# Network Science Project – Group 15

# Robustness and Cascading effects in Ryanair's Flight Routes

João Santos nº 102746
António Silva nº 102879
Tiago Basílio nº 103326

***Abstract* — Airline networks are complex systems where disruptions can lead to widespread failures, significantly affecting global connectivity. In this report, we analyzed the structural properties of Ryanair's flight routes to assess its robustness against node failures. Using numerous metrics, we explore the effects of attacks caused by node removals. This report provides insights into the vulnerability of airline networks and offers a framework for understanding how cascading failures propagate, highlighting the importance of resilience in network design and management.**

## I. Introduction

In this report, we studied a network consisting of Ryanair's Flight Routes [3]. It contains 176 nodes and 1242 edges. It is represented in the figure (fig. 1) below.

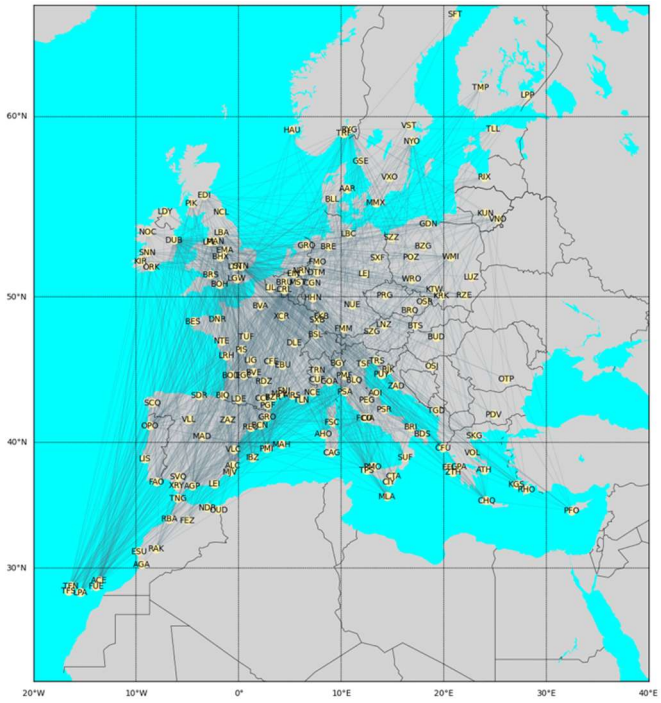


***Figure 1** – Ryanair's flight Routes*

In the first section of this report, we conducted a detailed analysis of the network's current structure. This involves calculating key metrics such as degree distribution, clustering coefficient and node centrality to uncover the most important airports, or hubs, in terms of connectivity. The centrality analysis informs our subsequent targeted attack, where we simulate the removal of the most influential airports to observe the cascading failures that follow.

In addition to the targeted attack, we also simulate a random attack, where nodes are removed without regard to their centrality or importance. By comparing the outcomes of these two types of attacks, we aim to uncover significant differences in how the network responds to planned versus random disruptions.

We aim to address several key questions regarding the robustness and vulnerability of the network:

- How well-connected is the network globally and regionally?
- What are the more important airports to the network?
- How does the removal of highly important airports impact the overall connectivity and structure of the network?
- What are the differences in network fragmentation between targeted and random attacks?
- To what extent can the network withstand cascading failures before a significant breakdown occurs?

## II. Methods

We implemented our project using Python3 coding language, with NetworkX library, that allowed us to generate graphs, along with other libraries that helped us do our research.

## III. Analysis of the Network

In this section of our report, we did an analysis of the network, focusing on numerous key metrics. These metrics help us assess the overall structure, connectivity, and potential vulnerabilities of the network, providing a deeper understanding of its robustness and susceptibility to cascading effects.

### *A. Number of Components*

A component of a network refers to a subset of nodes that are directly or indirectly connected to each other, meaning there is a path between any two nodes within the component.

Our network is one giant component, meaning that all airports are connected, forming a single, large, interconnected system, ensuring high levels of global connectivity, allowing passengers to travel between any two airports with relatively few layovers.

### B. Degree Distribution

The degree distribution of a graph is a crucial metric for characterizing a network because it reveals how connections are distributed among nodes, offering complementary insights into its structure and behavior. The degree of a node $i$, can be computed by:

$$k_i = \sum_j a_{ij}$$

where $a_{ij}$ corresponds to the adjacency matrix.

Knowing the degrees of every node in a network, it´s possible to compute its degree distribution:

$$P_k = \frac{N_k}{N} = \frac{1}{N}\sum_i \delta(k_i - k)$$

where $P_k$ is the probability of having a node of degree k.

To evaluate whether the degree distribution of the airport network follows a power-law or an exponential distribution, we performed a likelihood ratio test. The power-law exponent calculated was $\gamma = 3.81$, which suggests that the network doesn't perfectly align with traditional scale-free networks (which have $2 < \gamma \leqslant 3$).

We further tested the fit of the degree distribution using a likelihood ratio test, comparing the power-law and exponential models. The test returned a positive likelihood ratio (R = 1.76), indicating that the power-law model fits the data slightly better. However, the p-value (0.41) was not statistically significant, meaning we cannot confidently reject the exponential distribution as a possible explanation for the degree distribution.

Based on these results, the degree distribution of the airport network likely exhibits characteristics of both power-law and exponential-like behavior, which aligns with findings from other real-world complex networks.

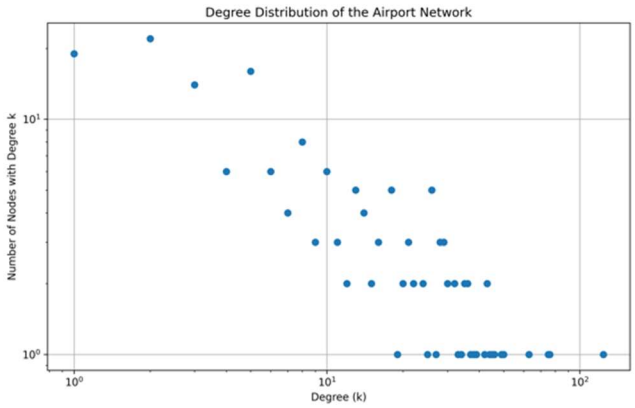


***Figure 2*** *– Degree distribution plot*

### C. Clustering Coefficient

The clustering coefficient is a measure of the degree to which nodes in a network tend to cluster together, quantifying how likely a node's neighbors are to be connected to each other, forming triangles. This is a very important measure to characterize the cohesiveness of a network, as it can reveal patterns of collaboration.

The clustering coefficient of a node $i$ can be computed by:

$$C_i = \frac{e_i}{(k_i(k_i - 1)) \div 2}$$

where $e_i$ corresponds to the number of edges among $i$'s neighbors.

Given the clustering coefficient of every node, it's possible to compute the network's clustering coefficient:

$$\langle C \rangle = \frac{1}{N}\sum_i C_i$$

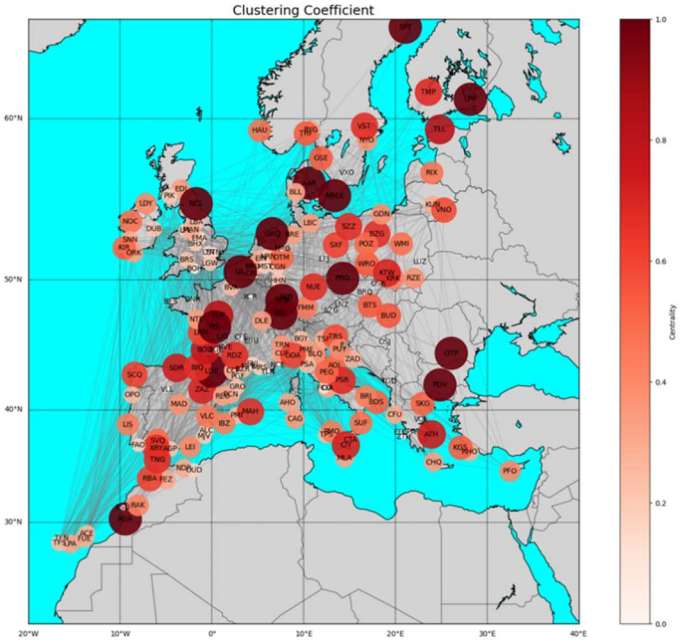


***Figure 3*** *- Clustering Coefficient of each Airport*

From this graph, we can see some airports like NCL (Newcastle Airport), PRG (Prague Airport) and PDV (Plovdiv Airport) present a very high clustering coefficient, forming a fully connected subgraph within their neighborhood, meaning that all airports connected to these highly clustered airports have direct flights to each other. These airports are likely local hubs in their regions, serving as major connectors within a small geographic area, playing an important role in regional traffic, helping distribute passengers between smaller destinations.

However, it is also possible to see the contrary, airports like LUZ (Lublin Airport), VLL (Valladolid Airport) and XCR (Paris-Vatry Airport) presenting very low clustering coefficients, meaning that they may not offer many direct flights between neighboring airports, relying on major hub airports for routing passengers to other destinations, leading to longer, multi-leg flights.

Our **network's clustering coefficient** is 0.3752.

### *D. Small-World effect*

The small-world effect refers to the phenomenon where most nodes in a large network can be reached from any other node through a small number of steps. In our network, we can verify that effect because:

- ACE (Lanzarote Airport) and SFT (Skellefteå Airport) are 4645 km apart; however, it only takes 2 flights to go from one airport to another.
- The diameter of the graph, that represents the maximum number of flights required to travel between two airports in the network, is 4.
- The average path length (APL) is the average over all shortest paths between all pairs of nodes in the graph, and can be computed by:

$$\langle L \rangle = \frac{1}{N(N-1)} \sum_{ik} L_{ik}$$

where $L_{ik}$ is the shortest path between node $i$ and $k$.

The **APL of our network** is 2.1696

The airport network exhibits the characteristics of a small-world network. The moderate clustering coefficient (0.3752) reflects relatively strong local connectivity among airports, while the short average path length (2.17) indicates that global connectivity is preserved, allowing efficient travel across the network. These results suggest that the network balances regional clustering and global reach, which are common features of small-world networks.

### *E. Node Centrality*

Node centrality is a measure of a node's importance in a network. It helps identify the most influential or central nodes based on their position and connections. There are several types of centralities:

**Degree centrality** is a measure of how many direct connections a node has in its network. It reflects the node's immediate influence, with higher degree centrality indicating that a node is more connected to others.

The degree centrality of a node $i$ is often normalized, and it can be computed by:

$$C_D(i) = \frac{1}{N-1} \sum_{j} a_{ij}$$

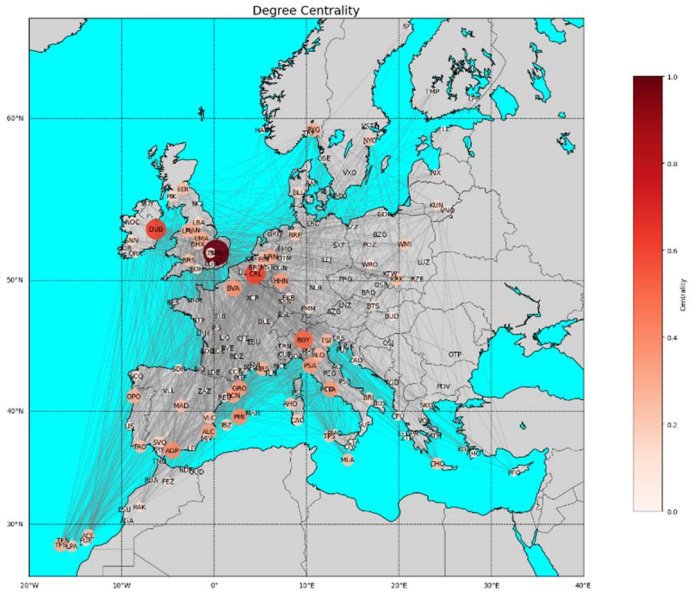


***Figure 4*** *- Degree Centrality of each Airport*

After computing the degree centrality of each airport, the airports with a higher degree centrality are:

- STN (London Stansted Airport) – 0.709
- DUB (Dublin Airport) – 0.434
- CRL (Charleroi Airport) - 0.429

These airports serve as major hubs, offering a wide range of direct routes to various locations. These hubs play a critical role in the overall network by facilitating a high volume of traffic and providing crucial links between different regions.

**Betweenness centrality** measures how often a node lies on the shortest path between other nodes in the network. Nodes with high betweenness centrality act as critical "bridges", controlling the flow of traffic between different parts of the network.

The betweenness centrality of a node $i$ can be computed by:

$$C_B(i) = \frac{1}{N^2} \sum_{s \neq t \neq i} \frac{\sigma_{st}(i)}{\sigma_{st}}$$

where $\sigma_{st}(i)$ is the number of shortest paths from node $i$ to $t$ that pass through $i$, and $\sigma_{st}$ is the total number of shortest paths from node $s$ to $t$.

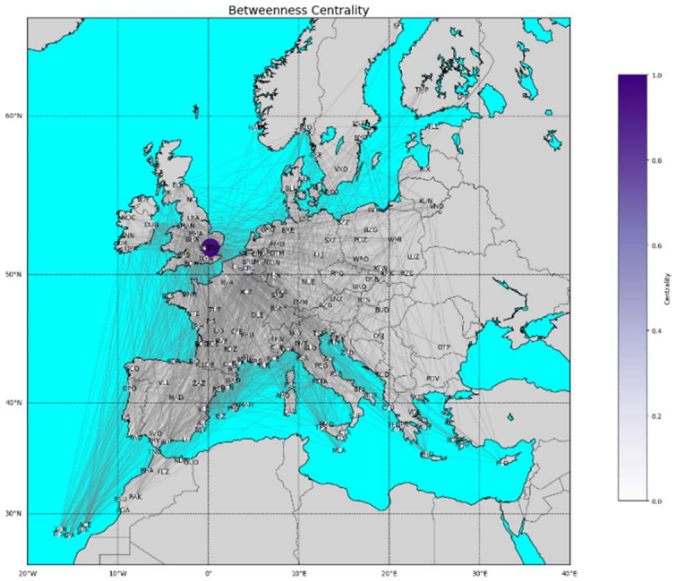


***Figure 5** - Betweenness Centrality of each Airport*

From the figure (fig. 5) above it is possible to see that STN (London Stansted Airport) has a much higher betweenness centrality (0.365) than the rest of the airports.

This airport is a crucial transit hub, acting as a key connector between different regions. It helps maintain the overall connectivity of the network by facilitating traffic between distant or less connected airports.

**Closeness centrality** measures how quickly a node can reach all other nodes in a network. A node with higher closeness centrality has short average paths to all other nodes, meaning it can efficiently access the entire network.

The closeness centrality of a node $i$ can be computed by:

$$C_c(i) = \frac{N-1}{\sum_{j(\neq i)} d_{ij}}$$

where $d_{ij}$ is equal to the distance between node $i$ and $j$.

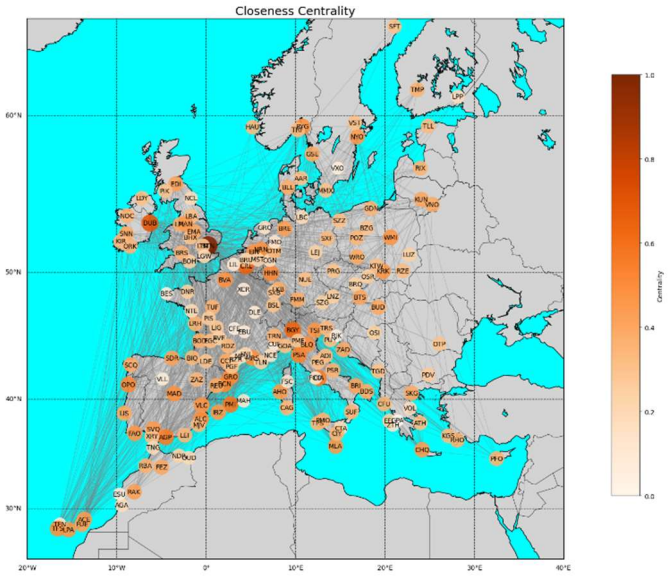


***Figure 6** - Closeness Centrality of each Airport*

From the figure above it is possible to conclude that almost all airports have a relatively high closeness centrality. This means that the network is well connected and efficient, and passengers can travel between any two airports with relatively few layovers or transfers.

### *F. Weak Ties*

In network science, weak ties refer to connections that are less frequent or less strong compared to others but play a critical role in maintaining the structure and flow within the network. In the context of an airport network, weak ties correspond to routes with fewer flights or less traffic. Although these routes may not be primary hubs for travel, they can play an essential role in linking regional airports with larger, central hubs. To identify weak ties in the network, we calculated the edge betweenness centrality for each route.

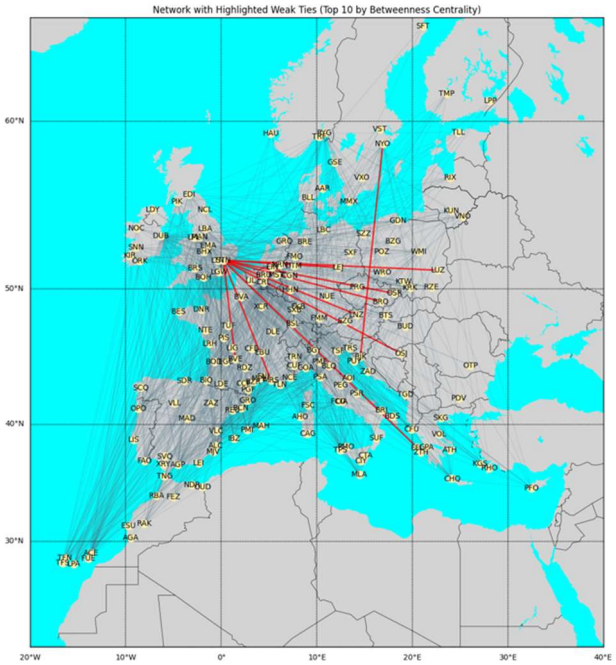


***Figure 7** – Ten weak-ties identified*

From the analysis, the route STN (London Stansted) to MRS (Marseille) has the highest betweenness value of 0.016, followed by several routes from STN to other smaller airports such as BRQ, BVE, and LEJ. This indicates that while these routes may not carry the highest volume of passengers, they serve a critical role in connecting fewer central airports to major hubs.

The removal of these ties could lead to network fragmentation, making certain regions less accessible or requiring passengers to make longer, less efficient journeys. For instance, routes like STN to MRS serve as vital connectors, and their absence could disrupt the flow of passengers between distant parts of the network, affecting resilience and robustness.

## IV. ATTACK SIMULATIONS

In this section, we simulate both targeted and random attacks on the network to assess its robustness under different disruption scenarios. This analysis is important because it allows us to evaluate how the network responds to the removal of key airports, versus random failures.

### *A. Targeted Attack*

In the context of network analysis, a targeted attack is a method used to test the robustness of a graph by deliberately removing the most important nodes. The idea is to remove the nodes that play the biggest role in maintaining the network's structure and connectivity.

The nodes will be removed based on their centrality. The 3 airports that present the highest node centrality are:

- STN (London Stansted Airport)
- DUB (Dublin Airport)
- CRL (Charleroi Airport)

#### *1) Number of Components*

Just like we mentioned in the initial analysis, the network is formed by a single, giant component. However, after removing the three airports, the network was altered significantly. While a large, connected component still remains, the removal of these key airports leads to the formation of 18 isolated components.

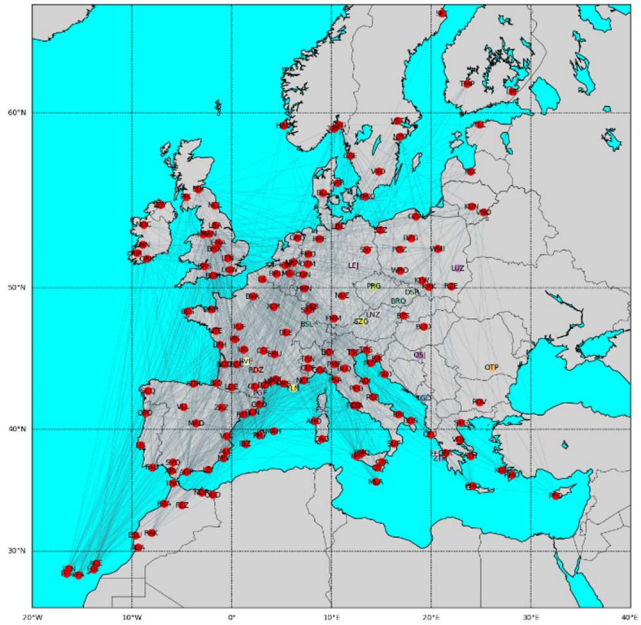

***Figure 8*** *– Isolated nodes after removing important nodes*

In red (fig. 8), it is possible to see the airports that belong to the giant component, and in other colors, are represented the isolated components.

The presence of these isolated components indicates that, although the network maintained a degree of robustness, evidenced by the survival of the giant component, the removal of central nodes still caused notable fragmentation.

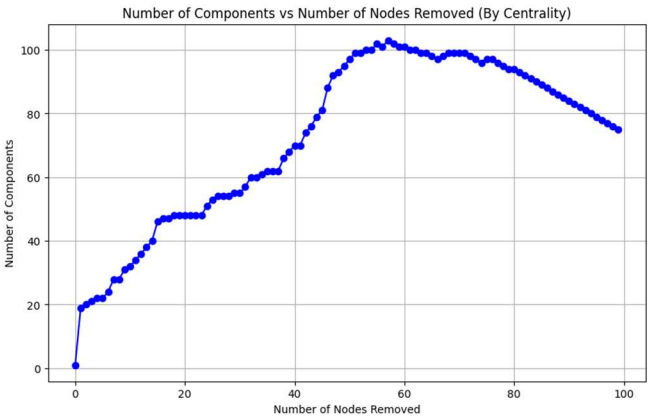


***Figure 9*** *– Number of Components based on Airports removed*

As it is possible to see from this graphic (fig. 9), there is a relatively steep increase in the number of components with the removal of a few nodes, suggesting that the graph is sensitive to disruptions, being likely to fragment quickly. At around 80 nodes removed, it is possible to see a decline, meaning that, as more nodes are removed, the network begins to break apart rapidly, because there is not a giant component anymore, only isolated components.

#### *2) Degree Distribution*

Before the attack, the degree distribution of the network exhibited a long-tail behavior, suggesting the presence of a few highly connected hubs alongside many airports with lower degrees.

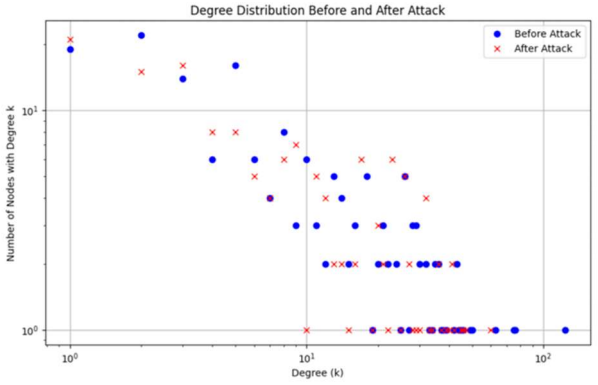


***Figure 10*** *– Plot comparing degree distribution before/after the attack*

The plot (fig. 10) shows a notable drop in the number of high-degree nodes post-attack (red crosses), indicating the removal of the network's core hubs. This is particularly evident in the right tail of the distribution, where nodes with degrees above $10^2$ largely disappear. This reflects the fragmentation caused by the removal of hubs that previously maintained a high number of connections.

The redistribution of degrees post-attack is evident in the increased concentration of nodes with lower degrees (1–10). Many nodes that initially had higher degrees lost several of their connections after the hubs were removed, reducing their degree and contributing to the overall flattening of the distribution.

When we fit the degree distribution to a power-law model after the attack the resulting exponent was unusually high ( $\gamma = 9$). When $\gamma$ is much larger than 3, the degree distribution decays very quickly. This means that the number of hubs becomes very small. The network behaves more like a random network because hubs, which play a crucial role in maintaining the structure of real-world networks, are almost nonexistent.

We can conclude that our network has become more homogeneous because most nodes have a low degree and there is a lack of hubs. Therefore, the network's behavior has become similar to that of an Erdos-Renyi random network.

### 3) *Average Path Length*

| Airports Removed | APL |
|---|---|
| None | 2.1696 |
| STN | 2.3045 |
| STN, DUB | 2.3496 |
| STN, DUB, CRL | 2.4041 |

As expected, it is possible to see the influence of STN airport in the network, as its removal leads to a significant increase in APL. Removing DUB and CRL also has a noticeable impact, although not as severe.

In the beginning (0 to ~30 nodes removed), the APL increases slowly and steadily (fig. 11), indicating that the network remains relatively connected, and the removal of the initial nodes does not drastically affect the overall connectivity. After that, there is a steeper rise, indicating that the removal of these many nodes is starting to fragment the network, making it harder to traverse from one airport to another. In the end, the graphic eventually stabilizes at 0, because, at this point, there are only isolated components, having no routes between the airports, and hence the APL becomes 0.

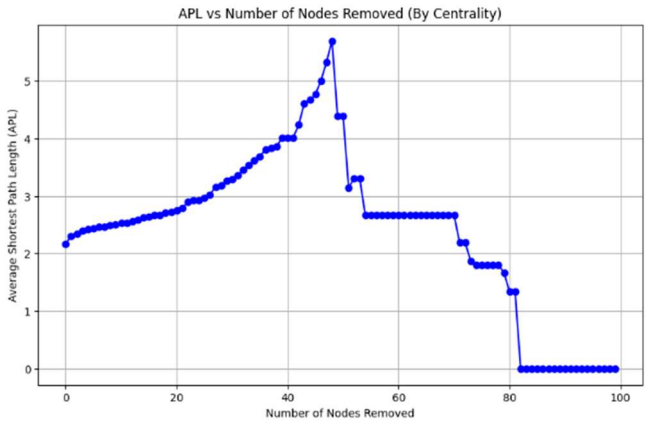


***Figure 11*** *– Average Path Length based on Airports removed*

### 4) *Clustering Coefficient*

The graphic below (fig. 12) demonstrates that the removal of key airports can significantly reduce the clustering coefficient of the network, suggesting that these airports are crucial for maintaining the network's interconnectedness and resilience.

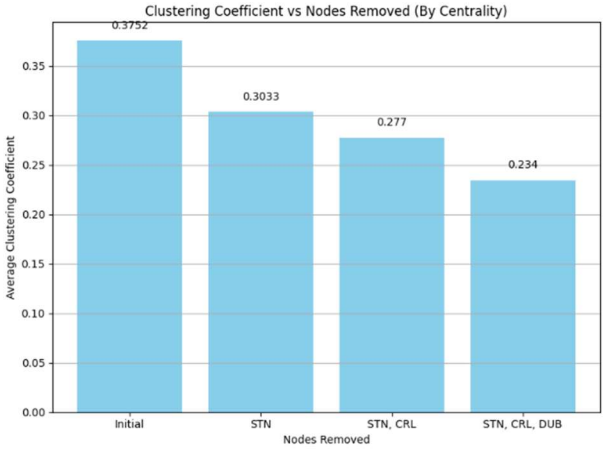


***Figure 12*** *– Clustering Coefficient after removing most important Airports*

The initial (fig. 13) sharp decline highlights the importance of hubs in maintaining the network´s clustered structure. Removing these hubs significantly disrupts the interconnectedness of the network. At approximately 50 airports removed, it is possible to see stabilization of the value of the clustering coefficient at 0, meaning that there are no more clusters in the network.

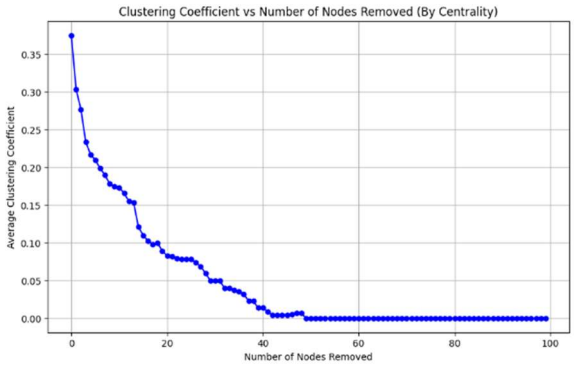


***Figure 13*** *– Clustering Coefficient based on Airports removed*

## B. *Random Attack*

In network science, a random attack refers to the process of removing nodes or edges from a graph. Contrary to a targeted attack, each node is equally likely to be selected for removal, without considering its importance, degree, or position in the network.

### 1) *Number of Components*

Initially (fig. 14) the network remains largely intact, showing resilience as only a few nodes emerge even after the removal of around 65 nodes. As more nodes are removed (between 65 and 150), the network gradually fragments, but

the increase in components is steady and moderate, indicating a slow disintegration.

It can be unusual for the number of components to decrease after having previously increased, however this happens due to the fact that the node removed is an isolated component, reducing the number of components present in the network.

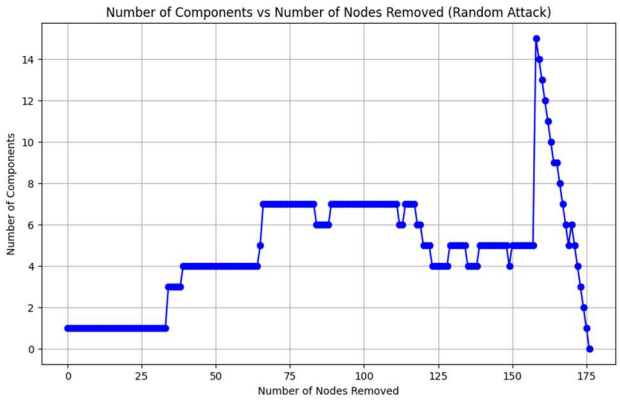


***Figure 14*** *– Number of Components based on Airports removed*

A critical point occurs after removing approximately 160 nodes, where there is a sharp increase in the number of components, likely due to the loss of key nodes that hold the structure together. Following this, the network experiences rapid fragmentation and collapse, reducing back to fewer components.

### *2) Degree Distribution*

In order to evaluate the degree distribution of our network, after performing a random attack, we analyze the evolution of the average power-law exponent (γ) over the course of 100 trials.

At the early stages (0 to ~65 nodes removed) of the random attack, the exponent remains relatively stable around 4, with minimal fluctuations. This indicates that the overall degree distribution of the network is preserved. Therefore, the network demonstrates strong robustness against random removal of small number of nodes.

As more nodes are removed, the network experiences increasing vulnerability as the removal of the higher-degree nodes (hubs) becomes more likely. The larger variability observed in the errors bars suggests that the outcome of each random trial is highly dependent on which specific nodes are removed.

A sharp decline in the average power-law exponent is observed after the removal of approximately 140 nodes. This rapid decrease indicates that the network's scale-free structure is no longer sustainable. At this point, the removal of critical hubs leads to fragmentation and the network becomes highly disconnected.

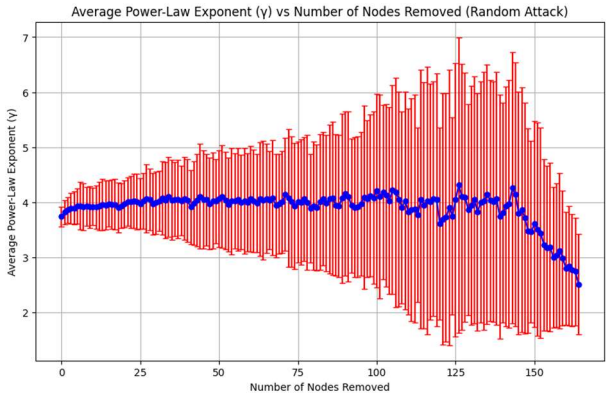


***Figure 15*** *– Power-law exponent based on Airports removed*

### *3) Average Path Length*

At the start (fig. 16), the APL decreases slightly as the network maintains its structure, even with the removal of up to around 125 nodes. This suggests that the network is robust, and its efficiency is not impacted in the early stages of the attack.

However, as more nodes are removed, the APL reaches a peak at about 135 nodes removed, indicating a decline in the network's efficiency, with longer paths between nodes. After about 160 nodes are removed, the APL drops sharply, signaling a rapid fragmentation of the network. At this stage, the giant, connected component disintegrates, resulting in smaller, more disconnected sub-networks where paths are much shorter, and eventually only isolated components, not connected to each other.

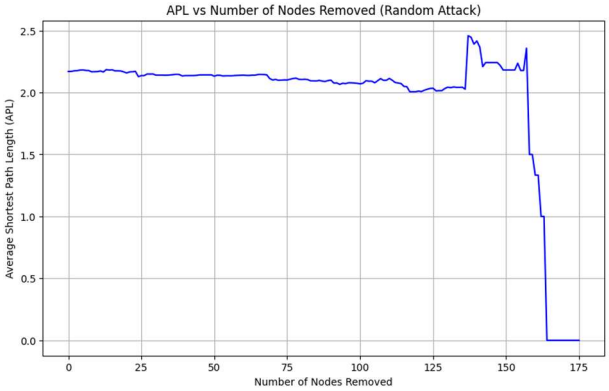


***Figure 16*** *– Average Path Length based on Airports removed*

### *4) Clustering Coefficient*

Initially (fig. 17), the clustering coefficient of the network is relatively high, indicating that the nodes in the network tend to cluster together.

As nodes are removed, the coefficient declines steadily at around 55 nodes removed, suggesting that a central node was removed, disrupting the interconnectedness within the network. Eventually, the clustering coefficient approaches zero, signifying that the remaining nodes are becoming more isolated from the network.

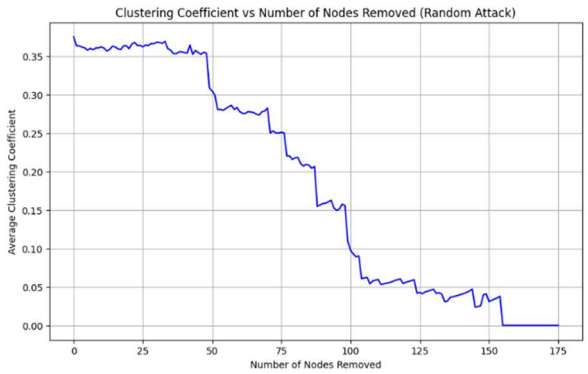


***Figure 17** – Clustering Coefficient based on Airports removed*

## V. CONCLUSION

This report provided a detailed examination of Ryanair's flight network, analysing its robustness and vulnerability to cascading failures under targeted and random disruptions. The results highlighted several key insights about the structure and behavior of the network, allowing us to answer the questions posed at the beginning of the report.

The Ryanair flight network is well connected both globally and regionally, as evidenced by the degree distribution and clustering coefficient. The degree distribution shows a mix of power-law and exponential-like behaviour, indicating that while there are a few highly connected hubs, many smaller airports have fewer connections. This structure supports both global reach through major hubs, and local connectivity through smaller airports. The presence of these highly connected hubs ensures efficient global connectivity, allowing passengers to travel between airports with a reduced amount of layovers.

Also, the clustering coefficient of the network (0.3752), reinforces this connectivity by indicating a moderate level of regional cohesion. Airports like Newcastle (NCL) and Prague (PRG) form highly clustered subgraphs, acting as regional hubs that connect neighbouring airports. This balance between global hub-and spoke connections and regional clusters suggests that the network is designed to facilitate efficient travel across both short and long distances.

The most important airports in the network, based on centrality metric, were identified as London Stanstead (STN), Dublin (DUB) and Charleroi (CRL). These airports have the highest degree centrality, betweenness centrality and closeness centrality, playing critical roles in maintaining the connectivity of the entire network. Stanstead, in particular, serves as a crucial hub, connecting various regions and facilitating efficient travel across Europe.

The removal of key hubs, like the ones mentioned in the previous paragraph, resulted in significant fragmentation of the network. After the targeted attack, the number of components increased dramatically, indicating that large parts of the network became isolated. While the network maintained connectivity through a remaining giant component, the loss of these key airports led to a reduction in the overall efficiency, as seen by the increase in average path length and the significant decline in the clustering coefficient. This suggest that these airports are crucial to preserving not only the connectivity but also the structural integrity of the network.

Comparing targeted and random attacks, we observed very different responses from the network. Targeted attacks on airports with high centrality caused rapid fragmentation, with the number of components rising sharply as few key airports were removed. This underscores the network´s vulnerability to planned disruptions. On the other hand, random attacks displayed more robustness initially. The number of components and average path length remained stable during the early stages of the attack, indicating that the network can withstand random failures of less critical airports. However, after a certain threshold, when more airports were removed, the network also began to break down, although more gradually, in comparison to the targeted attack.

The analysis further revealed that the network can withstand disruptions to a considerable extent before experiencing a total breakdown. However, the tipping point is reached relatively early in the targeted attack simulation, where the removal of around 80 airports leads to a substantial breakdown. In contrast, random attacks only severely impact the network after more than 160 airports were removed. This finding that Ryanair's network has built-in redundancy but remains vulnerable to attacks on its most central hubs.

In summary, this analysis confirms that while Ryanair's flight network demonstrates some resilience due to its small-world properties, it is highly vulnerable to cascading failures triggered by targeted disruptions. Key airports like London Stansted, Dublin, and Charleroi serve as vital connectors in the network, and their removal drastically impacts the network's overall structure. Conversely, random disruptions are less likely to lead to catastrophic failures unless a large number of airports are removed. This insight emphasizes the importance of designing airline networks with robust, resilient structures that can withstand both planned and random disruptions, thereby enhancing their capacity to handle real-world disruptions such as strikes, natural disasters, or technical failures.

## REFERENCES

**[1]** Gao, J., S.V. Buldyrev, S. Havlin, and H.E. Stanley, "Robustness of a network of networks," *Physical Review Letters*, vol. 107, no. 19, pp. 195701, 2011. Available: https://doi.org/10.1103/PhysRevLett.107.195701.

**[2]** Dong, G., J. Gao, R. Du, L. Tian, H.E. Stanley, and S. Havlin, "Robustness of network of networks under targeted attack," *Physical Review E*, vol. 87, no. 5, pp. 052804, 2013. Available: https://doi.org/10.1103/PhysRevE.87.052804.

**[3]** Open Flights, "Flight Route Database," 2021. Available: https://www.kaggle.com/datasets/open-flights/flight-route-database/data.

**[4]** F. Arranz, M. Torné, and C. Juárez, "Cascading Failures in Complex Networks," 2021. Available: https://upcommons.upc.edu/bitstream/handle/2117/344252/Cascading%20failures.pdf?sequence=3&isAllowed=y.

**[5]** R. Guimerà, S. Mossa, A. Turtschi, and L. A. N. Amaral, "The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles," *PNAS*, vol. 102, no. 22, pp. 7794–7799, 2005.

**[6]** Newman, M. E. "The structure and function of complex networks,", *SIAM Rev.* **45**, 167–256 (2003).

**[7]** Oriol Artime, Marco Grassia, Manlio De Domenico, James P. Gleeson, Hernán A. Makse, Giuseppe Mangioni, Matjaž Perc & Filippo Radicchi "Robustness and resilience of complex networks", Nature Reviews Physics vol. 6, 114–131 (2024).

**[8]** Motter, A. E. & Lai, Y.-C., "Cascade-based attacks on complex networks.", Phys. Rev. E 66, 065102 (2002).

**[9]** Albert, R. & Barabási, A.-L., "Statistical mechanics of complex networks.", Rev. Mod. Phys. 74, 47 (2002).