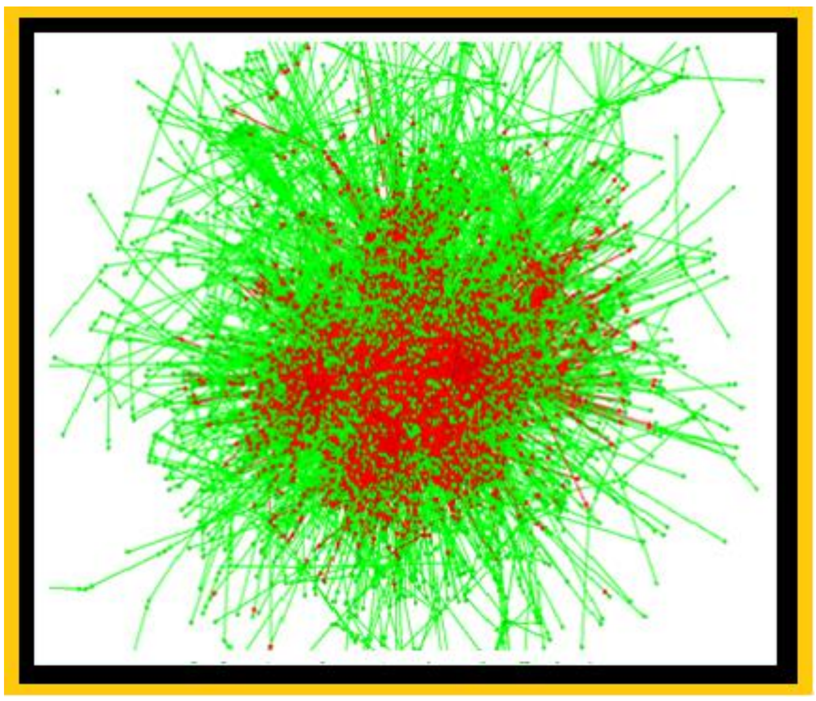
Result:

Starting with 10 highly connected randomly chosen nodes. Disease propagation can be seen along the red color nodes. In each strategy the majority of infected nodes are at center who are highly connected. The nodes on the periphery are less infected. It can be assumed that highly connected or big airports are valid target for disease propagation. Hence should be considered for cancellation.

Below images show the infected airports when only 25% of edges are removed for each strategy.

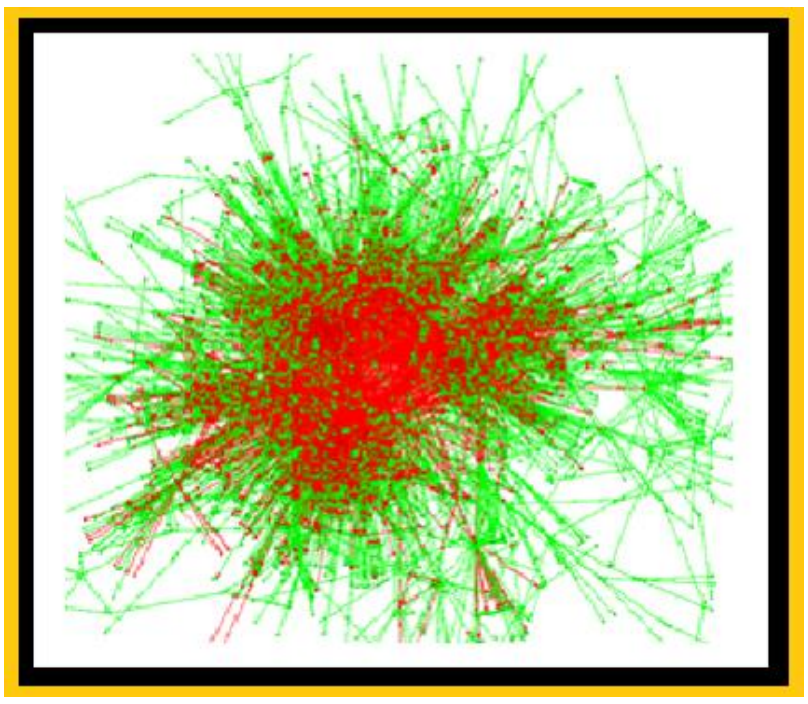
Betweenness Centrality:

In betweenness centrality strategy, the nodes in the center are more infected than the one on periphery, so the rate of infection is lower for those airports who are far from the central nodes. The population of red nodes are less dense than the one from clustering coefficient, when only 25% of edges are removed from each strategy edge removal pool.



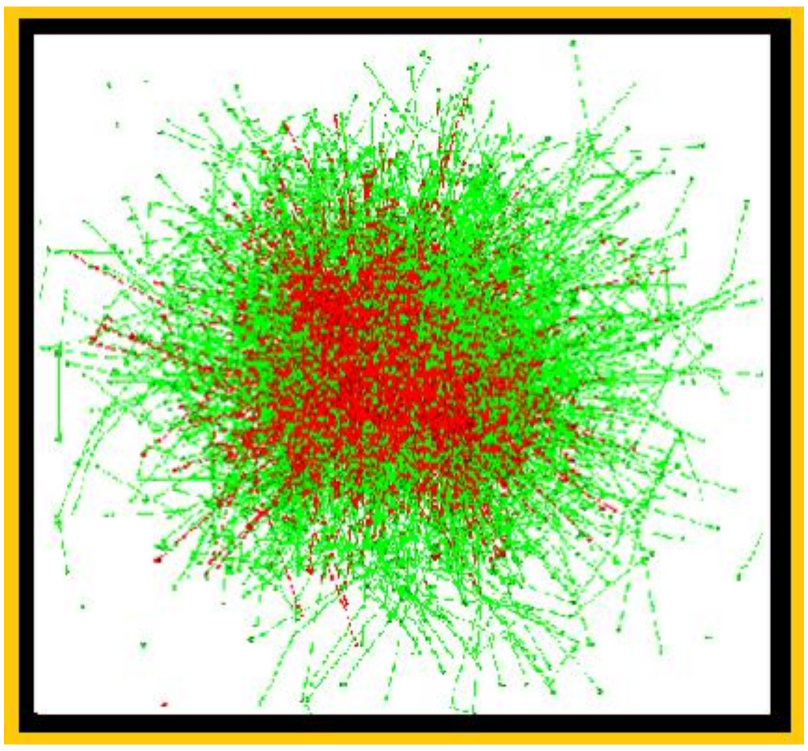
Clustering Coefficient:

In clustering coefficient strategy, the population of red nodes are denser than the one from betweenness strategies, however when more than 40 % edges are removed, the result is opposite when compared with betweenness strategy.



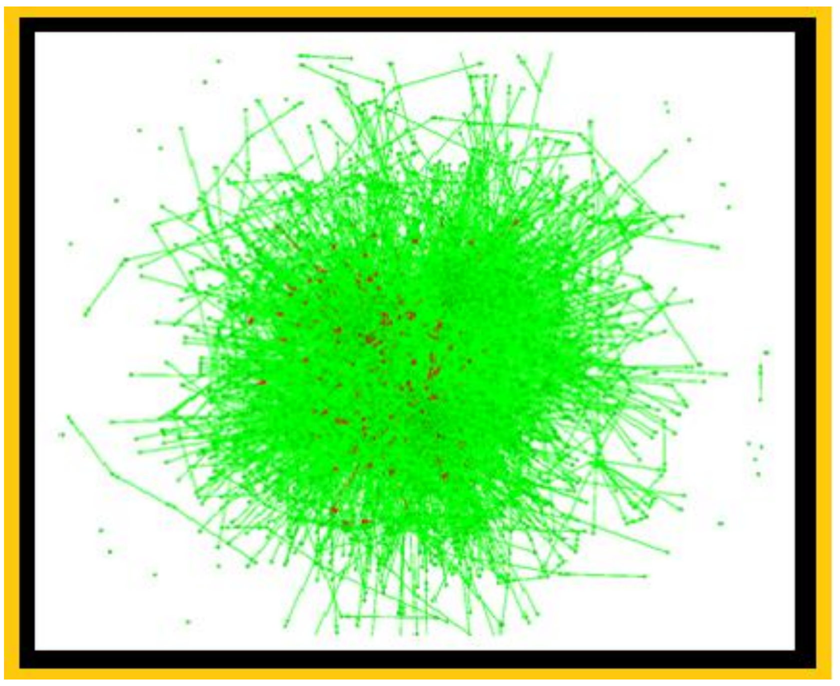
Jaccard Coefficient:

In jaccard Coefficient strategies, the count of infected nodes is not less than the other strategies but the infected nodes are more scattered even the infection reaches to few peripheral nodes.



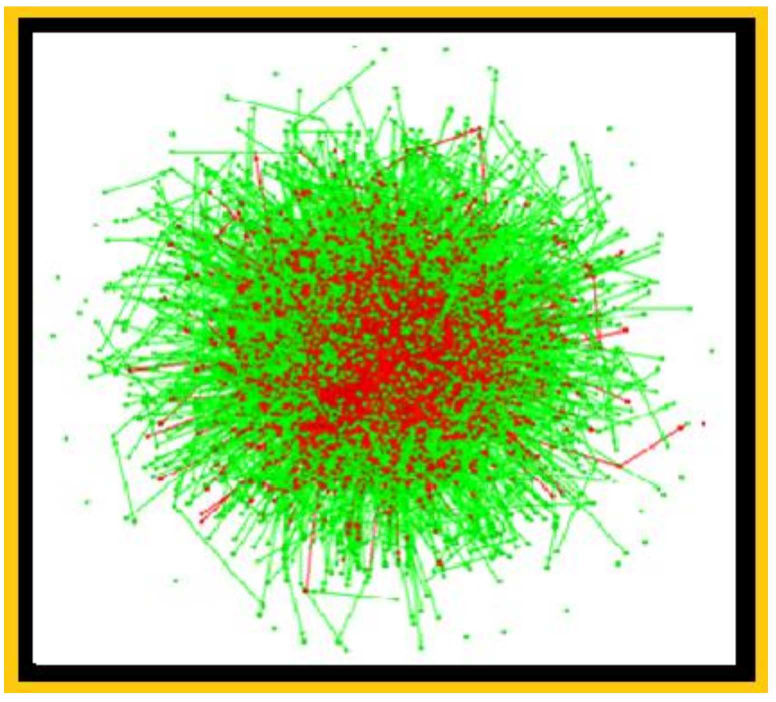
Hub Removal:

This is the best strategy where high degree nodes are removed and limiting the propagation to very few nodes but shutting down the major HUB (larger airport) is very expensive for humans.



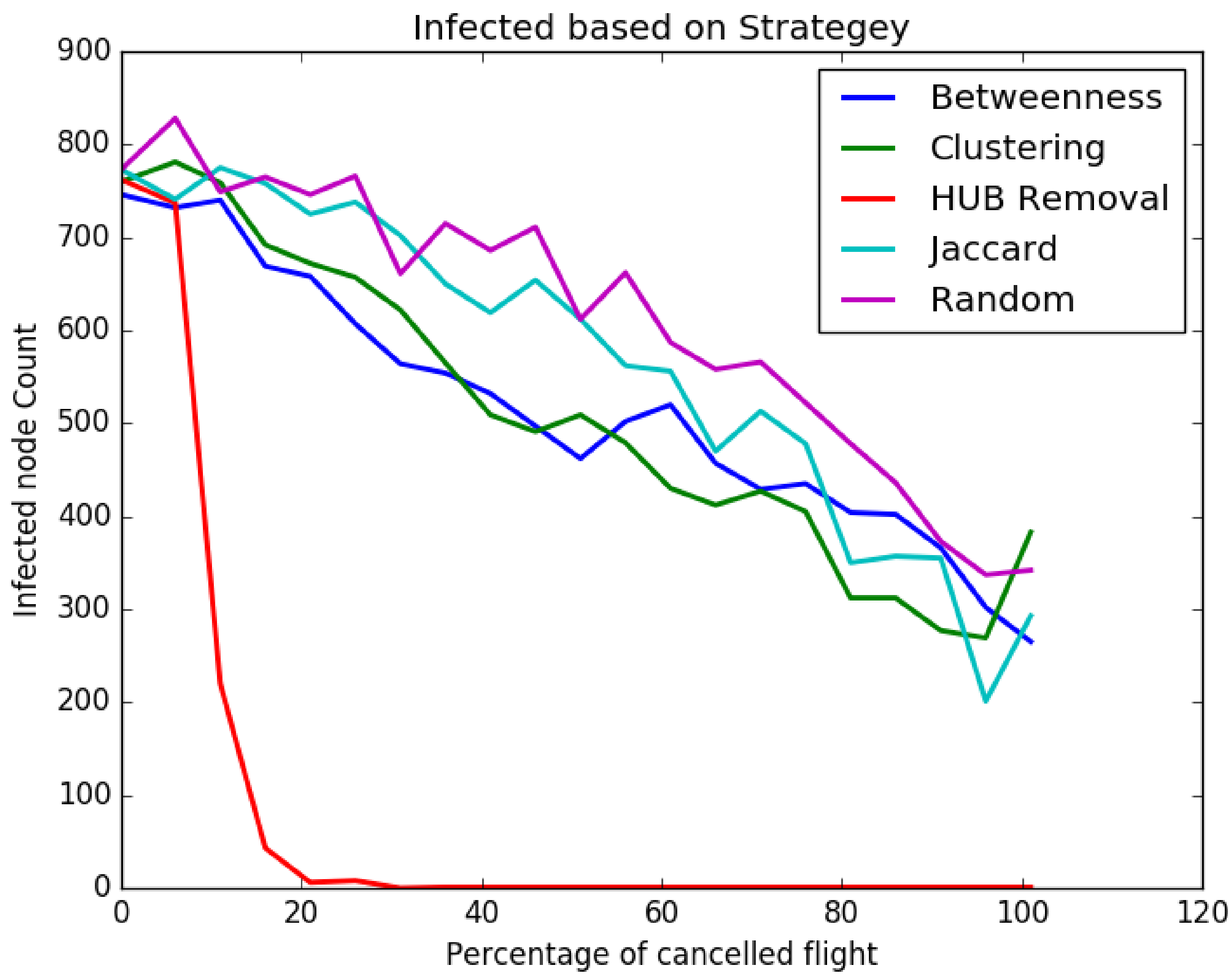
Random edge removal:

Random edge removal where any edge is randomly removed from the network here the count of infected nodes is also high and the infected nodes are also scattered. The infection reaches to peripheral nodes.



Comparison:

The below line graph clearly explains the test Hub Removal strategy result in least number of infected airports which can be predicted because since the hub itself is removed the propogation channel is closed and this limit the propogation of infection. With just 30% cancellation of flight betweenness comes out to be more efficient than the clustering coefficient and jaccard coefficient. With more than 40% cancellation of flight clustering coefficient comes out to be more efficient than jaccard coefficient and betweenness. The random edge removal strategies comes out to be more expensive as the count of infected nodes are high throughout the simulation also the infection is uniform throughout the network and reaches at the peripheral as well. Although in some simulation test we found random edge removal strategiey giving moderate result.



Conclusion:

The flight cancellation effectively limits the propagation of infection. The betweenness strategy is good with moderate number of cancellation while clustering coefficient is good when more than 40% of cancellation. We suggest that control group should chose the strategy intelligently after considering all parameters i.e. if the disease is severe and the number of days after first exposure is more than Hub removal can be good candidate algorithm. If the results are needed from the moderate number of flight cancellation betweenness is a good candidate measure.

References:

Marcelino J, Kaiser M. Critical paths in a metapopulation model of H1N1: Efficiently delaying influenza spreading through flight cancellation. PLOS Currents Influenza. 2012 Apr 23 . Edition 1. doi: 10.1371/4f8c9a2e1fca8

Chakrabarti, D., Wang, Y., Wang, C., Leskovec, J., and Faloutsos, C. 2008. Epidemic thresholds in real networks. ACM Trans. Inform. Syst. Secur. 10, 4, Article 13 (January 2008), 26 pages. DOI = 10.1145/1284680.1284681 http://doi.acm.org/ 10.1145/1284680.1284681