

ANALYSING THE SPREAD OF THE PANDEMIC OVER THE AIR TRANSPORT NETWORK

GROUP 28

Girish Sapkota	Ishraque Ahmed Mahi	Rai Muhammad Ibrahim Badar	Yat Syn Chan
Report Writing Algorithm Implementation 25%	Report Writing Algorithm Implementation 25%	Report Writing Algorithm Implementation 25%	Report Writing Algorithm Implementation 25%

Abstract

In this study we observed the spread of the pandemic over the air transport network using network theory. We used the dataset from openflights.org to create multiple airline networks where nodes represent airports and edges represent routes. We compared unweighted, weighted, and directed networks to determine which networks would closely depict the real world. The aim was to look at the importance of the airline network in the spread of a pandemic and observe the patterns of the spread related to the different characteristics of each node. We started our analyses from the node representing Wuhan, the origin of the pandemic and used the SIR model to simulate the spread of the pandemic. We tried different rates of infection and recovery to choose the ideal parameters that would closely shadow COVID 19. The SIR model was run 100 times, and the average time steps taken by each node were then compared with the degree distribution of that node and the distance of that node from Wuhan.

Introduction

Undoubtedly the Coronavirus disease is one of the biggest global catastrophes our generation has ever faced. 'Coronavirus (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus' (Who. int, 2022) ^[1], the very first case of COVID-19 was reported on Dec. 1, 2019 (John Hopkins Medicine, 2022) ^[2] and it wasn't until almost an entire year later Dec. 8, 2020, when the world saw the first vaccines being delivered (GOV.UK, 2022) ^[3]. This meant in the early days the prevention of the spread of the COVID-19 disease had to rely solely on non-pharmaceutical means. Experts found that the virus spreads mainly among people who are in close contact with each other (Who. int, 2022) ^[4], which consequently lead to governments announcing social distancing, mandatory face masks, lockdowns, and travel restriction measures to prevent the spread of the disease (Instituteofgovernment.org.uk, 2022) ^[5]. However, to date, there have been approximately 495 million cases and approximately 6 million deaths (Covid19.who.int. 2022) ^[6].

This leads us to question how effective the prevention measures were? Could we have stopped the pandemic at its root? Could we have avoided the crumbling of highly significant institutes? Could we have saved countless lives?

Contrary to popular opinion COVID-19 is not a 'once in a lifetime' or a 'once in a century' pandemic, the frequency and severity of spillover infectious diseases from wildlife host to human is steadily increasing (Smitham et al., 2022) ^[7], which makes another pandemic very likely. Analysing the spread of the COVID-19 virus could help us understand its spread and in turn help us generate more effective measures to control a future pandemic and ensure the same costly mistakes are not repeated.

Coronavirus was initially detected in China, in the city of Wuhan (Euro.who.int, 2022) ^[8] and was then transferred to the rest of the world. It was found that about 30,000 passengers fly from Wuhan daily to destinations throughout the world (Wilson et al., 2020) ^[9], furthermore in 2019 4.5 billion passengers were reported to have used the air transport network (Icao. int, 2022). This highlights the importance of the air transport network for the spread of disease and could also be a huge factor in determining disease control. Hence, we will focus on the spread of the pandemic over the air transport network. We will be using data from openflights.com and treating each airport as a node, all flights between airports will form an edge. We will be analysing the network's degree distribution and eigenvector centrality to compare weighted, unweighted and directed networks and find which network closely depicts the actual world airline network. The SIR model will be used to simulate the spread of the disease. Multiple infection and recovery rates will be tested to find the most ideal parameters for closely simulating COVID 19. The SIR model simulation will be run multiple times and an average will be calculated. We will compare the characteristics of each node, such as the distance of the country from Wuhan and the degree distribution against the time steps taken by the node to get infected in the simulations. Finally, we will generate a world map showing the spread of the pandemic i.e., the sequence in which countries got affected. These observations could then be used to carry out further work, help determine the most vulnerable countries and help identify any patterns between characteristics of each and how they influence the spread of the pandemic. This could in turn help in identifying the major factors in the spread of the pandemic and help identify the weaknesses, which could then be used to control a potential pandemic in the future.

Followed is the conducted literature review, its findings, the dataset and network presentation, the terminology used, the network analysis methodology, the results observed and finally the conclusions.

Related Work

'Revealing the structure of the world airline network' (Verma, T., et al, 2014) ^[12] is an article that uses standard and non-conventional removal measures to identify the different classes of clusters in the world airline network based on the physical proximity of airports and analyses the consequences of this fragmentation. The article observed and concluded that even though the network is highly connected and has direct flights, the scanty number of passengers flying from remote areas means that these flights aren't economical for the airlines and hence they instead use peripheral routes which results in star like network structures with peripheral hubs in the centre. Studying the spread of pandemic over the airline network could help identify these peripheral hubs and removing them could

significantly disrupt the network, hence, dramatically mitigating the spread of the pandemic.

Duan, W., et al, conducted a comprehensive review of epidemic models to provide insight into the literature of epidemic modelling and simulation (Duan, W., et al, 2015) ^[13]. They looked at three different kinds of models: mathematical, complex network and agent based. They ranked each model based on heterogeneity, implementation complexity, detail of results and amount of computing power required. They concluded mathematical models require low computing power and are simple, however they sacrifice heterogeneity and details. Agent Based Models on the other hand are much more complex and require high computing power but provide richer heterogeneity and detail. Complex models were found to be located in between and were found to be the most balanced for epidemic modelling. This justifies the use of the SIR model in our study, the SIR model is a differential equation model and falls in between mathematical and complex models which means that it doesn't sacrifice too much on details and heterogeneity even with the low processing power that is available to us.

A paper by Craig, B, et al studied the advantage of adopting a network perspective that explicitly accounts for the structure of interactions among individuals, focusing particularly on the class of SIR models (Craig, B, et al, 2020) ^[14]. The paper concluded that the SIR model commonly used makes assumptions about the pattern of interactions among individuals that are unlikely to hold true in the real world. The airline network used in this study tries to accommodate this. Moreover, we will also be studying additional metrics for each node such as their degree distribution and their distance from Wuhan in terms of flights to judge how realistic the results from the SIR model are. Reflecting on the article 'Revealing the structure of the world airline network,' it was observed that the airline network was highly connected so for the case of this study the assumptions that the SIR model makes in terms of the interaction among individuals may not be too unrealistic as almost every airport connects with every other airport in the airline network.

The research carried out in the article by Panayotis Christidis and Aris Christodoulou discusses the importance of understanding the origin point of the spread of the Covid-19 pandemic (Christidis, P et al., 2020) ^[15]. They discuss how Wuhan, the city of the outbreak, is a global aviation hub in central China. The article states, "Almost 1 million passengers flew out of Wuhan per month" and those outbound flights had an estimated 23,532 destination connections around the world. The aviation industry is extremely volatile as flight volumes drastically fluctuate during seasonal variations, especially during the months of January and February when flights peak due to the Chinese New Year. Additionally, the research paper summarizes that the modern air network system can connect any two points around the world in less than 48hrs, hence exacerbating the spread of any viral infections at an unprecedented rate. This is a clear indication of the importance of the air transport network in the spread of the pandemic over the world, it suggests that the airline network was one of the very first points of entry for the virus to spread, one of the most rapid ones and hence among the most significant ones.

Dataset and network presentation

The dataset used for this analysis has been gathered from openflights.org. Two files were used, 'airports-extended.txt' and 'routes.txt.' 'airports-extended.txt' contains details about all the airports that will form nodes in the network. Information like the country of the airport, latitude and longitude will help us shape the network in the form of a world map. 'routes.txt' contains the routes of different airlines between airports, these will form the edges of our network, the weight of the edges represents the flight routes that are flown by more than one airline.

For the first task we used Gephi to create the network, hence we had to reformat the dataset, so it was Gephi-compatible. First, we imported the files into Excel using comma delimiters to distinguish different pieces of data. For the routes file, we deleted all inessential data and gave only two columns' headings of source and target, as is required by Gephi. For the airports file, once again, we deleted all redundant data and only kept the id, name, country, latitude, and longitude and gave them the corresponding headers. The new files were saved as edges.xlsx and nodes.xlsx. When we imported edges.xlsx, the network showed 3425 nodes and 19256 edges, upon importing nodes.xlsx the number of nodes increased to 6974, which means that 3549 nodes were completely disconnected. This is not representative of the real world, and we deduced that this data was missing. We decided to only use the nodes with edges for our analyses, so the nodes that are in the edges.xlsx file.

For all the other tasks we had to use the SIR model. Gephi does not have the means to run the SIR model and hence we had to use Python. We used the NetworkX package to create our network. We altered the xlsx files by removing the headers and then exported them as csv files. First, we loaded the edge.csv file into our network to create the nodes and edges. Then for each node in the nodes.csv file, we checked if the node existed in the network already and if it did, we updated the network with its details such as the country it was located in, its latitude and longitude. Loading in the dataset this way ensured that we didn't include airports that were fully disconnected for the same reason as described above.

Tables 1 to 3 show the basic statistics of the networks generated using Gephi. Note that the network in Python shares the same statistics.

Terminology

- **Nodes:** Each node represents an airport in the air transport network.
- **Edges:** Edges represent the routes between airports. They are represented with lines connecting the nodes. If multiple airlines fly the same route, the weight of the edge increases.
- **Diameter:** The diameter represents the shortest number of flights required to travel between the 2 most distant nodes/airports in the network.
- **Average Path Length:** Average path length represents the average number of flights along the shortest routes for all possible pairs of nodes/airports.

- **Average Clustering Coefficient:** The clustering Coefficient measures the tendency of different nodes/airports in the network to cluster/group together. The average clustering coefficient is the average of all the clustering coefficients of all the airports/nodes in the network.



(Figure 1. World Map layout of our network)

Nodes	3,425
Edges	19,256
Diameter	13
Average Path	4.103
Average Clustering Coefficient	0.628

(Table 1. Undirected & Unweighted)

Nodes	3,425
Edges	37,595
Diameter	14
Average Path	4.146
Average Clustering Coefficient	0.467

(Table 2. Directed & Unweighted)

Nodes	3,425
Edges	37,595
Diameter	14
Average Path	4.146
Average Clustering Coefficient	0.467

(Table 3. Directed & Weighted)

Network Analysis Methodology

The goal of this study was to analyse the spread of the pandemic over the air transport network. Before we could start the analyses, we needed to decide which type of network would best suit the problem we are analysing, which meant deciding whether we should use a weighted or a directed network. To help us decide on the most suitable network we created 3 different networks: undirected and unweighted, directed, and unweighted, and directed and weighted. We continued by comparing the degree distribution and the eigenvector centrality of each of the networks.

The degree distribution of a network represents the number of connections each node has. For a directed network that means calculating the sum of the in and the out degrees. In the context of this study, the degree of the network represents the number of routes each airport has flights for. If an airport has a high degree i.e., a high number of routes you can fly, we would be able to deduce that this airport could potentially play a larger role in the spread of the disease, as more routes mean more people flying in and out of this airport. A higher degree would also suggest that a pandemic would be able to spread wider through this airport as more nodes can be reached through it.

The eigenvector centrality builds up on this idea. The significance value known as the eigenvector centrality of each node not only depends on the number of connections the node has but also on the importance of each connection. A connection to a node that itself has a high number of connections will increase the eigenvector centrality more than a connection to a node that does not have any other connections. To calculate the eigenvector centrality each node is assigned a score of 1. The scores are then repeatedly; computed for each node as a weighted sum of centralities of all the nodes in its neighbourhood and normalized by dividing each score by the largest score until the centrality scores stop changing. In the context of our problem, considering a single airport if it has a route to another airport that is itself mainstream and has loads of routes itself, it will be more significant in spreading the pandemic than a route to an airport that itself has no other routes or is otherwise isolated.

These statistics could clearly show the most significant airports in the spread of a pandemic and if all the different networks show the same degree and eigenvector centrality distributions then either of them could be used. Else we will have to look at how the distribution of these statistics reflects the real world, i.e., which network represents the airports that are actually very busy and hence represent the problem as precisely as possible, and then choose that network.

To continue with our analysis, we used the SIR model over our network. The SIR network distributes all nodes into 3 groups. The susceptible group contains all nodes that can be infected, the infected group contains all nodes that have been infected and the recovered group contains all nodes that have recovered or died. Typically, most nodes start as susceptible with only the source of the disease or the first case of the disease starting as infected. The infected node can infect any susceptible nodes that it has an edge with, any susceptible nodes that are infected

can in turn then affect other susceptible nodes that have a connection/edge with them. Each node that has an edge with an infected node can get infected with an infection rate β . Each infected node can recover with a recovery rate γ . The higher the value of β or γ is, the higher the rate of infection or recovery is, respectively. The SIR model derives the population of susceptible, infected, and recovered nodes as a function of time, where time is represented by time steps.

To run the SIR model, we used a function from the EoN module. The EoN module is a Python package for the simulation of epidemics on networks. It provides a function for SIR called the `fast_SIR`. The `fast_SIR` function takes in a networkx graph, the values for rates β and γ , and the name of the initially infected node. The function then returns the arrays containing the values of populations S, I, R over time. The figure below shows the differential equation of the `fast_SIR` model.

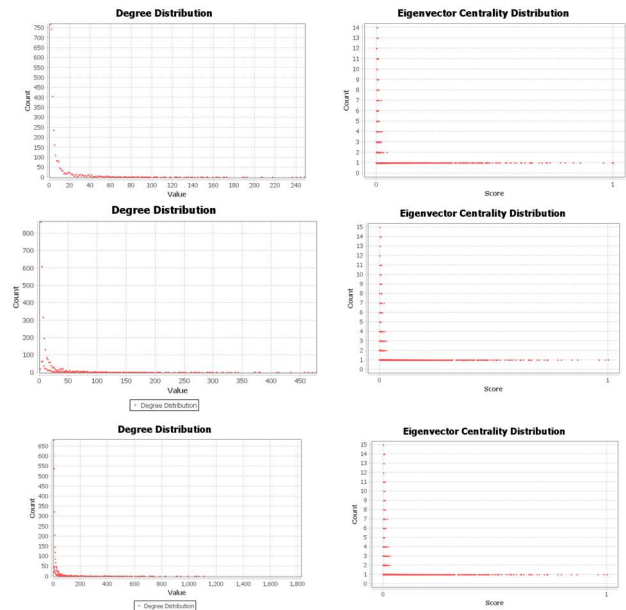
$$\mathbb{P}(X_n = k) = q \binom{n}{k} \sum_{i=0}^k \binom{k}{i} (-1)^i \frac{(1-p)^{n-k+i}}{1 - (1-q)(1-p)^{n-k+i}}.$$

For this study as we are studying the spread of Covid 19, we started with the airport corresponding to Wuhan as being initially infected and all other nodes being susceptible. We ran the SIR model for different values of β and γ to assess the dynamics of each case and how they affect the population of S, I and R over time.

Initially, with Covid 19 the rate of recovery was very low, and the rate of infection was very high, so we continued with a β value of 1 and a γ value of 0. We ran the SIR model again, but this time used the `Simulation_Investigation` object returned by the `fast_SIR` function to find the time when each node was infected. We ran the SIR model simulation 100 times and calculated the average time taken by each node to get infected. Then we used functions from networkx to calculate the degree of each node and plot this as a graph against the time it took for each node to be infected. We also calculated the distance of each node from the node corresponding to Wuhan and plotted a graph for that against the time it took for each node to get infected. Finally, we created a world map where the colour of each country represents the relative time it took for it to get infected, this was achieved with pandas, geopandas and folium. Pandas creates a data frame, consisting of all the countries and the time steps taken by them to get infected, we averaged the time steps taken to get infected for airports in the same country and assigned that value to the country. Geopandas provides the shapes and coordinates of each country which is merged on to the data frame, finally, folium is used to draw and colour code the map.

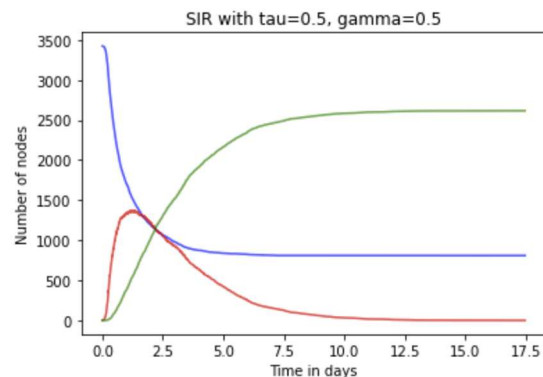
These results can help us judge if the number of routes of an airport and the distance of the airport from Wuhan in terms of the number of flights influences how fast the airport was infected. The world map could give an idea of the regions that would need more prompt restrictions to control the spread of the disease.

Results and discussion

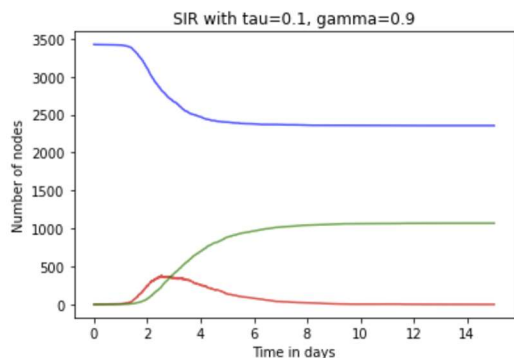


The figures above show the degree and the eigenvector distributions for all three networks collected from Gephi, the first row corresponds to the unweighted and undirected network, and the second row corresponds to the directed and unweighted network and finally the last row corresponds to the directed and weighted network. All networks follow the same pattern for both degree distribution and eigenvector centrality. Most nodes have a low degree distribution in all three networks, approximately $\frac{2}{3}$ of the nodes in the network have a degree of less than 20. The maximum degree differs for each network but that's since directed and weighted networks have more edges. The nodes with the highest degree for all graphs were very similar, airports in France, Dubai, Istanbul etc all corresponding to the busiest airports in the world (ACI World. 2022)^[11]. The eigenvector centrality showed similar patterns, most nodes in the graph had a low eigenvector centrality, approximately $\frac{1}{4}$ of the nodes in the network had an eigenvector centrality smaller than 0.1. The airport in France had the highest eigenvector centrality in all the networks.

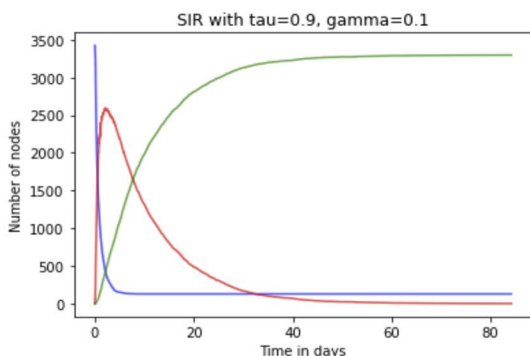
As all the networks shared the same patterns of degree and eigenvector centrality distribution, we decided to use the undirected and unweighted network as it is the easiest to implement and as it is compatible with the SIR function in the EoN library in python.



The figure above shows the result from the SIR model when the rate of recovery and the rate of infection were similar. The blue line represents the susceptible nodes, the red lines represent the infected nodes, and the green line represents the recovered nodes. The graph clearly shows a single wave of infection. The rate of infection at the start was very high. More than 1400 nodes were infected in less than 1.5 days. This is expected as the network had an average clustering coefficient of 0.628 i.e., each node was directly connected to more than half of the other nodes in the network allowing the disease to initially spread fast. Additionally, as the diameter was very small the two furthest nodes were only 13 flights away from each other. The rate of recovery increased gradually until all infected nodes were recovered. Most of the nodes recovered in 7.5 days. Approximately 1000 nodes remained susceptible and were never infected. This corresponds to how most of the nodes in the network had a low degree, which in turn means their chances of getting infected were low as they were not exposed to many infected nodes. This resulted in the disease dying out.



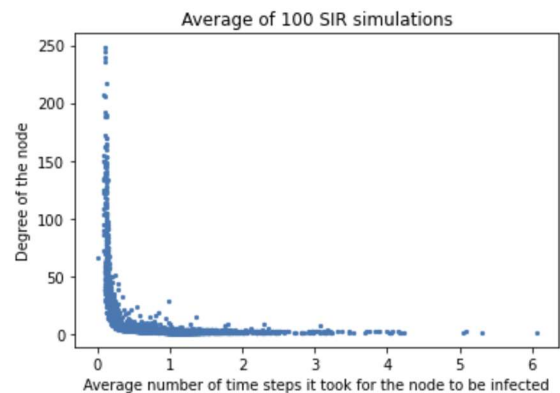
This figure shows the results from the SIR model when the rate of recovery was much higher than the rate of infection. Once again only one wave of infection was noticed but it was much less significant. At the start, the rate of infection was low. The airport from Wuhan had a degree of 66 so the spread was initially slow. Until approximately day 2 when most likely an airport with a very high degree and eigenvector centrality was infected. This increased the spread of the disease and the rate of infection peaked. However, due to the very high recovery rate, nodes likely recovered before they could infect other nodes and the infection died out and almost 70% of nodes were never infected i.e., left susceptible.



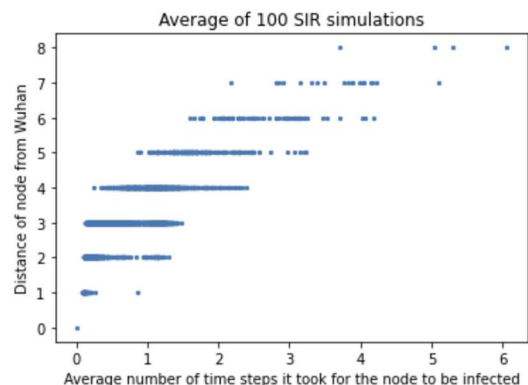
This figure shows the results from the SIR model when the rate of infection was high, and the rate of recovery was very low. The rate of infection was very high from the very start, almost all nodes in the network had been exposed to the disease in less

than 5 days. The number of infected nodes also peaked during this time, with almost 70% of the nodes being infected. The rate of infection declined exponentially after this. The rate of recovery was initially high when the number of infections was high and understandably plateaued when the number of infected decreased.

This clearly shows how important the rate of recovery and infection are. Hence each disease would have a different model of spread based on how it influences these parameters. For diseases where the rate of infection is like the rate of recovery, we can expect the infection to be high at the start, but recovery will follow at a little lower rate. We can expect a lot of the population to be infected. Comparatively, diseases where the recovery is high, and the rate of infection is low. We can expect more time before the disease starts to spread, the spread of the disease will not be significant as nodes will likely recover before they can infect others. Finally, with a disease that has a much higher rate of infection compared to recovery, we can expect almost all the nodes to get infected rapidly, with recovery being much slower comparatively.

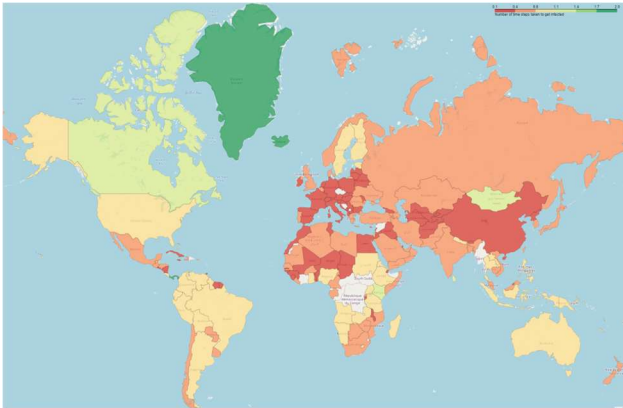


The figure above shows the average time steps it took for the node to be infected relative to the degree of the node. We can clearly observe that the nodes with a higher degree were infected faster than nodes with a lower degree. We can see some nodes that have a low degree but were still infected quickly, these could be nodes that have direct paths to Wuhan or have very small paths to Wuhan. The nodes with higher degrees being infected rapidly could also have caused this as the disease can spread to even more nodes through them at very early stages.



This figure shows the average time steps it took for each node to be infected over 100 SIR simulations relative to the distance of the node from Wuhan. A clear pattern can be observed. The

smaller the distance a node had from Wuhan the quicker it was infected. Some nodes with similar distances were infected later than other nodes, these nodes were likely not infected directly from the shortest path from Wuhan but from another neighbouring node that got infected later.



The figure above shows a world map which has been colour coded so that each country represents the average number of time steps taken for them to be infected over 100 SIR model simulations. The colour red represents countries that were infected the quickest and green represents the countries that were infected last. The world map shows the countries to get infected the earliest were China, Turkmenistan, Niger, Spain, France, Germany etc. It is understandable that China was infected rapidly considering that it was the initial source of infection. Countries like France and Germany had a high degree distribution and a high eigenvector centrality and hence, they were infected rapidly. This corresponds to how a high eigenvector centrality and a high degree distribution means that more people travel through here, i.e., more carriers of the disease and more chance for it to get infected. Overall the entire world was infected within two time steps; this corresponds to the research carried out by Panayotis Christidis and Aris Christodoulou (Christidis, P et al., 2020)^[15], they summarized that the modern air network system can connect any two points around the world in less than 48hrs, additionally, the paper 'Revealing the structure of the world airline network' (Verma, T., et al, 2014)^[12], also observed how the network was highly connected, the paper also discussed how a star-like network was also observed with the centre corresponding to a peripheral hub, this reinforces why France and Germany were affected so rapidly; Germany is home to Frankfurt, the world's leading financial institution and France is one of the most attractive tourist destinations in Europe with its iconic landmarks, architecture and food. They are the two most internationally connected countries in Europe to the rest of the world. The rest of the world shows a clear pattern with more industrial countries getting affected faster than others as they get more air passengers. Even though America has a higher degree and eigenvector centrality, it was still affected relatively late. This corresponds to how America has the most local airports which cluster together and hence increase each other's degree and eigenvector centrality but are still relatively less connected to the international world.

Conclusion

All the proposed analysis was completed successfully. A clear pattern can be observed in the results. Nodes with higher degrees and eigenvector centralities were infected faster. Additionally, nodes closer to Wuhan in terms of numbers of flights were also infected faster than the nodes further away from Wuhan. The entire network was infected in as little as two-time steps which suggests how well connected the network is. Without a doubt, it plays a major role in the spread of disease.

However, more work will be required to identify nodes in the network that could be removed to control the spread of the pandemic. We would need to test the resilience of the network, node removal techniques could be employed such as the ones in the article 'Revealing the structure of the world airline network' (Verma, T., et al, 2014)^[12].

Furthermore, the study faced countless challenges. We found over half the airports in the dataset were disconnected, which isn't representative of the real world. The dataset contains routes between airports but has no values to represent the frequency of flights or passengers that fly these routes, these could have indicated the weights in our network and routes with more flights would have contributed more towards the spread of the pandemic as more disease carriers would have been exported from the origin. The dataset has no timeline for the flights that were accomplished and were not restricted when the virus was discovered, so the simulation will not fully represent what happened. Furthermore, the SIR model used introduces even more challenges. SIR follows a uniform contagion probability, however different environments, climates, and variants could all contribute to different contagion probabilities. Additionally, SIR expects that all nodes will encounter each other with the same probability which does not hold for non-Erdos-Renyi random networks. Moreover, SIR assumes that a node can only infect once and once a node is recovered it cannot be infected again, however, these were not the cases with COVID 19, and it is highly likely these cases will not be true for the future pandemics either. Using a better dataset and an agent-based model as described in the report by Duan, W., et al could address these challenges and provide much more realistic and detailed results which will be more dependable (Duan, W., et al, 2015)^[13].

References

1. Who.int. 2022. *Coronavirus*. [online] Available at: <https://www.who.int/health-topics/coronavirus#tab=tab_1>.
2. John Hopkins Medicine, 2022. *What Is Coronavirus?* [online] Available at: <<https://www.hopkinsmedicine.org/health/conditions-and-diseases/coronavirus>>.
3. GOV.UK. 2022. *UK marks one year since deploying world's first COVID-19 vaccine*. [online] Available at: <<https://www.gov.uk/government/news/uk-marks-one-year-since-deploying-worlds-first-covid-19-vaccine>>
4. Who.int. 2022. Coronavirus disease (COVID-19): How is it transmitted?. [online] Available at: <<https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-covid-19-how-is-it-transmitted>>
5. Instituteforgovernment.org.uk. 2022. StackPath. [online] Available at: <<https://www.instituteforgovernment.org.uk/charts/uk-government-coronavirus-lockdowns>>
6. Covid19.who.int. 2022. WHO Coronavirus (COVID-19) Dashboard. [online] Available at: <<https://covid19.who.int/>>
7. Smitham, E. and Glassman, A., 2022. The Next Pandemic Could Come Soon and Be Deadlier. [online] Center for Global Development | Ideas to Action. Available at: <<https://www.cgdev.org/blog/the-next-pandemic-could-come-soon-and-be-deadlier>> .
8. Euro.who.int. 2022. About the virus. [online] Available at: <<https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/novel-coronavirus-2019-ncov>>
9. Wilson, M. and Chen, L., 2020. Travellers give wings to novel coronavirus (2019-nCoV). *Journal of Travel Medicine*, 27(2).
10. Icao.int. 2022. The World of Air Transport in 2019. [online] Available at: <<https://www.icao.int/annual-report-2019/Pages/the-world-of-air-transport-in-2019.aspx>>
11. ACI World. 2022. The top 10 busiest airports in the world revealed - ACI World. [online] Available at: <https://aci.aero/2022/04/11/the-top-10-busiest-airports-in-the-world-revealed/> [Accessed 20 April 2022].
12. Verma, T., Araújo, N. and Herrmann, H., 2014. Revealing the structure of the world airline network. *Scientific Reports*, 4(1).
13. Duan, W., Fan, Z., Zhang, P., Guo, G. and Qiu, X., 2015. Mathematical and computational approaches to epidemic modeling: a comprehensive review. *Frontiers of Computer Science*, 9(5), pp.806-826.
14. Craig, B., Phelan, T., Siedlarek, J. and Steinberg, J., 2020. Improving Epidemic Modeling with Networks. *Economic Commentary (Federal Reserve Bank of Cleveland)*, pp.1-8. W
15. Christidis, P. and Christodoulou, A., 2020. The Predictive Capacity of Air Travel Patterns during the Global Spread of the COVID-19 Pandemic: Risk, Uncertainty and Randomness. *International Journal of Environmental Research and Public Health*, 17(10), p.3356.