

REPUBLIC OF TURKEY
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

**CENTRALITY AND CONNECTIVITY ANALYSIS OF
AIRPORT NETWORKS: CASES OF TURKEY AND EUROPE**

Cem ERSÖZ

DOCTOR OF PHILOSOPHY THESIS

Department of Statistics

Statistics Program

Supervisor

Prof. Dr. Filiz KARAMAN

May, 2023

REPUBLIC OF TURKEY
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

**CENTRALITY AND CONNECTIVITY ANALYSIS OF AIRPORT
NETWORKS: CASES OF TURKEY AND EUROPE**

A thesis submitted by Cem ERSÖZ in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY is approved by the committee on 11.05.2023 in Department of Statistics, Statistics Program.

Prof. Dr. Filiz KARAMAN
Yıldız Technical University
Supervisor

Approved By the Examining Committee

Prof. Dr. Filiz KARAMAN, Supervisor
Yıldız Technical University

Doç. Dr. Serpil KILIÇ DEPREN, Member
Yıldız Technical University

Doç. Dr. Seda BAĞDATLI KALKAN, Member
İstanbul Ticaret University

Doç. Dr. Gülder KEMALBAY, Member
Yıldız Technical University

Doç. Dr. Çağlar ÜÇLER, Member
Özyeğin University

I hereby declare that I have obtained the required legal permissions during data collection and exploitation procedures, that I have made the in-text citations and cited the references properly, that I haven't falsified and/or fabricated research data and results of the study and that I have abided by the principles of the scientific research and ethics during my Thesis Study under the title of Centrality and Connectivity Analysis of Airport Networks supervised by my supervisor, Prof. Dr. Filiz KARAMAN. In the case of a discovery of false statement, I am to acknowledge any legal consequence.

Cem ERSÖZ

Signature

ACKNOWLEDGEMENTS

I would like to thank my supervisor Prof. Dr. Filiz KARAMAN on guiding me throughout my PhD studies. I would also like to thank my family and my colleagues Hüseyin Önder Aldemir, Sena Kılıç and Özge Yanıkoğlu at Özyeğin University who supported me during this tiring period, without their help, completion of this thesis would have taken much longer.

Cem ERSÖZ



TABLE OF CONTENTS

LIST OF SYMBOLS	vi
LIST OF ABBREVIATIONS	vii
LIST OF FIGURES	viii
LIST OF TABLES	x
LIST OF MAPS	xi
ABSTRACT	xii
ÖZET	xiv
1 INTRODUCTION	16
1.1 Literature Review	16
1.1.1 Analysis of Air Transport Networks Using Complex Networks Theory 18	
1.1.2 European Air Transport Network.....	23
1.1.3 Turkish Air Transport Network.....	24
1.2 Objective of the Thesis	25
1.3 Hypothesis	25
1.4 Data	26
2 METHODOLOGY	27
2.1 Graph Theory	27
2.2 Measures of Centrality and Connectivity	32
2.3 Network Correlations	35
2.4 Random Graphs	36
2.5 Small-World Networks	39
2.6 Scale-free Networks	40
2.7 Community Detection in Complex Networks.....	42
3 ANALYSIS OF THE TURKISH AIRPORT NETWORK	44
3.1 Turkish Airport Network	44
3.2 Topological Analysis of the Domestic Network.....	48
3.3 Connectivity Performance	53
3.4 Comparison of the Domestic Network with Simulated Models	57
3.5 Topological Analysis of the Overall Network.....	60
4 ANALYSIS OF THE EUROPEAN AIRPORT NETWORK	63
4.1 Sample Information	63
4.2 Connectivity and Centrality Analysis.....	64

4.3 Network Correlations	66
4.4 Identification of Community Structure.....	70
5 RESULTS AND DISCUSSION	74
5.1 Discussion of the Turkish Airport Network.....	74
5.2 Discussion of the European Airport Network.....	76
5.2.1 Previous Studies on the EAN	76
5.2.2 EAN and Simulated Networks.....	78
5.2.3 Configuration of the EAN	78
5.2.4 The Development of Air Traffic Market and Strategies of Airline Businesses	79
5.3 Conclusion	80
REFERENCES	82
A APPENDIX	87
PUBLICATIONS FROM THE THESIS	91

LIST OF SYMBOLS

C	Global clustering coefficient
c	Local clustering coefficient
C_B	Betweenness centrality
C_{Bw}	Weighted betweenness centrality
C_C	Closeness centrality
D	Diameter
k	Degree centrality
L	Average path length
R	R statistical programming language
S	Strength

LIST OF ABBREVIATIONS

DGCA	Directorate General of Civil Aviation
EAN	European Airport Network
GDP	Gross Domestic Product
IATA	International Air Transport Association
TAN	Turkish Airport Network



LIST OF FIGURES

Figure 1.1 A point-to-point configuration with 10 vertices requires 45 edges to be fully connected, as the number of edges required in this type of system is calculated by the formula $N(N-1)/2$, where N is the network size.	20
Figure 1.2 A hub-and-spoke network with 10 nodes and 9 edges, where the hub has the highest degree and the spokes have lower degrees. To achieve full connectivity in this type of network, $N-1$ edges are required.....	20
Figure 2.1 An undirected graph with 8 nodes and 7 edges	28
Figure 2.2 A directed graph with 8 nodes and 7 edges.....	28
Figure 2.3 A loop and a multiple edge	29
Figure 2.4 A weighted graph	29
Figure 2.5 A cyclic graph or a cycle of order 10	30
Figure 2.6 A tree is a graph that is both connected and acyclic.....	30
Figure 2.7 A directed graph with 4 vertices	31
Figure 2.8 3 random graphs with 10 vertices. (a) $p=0.2$, (b) $p=0.3$ and (c) $p=0.7$	37
Figure 2.9 Degree distribution of a random graph with 10000 vertices and a connection probability of $p=0.002$	39
Figure 2.10 A scale-free network with 20 vertices generated by simulation....	41
Figure 2.11 Degree distribution of the scale-free network in figure 2.10	42
Figure 3.1 Degree distributions of samples	51
Figure 3.2 Cumulative degree distributions of samples shown on log-log scale	51
Figure 3.3 Normalized betweenness-rank distributions of samples	53
Figure 3.4 Closeness distribution of three samples (all follow power laws)	54
Figure 3.5 Relation between the degree and clustering coefficient of the nodes	57
Figure 3.6 Relation between the degree and closeness centrality of the nodes of sample 3.....	57
Figure 3.7 Visual comparison of the domestic network with simulated models	59
Figure 4.1 Distributions of edge weights and route distances in the EAN. The vertical axes represent the counts of the number of flights for both	

figures. In (a), the average number of passengers handled on the routes is 191,585, but the histogram of route traffic peaks around 107,000, (b) indicates that the average route distance is 1,326 km, but 50% of the routes are shorter than 1,173 km 64

Figure 4.2 (a) Power-law degree distribution where only a small number of airports have a large number of connections. The cumulative degree distribution in (b) illustrates that 25% of airports in the network have 3 or fewer connections, 50% have 9 or fewer connections, and only the top 25% of airports have more than 27 connections. 65

Figure 4.3 The Pearson correlation matrix is a graphical representation of the correlation between different measures of the network. The size of the circles indicates the strength of the correlation coefficient, with larger circles representing higher absolute values. The matrix includes statistically significant ($p < 0.05$) correlations, marked with circles, as well as statistically insignificant correlations, marked with a cross. correlation matrix of the network. 68

Figure 4.4 (a) Degree vs strength on log-log scale and (b) degree vs weighted betweenness. 69

Figure 4.5 The top 30 airports in the European Air Network ranked by their weighted betweenness, with the size of the circle representing the airport's degree, and the color indicating the country in which it is located. The y-axis represents the airport's strength. There is a clear positive correlation between the weighted betweenness and strength of airports, meaning that the busiest airports tend to also have the highest weighted betweenness values.....70

Figure 4.6 (a) modularity and (b) number of communities versus gamma. 71

LIST OF TABLES

Table 3.1 Sampling summary	45
Table 3.2 Airports with the highest aircraft movements	46
Table 3.3 Airports with the highest domestic aircraft movements	46
Table 3.4 Top 5 busiest domestic routes in terms of aircraft traffic	47
Table 3.5 Top 5 busiest international routes in terms of aircraft traffic	48
Table 3.6 Degree rankings of domestic network.....	50
Table 3.7 Eigenvector rankings of domestic network	52
Table 3.8 Normalized betweenness rankings of domestic network.....	52
Table 3.9 Closeness rankings of domestic network.....	54
Table 3.10 Parameters of domestic network and simulated models	58
Table 3.11 Overall network measures	62
Table 4.1 Ranking of airports and routes in terms of passenger traffic.....	66
Table 4.2 Degree (k), eigenvector (k_E), betweenness (C_B) and weighted betweenness (C_{BW}) rankings of airports. Airports are shown with their IATA codes. Full names of the airports can be found in appendix A	67
Table 4.3 7 communities of the EAN identified by modularity maximization ..	72
Table 5.1 Comparison of airport networks of Turkey and other countries/regions.....	75
Table 5.2 Summary of the previous studies and this study on the EAN	77
Table 5.3 Comparison of the EAN with network models that have the same number of nodes and edge density (0.05)	79

LIST OF MAPS

Map 3.1	Operational airports in Turkey in 2020 shown with their IATA codes	45
Map 3.2	Domestic routes of sample 3	49
Map 3.3	All routes flown in sample 3	61
Map 4.1	Community structure of the EAN. Airport labels are scaled for weighted betweenness centrality.....	73



Centrality and Connectivity Analysis of Airport Networks: Cases of Turkey and Europe

Cem ERSÖZ

Department of Statistics

Doctor of Philosophy Thesis

Supervisor: Prof. Dr. Filiz KARAMAN

Complex networks are a powerful framework for understanding the structure and behavior of complex systems, ranging from biological systems to social networks, transportation systems, and the internet. This research aims to analyze the airport networks of Turkey and Europe using complex network theory. The research question focuses on identifying the topological characteristics of air transport networks and how they evolve over time. The importance of air transport in the economic and social development of countries is highlighted. Different data sources are utilized to analyze the structure of Turkey's Airport Network (TAN) and the European Airport Network (EAN). For Turkey, FlightRadar was used to collect data during the initial phase following the first Covid-19 outbreak. In order to understand the structure of the TAN, it is assessed for centrality and connectivity. TAN is compared with simulated network models for classification. It is also compared with other regional and national networks. In the case of Europe, the study uses 2019 passenger data from Eurostat to explore the structure of the EAN. Network measures are computed, and the communities of the EAN are identified using modularity maximization. Furthermore, binary centrality

measures of the network are compared with their weighted counterparts which employ passenger demand and/or route distances as edge weights and a notable distinction is observed between the two. Finally, evolution of the network is analyzed using the findings of previous research papers. The results of this study help to better understand the air transport networks in Turkey and Europe, their similarities, and differences with other air networks worldwide, and can be used to inform policymakers and stakeholders in the aviation industry.

Keywords: Complex networks, Air transport networks, weighted betweenness, community detection



Havaalanı Ağlarının Merkezilik ve Bağlantı Analizi: Türkiye ve Avrupa Örnekleri

Cem ERSÖZ

İstatistik Anabilim Dalı

Doktora Tezi

Danışman: Prof. Dr. Filiz KARAMAN

Karmaşık ağlar, biyolojik sistemlerden sosyal ağlara, ulaşım sistemlerine ve internete kadar geniş bir yelpazede karmaşık sistemlerin yapısını ve davranışını anlamak için güçlü bir çerçeve sunar. Bu araştırma, karmaşık ağ teorisi kullanarak Türkiye ve Avrupa'nın havaalanı ağlarını analiz etmeyi amaçlamaktadır. Araştırma sorusu, hava taşımacılığı ağlarının topolojik özelliklerini ve zaman içinde nasıl evrimleştiğini belirlemeye odaklanmaktadır. Ülkelerin ekonomik ve sosyal gelişmesinde hava taşımacılığının önemi vurgulanmaktadır. Türkiye Havaalanı Ağı (TAN) ve Avrupa Havaalanı Ağı'nın (EAN) yapılarını analiz etmek için farklı veri kaynakları kullanılmaktadır. Türkiye için veriler, Covid-19'un ilk dalgasından sonra açılış döneminde FlightRadar'dan toplanmıştır. Bağlantı performansı ve merkezilik ölçüleri olan derece, aradalık ve yakınlık, TAN'ın yapısını ortaya çıkarmak için hesaplanmıştır. Ağ, sınıflandırma için simüle edilen ağ modelleriyle karşılaştırılmıştır. TAN, diğer bölgesel ve ulusal ağlarla da karşılaştırılmıştır. Avrupa için, çalışma Eurostat'ın 2019 yolcu verilerini kullanarak EAN'ın yapısını keşfetmektedir. Bağlantı ve merkeziyet ölçüleri hesaplanmış ve EAN'ın topluluk yapısı Leiden algoritması kullanılarak incelenmiştir. Ayrıca, havaalanlarının ikili aradalık merkeziliği, havaalanları arasındaki coğrafi mesafe ile yolcu trafiğinin oranını kullanan ağırlıklı bir aracılık ölçüsü ile karşılaştırılmıştır. İki merkezi ölçüsü arasında önemli bir fark gözlemlenmiştir. Son olarak, ağın evrimi önceki araştırma

makalelerinin bulguları kullanılarak analiz edilmiştir. Bu çalışmanın sonuçları, Türkiye ve Avrupa'daki hava taşımacılığı ağlarının dünya genelindeki diğer hava ağlarıyla benzerliklerinin ve farklılıklarının daha iyi anlaşılmasında yardımcı olacağı düşünülmektedir. Bunun yanında, bulgular havacılık endüstrisindeki politika yapıcıları ve paydaşları bilgilendirmek için kullanılabilir.

Anahtar Kelimeler: Kompleks ağlar, Hava taşımacılığı ağları, ağırlıklı ara konum ölçüsü, komünite tespiti



1

INTRODUCTION

1.1 Literature Review

The concept of “*complexity*” refers to a situation in which a vast number of variables interrelate in what appears to be a disorganized manner [1]. The action of individual components within a complex system may be unpredictable, however, the system as a whole tends to display a specific degree of organization and patterns that can be identified and estimated through the application of statistical concepts. For example, motion of individual atoms of a gas in a container cannot be predicted precisely, however, macroscopic properties of the gas such as pressure, temperature etc. which depend on the motion of atoms can be calculated using statistical mechanics. Such problems can be described as having “*disorganized complexity*”. There are also many complex systems in which the variables aren’t as many as the number of atoms in a gas, but also not as few as to describe the system as exhibiting simplicity. In these systems, the variables may work in tandem for their mutual benefit, resulting in parts of the system that are in some way interdependent or correlated. For instance, the price of a particular product is determined by the complex interaction of numerous economic factors, while the behavior patterns of an organization like a labor union is an outcome of the interaction of its members. These systems are referred to as demonstrating “*organized complexity*”, which characterizes majority of the social and economic networks.

Social and economic networks can be considered as complex systems due to their ability to be divided into hierarchical subsystems. For instance, a business organization has clear distinct parts from top management to most elementary labor units (staff) which interact continuously in a command-report manner. A system, such as an economy, that can be modeled with an input-output matrix i.e., price, demand quantity, supply quantity etc. variables is said to be *decomposable*

[2]. Many complex systems exhibit these two characteristics of being structured hierarchically and decomposability into subsystems. A complex system will always have a certain goal or a final state to achieve by means of different processes. For example, for a governmental organization, the goal could be generating high employment rate, or for a transportation system, the final state could be to link as many routes as possible to the main hub. Finding the required equations for processes that will lead the system to reach its final state is the science of complex systems.

The study of social networks started as early as the 1920s, with a focus on the interactions between various entities, including communication among members of social groups, trade, and the flow of goods between regions and countries, and commercial activities among businesses. A complex network is identified by a multitude of elements that are arranged in non-uniform configurations, with a dynamic framework that varies over time [3].

In scientific research, it is usually assumed that many natural occurrences as well as socio-economic variables are distributed so as to obey the normal curve. However, data coming from complex systems research shows that many indicators like people's incomes, forest fires, heart rate etc. obey power-law distributions. Power law distributions occur so often that many scientists nowadays consider them as a norm just like the normal distribution [4].

Systems such as the internet, references related to scientific research, biological networks, and others might be regarded as grids exhibiting complexity. These networks exhibit two primary features: a constant expansion pattern as new edges grow the network continuously and new edges generally arise from nodes that are already well-connected. These traits are fundamental to the formation of "scale-free networks", suggesting that the growth of these networks adheres to patterns of self-organization patterns that must be examined holistically, instead of a component-by-component basis [5].

Scale-free networks exhibit the property of being resilient to random removal of vertices, that is, the network maintains its basic properties under random attacks. This is due to most of the vertices having low-degrees in a network that has scale-

free characteristics. This type of networks' degrees are distributed according to a power-law, nodes are clustered and the system possesses a few hubs with many connections and many nodes with fewer connections [4].

1.1.1 Analysis of Air Transport Networks Using Complex Networks Theory

According to research [6], complex networks such as transportation, power grids, and communication infrastructures play a crucial role in world's socioeconomic elements, including employment rate and GDP growth. In the year 2018, the worldwide aviation sector generated 65.5 million job opportunities, out of which 10.2 million were categorized as direct employment. The industry directly contributed \$704.4 billion to the world's GDP. Additionally, the industry made a total contribution, encompassing both direct and indirect effects, of 3.6% to the world's GDP [7].

The conventional approach to modeling aviation networks is to consider them as a set of airports, where each airport is represented as a node in the network [8]. An alternative approach has been adopted, which uses cities as nodes rather than airports and collapses cities with multiple airports into a single node [9]. These networks exhibit several characteristics [10]:

1. The size of the network is limited and varies depending on the extent of the research, with a maximum of only a few thousand nodes.
2. Nodes are stationary and their positions are fixed over time.
3. Edges (representing flights, passengers, and cargo) are typically weighted and mostly bidirectional.

The network of airports connected by flights is a fundamental aspect of the aviation industry. Enlarging the global aviation network can have a positive impact on the economy, employment, and the general wellbeing of people. Consequently, comprehending the structure of air transport networks is essential for air transport organizations to adjust their business tactics to changing market dynamics.

Operating an air transport network necessitates substantial investment because of the expenses involved in constructing airports and running aircraft. In addition, this network is vulnerable to external factors like natural calamities or acts of terrorism. Therefore, it is necessary to identify the airports that are of significant importance to the network [11]. To assess the robustness of an aviation network, it can be analyzed under different attack (vertex removal) scenarios. After the attack, previous network metrics can be compared to new ones and the vulnerability of the network can be calculated [12]. For example, removing Van airport will probably not have much impact on the overall operation of Turkish Airlines flight network. On the other hand, if a major hub like Istanbul airport is removed from the network, operation of many airlines will be affected and there will likely be many flight delays and cancellations resulting in financial loss.

The principles of complex network theory have impacted various fields of scientific research. In recent times, there has been an increasing focus on applying these principles to air transport networks, resulting in the identification of various topological structures and measures. In an air transport network, airports are connected by edges that represent flights, passenger numbers, cargo volumes, or other factors. These edges can be one-directional, bidirectional, or weighted [13]. However, since most airlines operate return flights, it is common to treat the network as undirected without assigning specific directions to the edges [8].

There are two primary network types connecting airports in terms of topology; “point-to-point” and “hub-and-spoke”. In the former case (Figure 1.1), each airport is directly linked to all other airports through different aircraft. In order to be fully connected, this system’s number of links increases with the second power of the network size. In contrast, a hub-and-spoke network (Figure 1.2) possesses a limited number hubs that connect to nodes with lower connectivity. This system needs smaller number of links compared to a point-to-point configuration of the same size to achieve connectivity, which makes it more economical for airlines [14]. Hub-and-spoke networks are categorized as scale-free networks, in which the degrees are distributed as a power law.

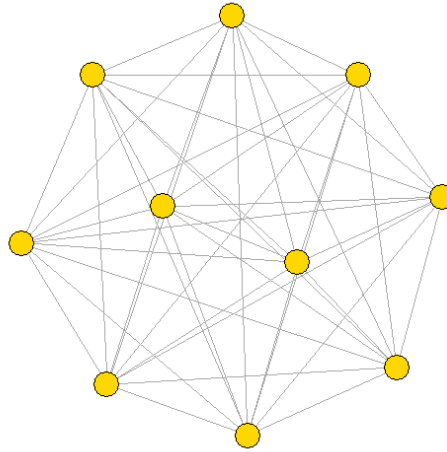


Figure 1.1 A point-to-point configuration with 10 vertices requires 45 edges to be fully connected, as the number of edges required in this type of system is calculated by the formula $N(N-1)/2$, where N is the network size.

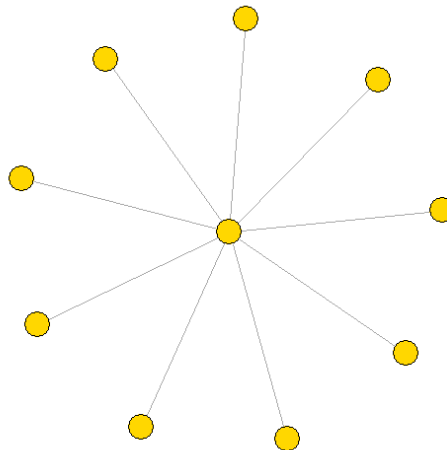


Figure 1.2 A hub-and-spoke network with 10 nodes and 9 edges, where the hub has the highest degree and the spokes have lower degrees. To achieve full connectivity in this type of network, $N-1$ edges are required

A very significant characteristic of airports in an aviation network is their *centrality*. There are different centrality measures for airports such as *degree centrality* and *betweenness centrality*. In a hub and spoke network, hub airports are the highest degree airports and most of the airports have low degrees. In a point-to-point network, airports will have similar degrees. The connectivity of airports is another issue that needs to be analyzed in an air transport network. It is essential

to start the network analysis by computing centrality and connectivity measures to assess the importance of airports and to determine the type of the network.

The study conducted by Guimera and Amaral in 2004 revealed that the worldwide network of cities with airports demonstrated characteristics of a “small-world network” with a large “clustering coefficient” and a small “characteristic path length” [15]. The network was understood to have a community structure which was thought to arise due to geopolitical reasons. The number of connections of the airports were found to follow a truncated power law $P(k) \sim k^{-\gamma} f(\frac{k}{k_x})$ with $\gamma = 1.0 \pm 0.1$ and $k_x \sim S^{0.4}$ where S is the size of the network. Similarly, betweenness centrality of the network was also observed to obey a truncated power law. Another finding of the research was that cities with high betweenness centrality were not necessarily the ones with the highest degrees, indicating that there were many cities having small degrees but large betweenness values. Researchers considered this issue to be arising from preferential attachment and geo-political constraints.

Another study on the global air transport network in 2014 by Lordan et al [12], used airports instead of cities and the airport network was taken to be undirected due to most flights being bi-directional. The study revealed that the network has a low average path length, which confirms the small-world characteristic, and a high clustering coefficient. Additionally, both the degree and betweenness distributions of airports followed a truncated power law.

A 2003 research done by Chi et al [16], where the directed and unweighted United States flight network were analyzed, revealed that the network had a much higher clustering coefficient than a random network with the same size and average degree. Daily network evolution was analyzed using a 1-week time window and it was observed that in-degree, out-degree and overall degree distributions were almost identical and followed power-laws.

A case study on Australian airport network done by Hossain and Alam [17] analyzed both the unweighted and the weighted route network. It was seen that Australian Airport Network had the characteristics of a “small-world network”, with a longer characteristic path length than a “random network” that has the

same number of nodes. The cumulative degree distribution was identified as a power law and network's most connected airports were not the ones with the highest betweenness and closeness centralities pointing to bottlenecks in the network. A new approach adapted in this study was the calculation of weighted clustering coefficient which was observed to be smaller compared to the unweighted case. This finding was thought to occur due to topological clustering originating from low weight links confirming the hub-and-spoke configuration.

A Chinese airport network analysis conducted on a daily basis for a week in 2003 by Li and Cai [10] used both weighted and unweighted networks and demonstrated that the network exhibits both small-world and scale-free features. A high clustering coefficient and a small average path length compared to a random network with the same features were identified. In-degree, out-degree and undirected degree distributions were observed to follow two-stage power laws which have two different exponents with transition points occurring at a certain degree value. Furthermore, the negative degree-degree correlations of the vertices indicated that the network was disassortative i.e., the high degree vertices were usually connected to the low degree vertices signaling a hub-and-spoke structure.

Another city-based Chinese aviation network study by Lin in 2011 firstly reviewed the growth mechanisms of the air network [18]. The author claimed the network to have mostly developed under the influence of the government and civil aviation market in China to have never been fully liberalized as in US and Europe. The paper also emphasized on the influence of the geographical distance on aviation networks and that spatial effects should not be neglected when analyzing the structure of these networks. Hub-and-spoke (scale-free) structure and small-world characteristics of China's flight network was demonstrated. Moreover, it was shown that effects of distance should be measured separately for short-range and medium- to long-range flights. The network was found to be organized hierarchically as the degree-clustering coefficient correlation followed a power law.

In their study, Lordan and Sallan analyzed the robustness of world air transport network [19] which they define to be "resilience facing vertex isolation". In other

words, they analyzed the connectivity of airport network by removing airports randomly and by removing specific airports. By simulating different attack scenarios, what they found was the aviation network, classified as scale-free, was resilient to random removal of nodes (airports) however, not to removal of nodes that were selected based on centrality or importance.

1.1.2 European Air Transport Network

In this section, the focus is on studies that have been conducted on the European air network. Several papers investigated the European air network using the lens of complex network theory. One such study conducted by Cardillo et al. [20] utilized one-day flight data to examine the cumulative degree distribution, average clustering coefficient, average path length, and average size of the "European air transportation multiplex network." This network was created by merging individual networks of European commercial airlines. The authors discovered that the characteristics of the network were a result of its multi-layered structure.

Another study, by Sun, Wandelt, and Linke [21], analyzed the seasonal and weekly changes in the European air transportation system using both the air navigation route network and airport network. The researchers used 217 days of seat data and found that the airport network displayed differences in terms of clustering and average path length depending on the time of year (summer/winter) and day of the week (weekday/weekend). The study also identified hub nodes in the network that could potentially serve as bottlenecks.

Pien et al. [22] revealed that the traditional measure of betweenness centrality is not adequate for assessing the robustness of the Europe's aviation network. As a result, the authors proposed a new network measure called the "relative area index", which is better suited for measuring the importance of nodes in the event of network disruptions.

Wandelt, Sun, and Zhang [23] examined the topological changes in the European network between 2002 and 2013 using passenger traffic as the edge weights. The authors found that passenger traffic increased significantly during this period and had a distinct seasonal pattern. Additionally, the study revealed that the degree

distribution of the network generally followed a truncated power law, indicating that the network had a scale-free property.

Finally, Gurtner et al. [24] conducted a study on the 2010 European air traffic network over a 28-day period, looking for communities within the network. The study weighted the network edges using number of flights between the nodes. By perturbing the communities found by different algorithms, the authors were able to assess the robustness of these communities. The study also showed that a method called multiresolution modularity maximization was more effective than other methods in uncovering the community structure of the network.

1.1.3 Turkish Air Transport Network

Several studies have examined the domestic air transport network in Turkey, including those conducted by Song and Yeo [25], Depren and Yavuz [26], and Erdem et al. [27]. However, each of these studies had certain limitations and failed to provide a comprehensive understanding of the network. For example, Song & Yeo provided an overview of the general characteristics of the network but did not analyze its structure and topology. Depren & Yavuz conducted a air transport network analysis but only took into account 38 airports as nodes and did not account for recent changes in the network, such as the closure of Istanbul Atatürk Airport. Erdem et al. focused only on densely populated cities, limiting the generalizability of their results. Therefore, there is a need for a more comprehensive study to better understand Turkey's air transport network.

Turkish air transport network has garnered considerable attention for various reasons, primarily due to its unique geographic shape, which takes on a rectangular form, making travel across the country by any other means other than air travel quite cumbersome [28]. Consequently, air transportation has become the preferred mode of travel between cities, especially for those traveling to and from Istanbul, the country's economic hub, and Ankara, the political capital. In the past few years, the establishment of airports in nearly all cities has led to a rise in air travel between them, causing the air network to expand and become more intricate. The development of the tourism industry in Turkey, alongside the increasing popularity of cultural and historical sites in various cities, has resulted

in increased passenger traffic, both domestically and internationally, from central cities to other parts of the country [29]. Lastly, the growth of education and health opportunities in all cities, as well as the influx of immigrants from various countries, have fueled the demand for transportation to all corners of the nation [30]. As a result, airlines operate in an intensely competitive environment, where proper route and network planning has become critical to their success.

1.2 Objective of the Thesis

This research intends to identify the topology of the Turkish and European air transport networks by employing the tools of complex networks theory by utilizing flight and passenger traffic data. Centrality and connectivity performance of airports in the network is to be assessed by examining the topology of the airport network and determining its related distributions. Node- and edge-based centrality and connectivity measures are to be computed. Optimal community structure of the network is to be identified.

Another objective of this thesis, the computation of a weighted node betweenness, using the ratio of route distance to passenger traffic as an effective distance between nodes, is to be conducted to determine the airport hub performance. Additionally, a comparison of this measure to the binary betweenness centrality is to be presented to highlight the shortcomings of the latter in assessing the centrality of airports.

1.3 Hypothesis

The first hypothesis of this thesis is that the networks under consideration are both “small-world” and “scale-free” networks. The thesis aims to prove this by analyzing the structural features of the networks.

The second hypothesis of this thesis is that binary betweenness centrality is not sufficient to measure the hub performance of an airport in an air transport network. The thesis aims to prove this by introducing a weighted node betweenness centrality that considers the geographical distance and passenger demand between the airports. Centrality measures of the networks are to be compared using correlation analysis between degree, strength.

1.4 Data

To examine the air transport network of Turkey, we gathered data on domestic and international passenger flights from FlightRadar for three different periods: the final weeks of July, August, and September. These periods were chosen because they have similar flight schedules. In the initial stages of the Covid-19 pandemic, flight schedules were adjusted frequently, so monthly intervals were utilized to monitor variations in the flight schedules. The FlightRadar data was analyzed using R. The study focuses on the summer season during the Covid-19 pandemic, which is an external shock and a unique scenario for the air transport industry. As a result, the findings are not generalizable. The three samples collected were: Sample 1: 29.07.2020 - 05.08.2020, Sample 2: 26.08.2020 - 01.09.2020, and Sample 3: 25.09.2020 - 01.10.2020.

For the study of the Europe's air transport network, the data was collected from Eurostat's website. The website contains individual country datasets that show the passenger, seat, and flight traffic of airports from 35 European countries. The datasets were in the form of monthly and yearly time series and were filtered to only include data from the year 2019. The filtered datasets were used to create a passenger traffic-weighted network for each country. These 35 networks were combined and filtered to only include traffic between and within European countries. To keep calculations simple, the final network was considered undirected, and the edge weights were based on cumulative traffic on the routes. In cases where airport pairs had slightly different values for cumulative traffic on the routes, the maximum value was taken as the edge weight.

All computations and visualizations were done using R's Igraph and Ggplot packages. The Airportr package was used to obtain airport coordinates, and the Geosphere package was used to calculate distances between airport pairs using the ellipsoid distance function for high precision. The LeidenAlg package was used to identify community structure by maximizing modularity. It's worth noting that all software used in this study is open-source and freely available on the internet.

2.1 Graph Theory

With the advent of computers with fast processors and high memory capacities and the availability of large amounts of data, it is now possible to model and simulate systems of organized complexity in order to explain their characteristics [3]. Representing complex systems through network models is a widely adopted approach. In this approach, the variables of the system are depicted as nodes or vertices, and their interactions are represented as edges or links. Graph theory concepts can be employed to mathematically formulate and analyze these networks. Some characteristics of “random graphs”, “small-world networks” and “scale-free networks” are essential for analyzing complex networks, which will be reviewed in this chapter.

A mathematical representation of a network called a graph, denoted by G , is made up of two distinct sets, V and E . The set V is composed of non-empty items, also referred to as vertices or nodes, while the set E is composed of pairs of subsets of V , also known as edges or links. These names will be used interchangeably in this thesis. Figure 2.1 shows the vertices $V = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ as the dots and edges $E = \{\{x_1, x_2\}, \{x_1, x_3\}, \{x_1, x_4\}, \{x_1, x_5\}, \{x_5, x_6\}, \{x_4, x_6\}, \{x_7, x_8\}\}$ as the lines connecting the vertices (x_i — x_j notation may also be used for edges). If $E=0$, then there are no edges between the vertices [31].

A graph is *undirected* if edges $\{x_i, x_j\}$ and $\{x_j, x_i\}$ are the same (symmetric relationship) as the one in figure 2.1, and, if the relationship between the vertices is one way, it is named a *directed graph* or a *digraph* (Figure 2.2). Arrows are drawn in order to show the directions of the edges on graphs. A particular directed edge from x_i (*head*) to x_j (*tail*) is written as (x_i, x_j) or $x_i \rightarrow x_j$ [32].

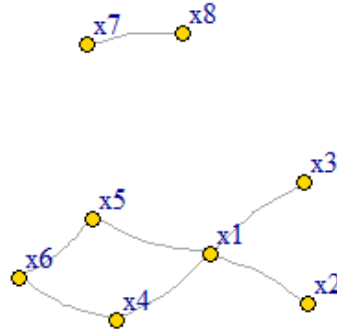


Figure 2.1 An undirected graph with 8 nodes and 7 edges

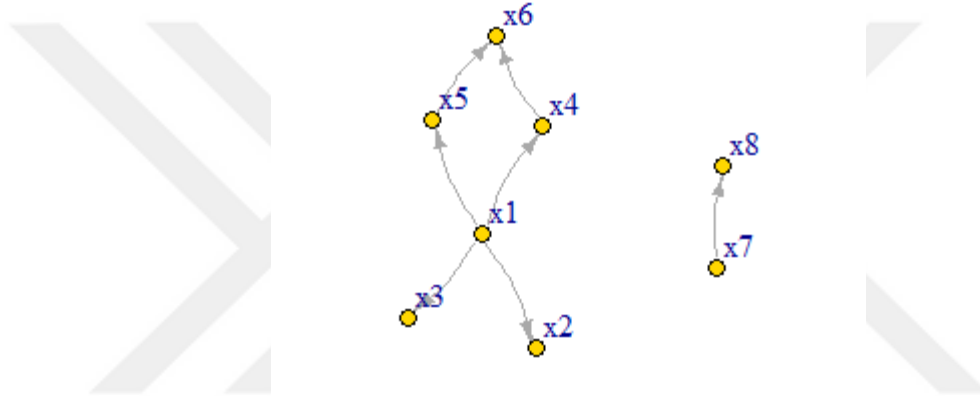


Figure 2.2 A directed graph with 8 nodes and 7 edges

A graph is considered simple if it does not contain any loops or multiple edges. (Figure 2.3 shows a loop and a multiple edge). Two vertices x_i, x_j are adjacent if $\{x_i, x_j\} \in E$ and an edge $e = \{x_i, x_j\}$ is incident to x_i, x_j . The degree k of a vertex is the number of edges that are connected to it. For directed graphs, every vertex has an in degree and an out degree. For example, vertex x_4 in Figure 2.2 has 1 in degree and 1 out degree.

A graph $G_1 = (V_1, E_1)$ is called a *subgraph* of a graph $G_2 = (V_2, E_2)$ if $V_1 \subseteq V_2$ and $E_1 \subseteq E_2$. Given a graph $G = (V, E)$, a *matching* is a subgraph of G , where every vertex has degree 1.

A *weighted graph* is a simple graph in which each edge is associated with a real number showing the *weight* or *strength* of that edge (Figure 2.4). In other words,

a weighted graph is made up of a graph $G = (V, E)$ and its corresponding *weight function* w .

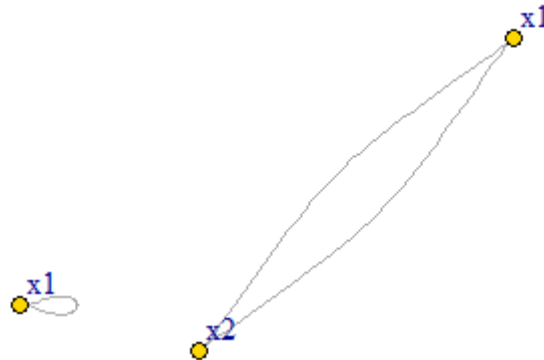


Figure 2.3 A loop and a multiple edge

A graph $G_1 = (V_1, E_1)$ is called a *subgraph* of a graph $G_2 = (V_2, E_2)$ if $V_1 \subseteq V_2$ and $E_1 \subseteq E_2$. Given a graph $G = (V, E)$, a *matching* is a subgraph of G , where every vertex has degree 1

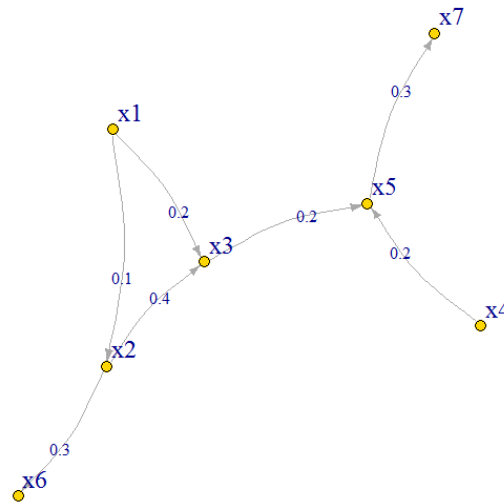


Figure 2.4 A weighted graph

A *walk* is a sequence of vertices $x_1—x_2—\dots—x_k$ of length $k-1$ that are connected by edges. A *path* is a walk where all vertices v_i are different. A *closed walk* is a walk which starts and ends at exactly the same vertex, If $k \geq 4$ and $x_1, x_2, \dots, x_{k-1}$ are all different then it is called a *cycle*. A cycle's *order* or *size* is determined by the number of vertices it has. (Figure 2.5). Average degree of a cycle is always 2.

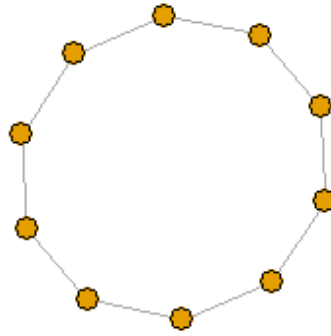


Figure 2.5 A cyclic graph or a cycle of order 10

Vertices x_i and x_k are *connected* if there is a path from x_i to x_k . A graph is connected when every pair of vertices in the graph are connected. A connected and acyclic graph is called a *tree* (Figure 2.6). A *leaf* is a vertex with degree 1 in a tree. Any connected subgraph of a tree is a tree. A tree with n vertices has $n-1$ edges and therefore is of order n . Average degree of a tree of order n is $2 - \frac{2}{n}$ which approaches to 2 as the tree size gets larger.

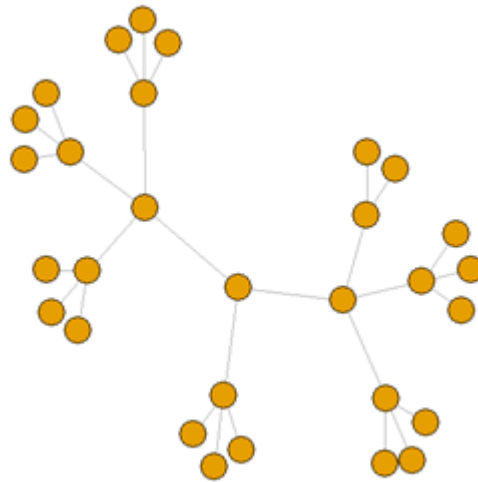


Figure 2.6 A tree is a graph that is both connected and acyclic

A popular method of depicting a graph G is by drawing it, but it can also be represented mathematically using an adjacency matrix. An adjacency matrix of a graph $G = (V, E)$, where $V = \{x_1, x_2, \dots, x_n\}$ denotes the nodes, is a square matrix $A = \{a_{ij}\}$ that shows the connections between nodes using 1s and 0s:

$$a_{ij} = \begin{cases} 1 & \text{if } \{x_i, x_j\} \in E \\ 0 & \text{otherwise} \end{cases} \quad (2.1)$$

For example, the directed graph in Figure 2.7 can be represented with the adjacency matrix:

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

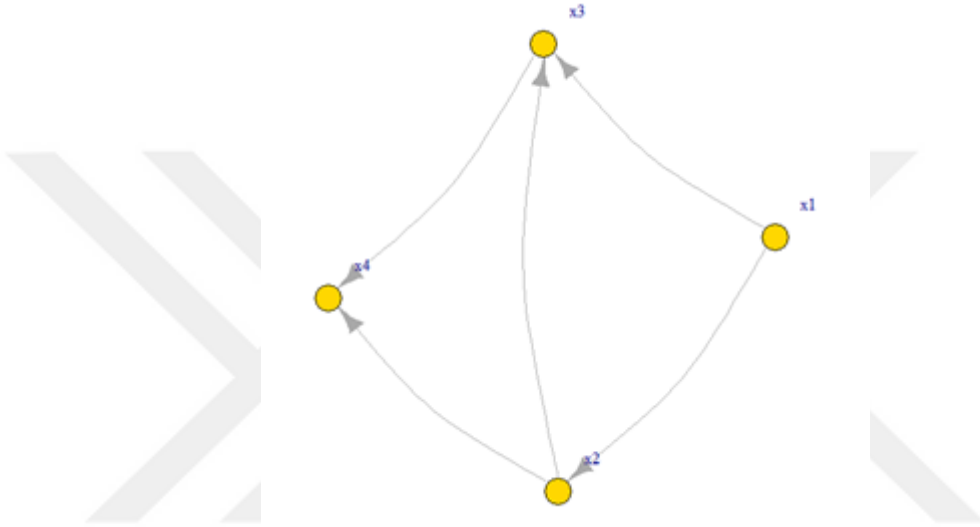


Figure 2.7 A directed graph with 4 vertices

If G is a weighted graph, 1 is replaced by the weight of the corresponding edge.

Usually, there will be more than one walk between the vertices of a connected graph. For example, in Figure 2.7, a *directed walk* starting from x_1 and ending at x_4 can be completed by the walks $x_1 \rightarrow x_3 \rightarrow x_4$ or $x_1 \rightarrow x_2 \rightarrow x_4$ or $x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4$. For a given pair of vertices x_i and x_j in a graph $G = (V, E)$, the number of walks of length k between the two vertices is equal to the k^{th} power of the (i, j) element of the adjacency matrix ($a_{ij}^{(k)}$). For instance, the number walks of length 2 between x_1 and x_4 of the graph in Figure 2.7 can be found by taking the square of the adjacency matrix.

$$A^2 = \begin{bmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Here, $A_{14} = 2$ is the number of walks of length 2 between x_1 and x_4 which can also be shown graphically.

Finding the *shortest path* between any pairs of nodes of a network is another problem in graph analysis. One way to find this is to calculate the powers of adjacency matrix until $a_{ij}^{(k)}$ exceeds zero.

2.2 Measures of Centrality and Connectivity

To determine the significance of nodes in a network, various performance measures are applied. The most fundamental centrality measurement for a graph is degree centrality, which simply signifies the overall number of connections (incoming and outgoing) a node has with other nodes. The degree (k) of node i can be calculated as follows:

$$k_i = \sum_{j=1}^v a_{ij} \quad (2.2)$$

where a_{ij} are the components of the adjacency matrix. The mean degree of nodes in a graph can be determined through the following calculation:

$$\langle k \rangle = \frac{1}{n} \sum_{i=1}^n k_i \quad (2.3)$$

The *strength* s of a node is computed with the following:

$$s_i = \sum_{j=1}^v a_{ij} w_{ij} \quad (2.4)$$

Here, w_{ij} shows the weight of the edge linking nodes i and j [33].

An extremely significant descriptive statistic of a graph is its degree distribution function $p(k)$ that defines the likelihood of a vertex chosen randomly has a degree equal to k . In a graph with n vertices and n_k of them having a degree of k , the degree distribution function $p(k)$ can be expressed as:

$$p(k) = \frac{n_k}{n} \quad (2.5)$$

In a similar manner, the cumulative degree distribution $P(k)$ which signifies the likelihood that a randomly chosen vertex has k or more degrees, can be calculated as follows:

$$P(k_i \geq k) = \sum_k^{\infty} p(k) \quad (2.6)$$

Eigenvector centrality is an alternative vertex centrality metric that expands upon degree centrality. In degree centrality, all edges are considered to hold equal value, however, eigenvector centrality assigns higher value to connections with nodes that already possess high connectivity [34]. This measurement is computed as follows:

$$k_E(i) = \frac{1}{\lambda} \sum_{j=1} a_{ij} k_E(j) \quad (2.7)$$

where λ represents the eigenvalue of the eigenvector $\mathbf{k}_E$ of the adjacency matrix.

The shortest path length d_{ij} measures the shortest distance between vertices i and j in a network, and it is equivalent to the number of edges in the shortest path connecting them. To assess the network's connectivity, the average of all shortest paths, also known as the average shortest path length L , or the characteristic path length, can be computed as follows [35]:

$$L = \frac{1}{n(n-1)} \sum_{i \neq j} d_{ij} \quad (2.8)$$

A lower value of L indicates a network with better connectivity, while a network with a large value of L will have some vertices that are farther apart from each other, which reduces the network's performance. The connectivity of a graph can also be assessed using its diameter. The diameter (d) of a graph is defined as the longest shortest path length between vertices [19].

The clustering coefficient of a vertex measures the likelihood of two neighboring vertices being connected to each other. It is calculated by taking the number of existing triangles involving a certain vertex and dividing it with the maximum

number of triangles that can be formed. The formula to calculate the local clustering coefficient of a vertex (c_i) is as follows:

$$c_i = \frac{1}{k_i(k_i-1)} \sum_{i,j,k} a_{ij} a_{jk} a_{ik} \quad (2.9)$$

The clustering coefficient of a vertex determines the degree to which a vertex is part of a clique-like group, or the proportion of triangles that involve a certain vertex compared to the maximum number of triangles that could be formed. The global clustering coefficient of the network is computed by taking the mean value of all local clustering coefficients and is expressed as:

$$C = \frac{1}{n} \sum c_i \quad (2.10)$$

This global clustering coefficient measures the cohesiveness of a group of vertices and serves as a metric for evaluating interconnectivity in the network. A fully connected or a complete graph will have a value of 1 for the clustering coefficient, which signifies that each vertex in the network is connected to all other vertices. Additionally, there is another metric, the weighted clustering coefficient c_i^w [36] that considers the intensity of the interactions within triangles in the network. It is calculated by summing the weighted connections between vertices i , j , and k and dividing by the number of possible connections:

$$c_i^w = \frac{1}{s_i(k_i-1)} \sum_{j,k} \frac{(w_{ij} + w_{ik})}{2} a_{ij} a_{jk} a_{ik} \quad (2.11)$$

The betweenness centrality (C_B) of a node indicates the importance of that node in the network by calculating how often it falls on the shortest path between other nodes. It is expressed mathematically as the sum of the number of shortest paths between all other node pairs (excluding the node itself) that pass through it, divided by the total number of shortest paths between those nodes:

$$C_B(i) = \sum_{s \neq t \neq i} \frac{Q_{st}(i)}{Q_{st}} \quad (2.12)$$

where the overall count of shortest paths connecting any two nodes s and t is represented by Q_{st} and $Q_{st}(i)$ represents the number of shortest paths between s

and t that go through node i [37]. This centrality measure provides insight into how nodes perform as hubs in the network. However, in a weighted network where some edges are more preferred due to their capacity to handle more traffic, a more accurate representation of betweenness centrality is to use effective shortest paths as weights. This approach was suggested by Dall'Asta et al. [38]. The effective distance between nodes i and j in an air transport network can be determined by considering both the edge weight (w_{ij}) and the geographical distance (d_{ij}) between the nodes. The effective distance is calculated by dividing the geographical distance by the edge weight, so that a longer route distance may be offset by a higher flow of traffic:

$$l_{ij} = \frac{d_{ij}}{w_{ij}} \quad (2.13)$$

When determining the weighted betweenness centrality in a network, instead of finding the shortest path based on network topology, the betweenness algorithm seeks the path that minimizes the effective distance between nodes.

2.3 Network Correlations

The manner in which nodes are linked in a network can be evaluated by degree-degree correlation. When nodes with a large number of connections tend to connect to other nodes with a large number of connections, the network is said to have assortative mixing. Conversely, if high-degree nodes have a tendency to link to low-degree nodes, the network is said to have disassortative mixing [39]. We aim to quantify the degree of assortativity within the air transport network by utilizing the assortativity coefficient ρ [40]:

$$\rho = \frac{\langle kk' \rangle_e - \langle k \rangle_e^2}{\langle k^2 \rangle_e - \langle k \rangle_e^2} \quad (2.14)$$

where $\langle \dots \rangle_e$ means averaging over all edges. In a disassortative network, the nodes with low degree of connections typically link to high-degree nodes, resulting in a hub-and-spoke structure [41]. When the coefficient is near -1, the network can be categorized as having a hub-and-spoke design, and if it is close to

+1, it can be characterized as a point-to-point network where links are established directly, bypassing the need for hubs.

The relationship between the degree and clustering coefficient is a crucial aspect of network analysis. If the correlation between these two values follows a power law relationship $C(k) \sim k^{-\alpha}$, this suggests that the network is organized in a hierarchical manner, with smaller groups of nodes merging to form larger groups, indicating a hub-and-spoke structure [18].

2.4 Random Graphs

In a random graph, the edges between the nodes are randomly distributed [42]. If there are N nodes in the graph and the probability of an edge being formed between any two nodes is p , then the average number of edges, n , would be equal to:

$$E(n) = p \frac{N(N-1)}{2} \quad (2.15)$$

Therefore, the likelihood of having a graph G that has N nodes and n edges is calculated as:

$$P(G) = p^n (1-p)^{\frac{N(N-1)}{2} - n} \quad (2.16)$$

The progression of a random graph model is displayed in Figure 2.8. Beginning with 10 nodes, the connections between each vertex are established based on probabilities of 0.1, 0.3, and 0.7 through computer simulations. As the probability (p) increases, it is clear from the figure that the number and order of trees and cycles also increase. For instance, when $p=0.1$, there is only one cycle of 3 edges, while when $p=0.3$, there are 4 cycles of 3 edges. As p approaches 1, the graph becomes a complete graph, with all nodes fully connected, possessing $n = \frac{N(N-1)}{2}$ edges.

A remarkable property of random graphs was uncovered by Erdős and R enyi, which showed that the appearance of a graph property Q occurs in a sudden manner, similar to a phase transition. This transition occurs at a critical connection probability, $p_c(N)$, and goes from not having the property to having it [43].

Researchers can investigate the critical probability that leads to the presence of a certain type of subgraph in the graph, for example, the critical probability for the existence of a tree with 3 vertices or a subgraph with k vertices and l edge. The expected number of subgraphs present in a graph has been mathematically demonstrated to be expressed as follows [44]:

$$E(X) = C_N^k \frac{k!}{a} p^l \approx \frac{N^k p^l}{a} \quad (2.17)$$

where a represents the number of graphs that are isomorphic to each other.

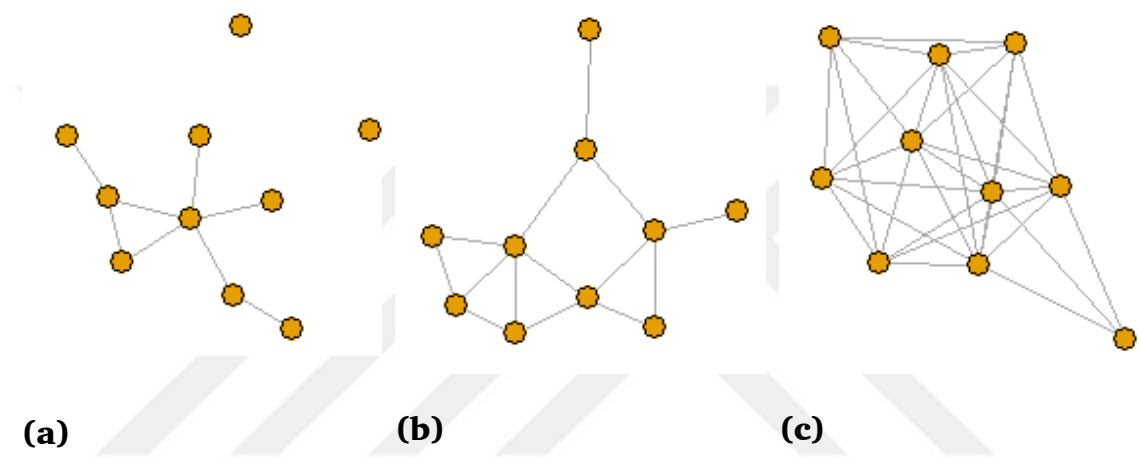


Figure 2.8 3 random graphs with 10 vertices. (a) $p=0.2$, (b) $p=0.3$ and (c) $p=0.7$

Additionally, it has been established that the critical probability for a subgraph to have at least one subgroup consisting of k vertices and l edges can be represented by the equation:

$$p_c(N) = cN^{-k/l} \quad (2.18)$$

with c being a constant. The critical probability for the occurrence of different types of subgraphs has been determined to have the following implications:

1. For a graph to contain a tree of order k , the critical probability is given by

$$p_c(N) = cN^{-\frac{k}{k-1}};$$

2. For a graph to contain a cycle of order k , the critical probability is given by

$$p_c(N) = cN^{-1};$$

3. For a graph to contain a complete subgraph of order k , the critical probability is given by $p_c(N) = cN^{-\frac{2}{k-1}}$.

This result implies that, at the critical point of supercritical probability (p_c approximately equal to $1/N$) where the average degree remains constant, the graph experiences a significant change in its physical attributes. It shifts from being composed of multiple small, disconnected clusters of trees or cycles to having a large cluster where many vertices are interlinked [45].

Another important aspect of random graphs is their degree distribution [46]. In a random graph with N nodes, each with a connection probability p , the degrees of vertices follow a binomial distribution:

$$P(k) = C_{N-1}^k p^k (1-p)^{N-1-k} \quad (2.19)$$

which gives the probability that a particular node will have k edges. When the number of vertices (N) is large, this probability distribution can be approximated using Poisson distribution:

$$P(k) \approx e^{-pN} \frac{(pN)^k}{k!} = e^{-\bar{k}} \frac{\bar{k}^k}{k!} \quad (2.20)$$

where $\bar{k}$ is the expectation value of Poisson distribution as shown in Figure 2.9.

The diameter of a random graph with N vertices and connection probability p has been found to typically be around:

$$d = \frac{\ln(N)}{\ln(pN)} = \frac{\ln(N)}{\ln(\bar{k})} \quad (2.21)$$

Additionally, research has shown that the average path lengths in real world networks and random graphs of the same size and edge density are very similar [42].

Another aspect of interest in complex networks is their level of clustering. The clustering coefficient for a random graph is expressed as:

$$C_{rand} = p = \frac{\bar{k}}{N} \quad (2.22)$$

where p is equal to the likelihood of any two vertices being connected. Despite this prediction by random graphs, real-world networks are often composed of large clusters, deviating from the scheme.

In this study, the Erdős-Renyi random graph model is used to simulate a random network [42].

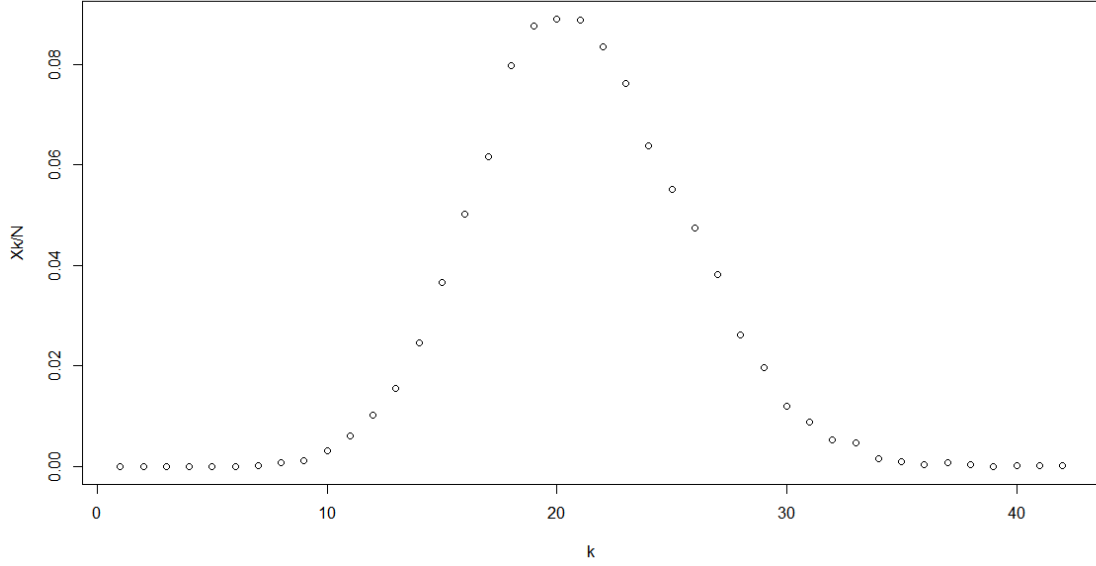


Figure 2.9 Degree distribution of a random graph with 10000 vertices and a connection probability of $p=0.002$

2.5 Small-World Networks

A small-world network is characterized by having many nodes and a relatively small average path length between them [47]. This property is believed to be common in many real-world networks [42]. The average path length of a small-world network has the form:

$$l(N, p) \sim \frac{N}{K} f(pKN^d) \quad (2.23)$$

with N as the number of nodes, p as the connection probability, K as the number of first neighbors of a node used to calculate the average path length, and d as the network's dimension. The universal scaling function, $f(u)$, for a small-world network has two different forms based on the value of u :

$$f(u) = \begin{cases} \text{constant} & \text{if } u \ll 1 \\ \frac{\ln(u)}{u} & \text{if } u \gg 1 \end{cases} \quad (2.24)$$

Small-world networks have a distinctive property of having large clustering coefficients [47]. The clustering coefficient is known to remain constant regardless of the network size. The degree distribution in these networks is comparable to random graphs where each vertex has a similar number of edges. One illustration of a social small-world network is the theory that everyone is connected by six or fewer connections, also known as "six degrees of separation" [48].

This study employs the Watts-Strogatz model to create a simulation of a small-world network [47].

2.6 Scale-free Networks

Real world networks have been observed to have degree distributions that follow a power-law distribution $P(k) \sim k^{-\gamma}$, which is different from random graphs where the degrees are approximately distributed according to a Poisson distribution [49]. Because of the distinct scale-free quality of power laws, networks that exhibit this property are referred to as scale-free networks [42]. Studying the way these networks change and grow over time is crucial in creating models that accurately represent the structure of real-life complex networks.

The procedure of creating a scale-free network model begins by understanding the real-life network's growth pattern and how connection probability between vertices is dependent on the vertex's degree. In typical real-life networks such as the internet, the network size increases constantly with the addition of new nodes and the connections grow exponentially. The preferential attachment behavior in these networks means that the probability of an edge existing between two nodes depends on the degree of the vertices, where a vertex with high connections is more likely to add another edge and increase its degree than a poorly connected vertex. These growth and preferential attachment properties serve as the foundation for the scale-free network model. An algorithm for the SF model is created by the following steps:

1. Initialize the network with m_0 vertices.

2. At each timestep, a new vertex with m (less than or equal to m_0) edges is added to the network.
3. The probability that the new vertex will be connected to vertex i is determined by its degree k_i , with the formula:

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} \quad (2.25)$$

4. After t time steps, the network will consist of $N = t + m_0$ vertices and mt edges.

Through simulation, it can be demonstrated that the resulting network is scale-free. Figure 2.10 demonstrates a network that was created through the above process. The network consists of some vertices with a high degree and many vertices with a low degree. This observation suggests a power law distribution in the degree of the network, which is shown in Figure 2.11 representing the probability of a vertex to have k edges. The aim is to identify the scaling exponent γ of this distribution. Real networks have been found to have varying γ values ranging from 1 to 3.

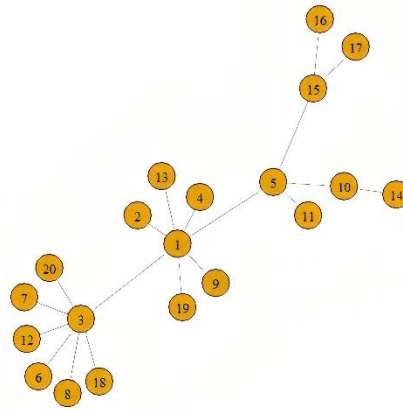


Figure 2.10 A scale-free network with 20 vertices generated by simulation

It was discovered through empirical research that the average path length in a scale-free network follows a logarithmic function with regards to the size of the network, N [42].

$$l = A \log(N - B) + C \quad (2.26)$$

The solution for the average path length of a scale-free network when it approaches infinity has been found to be in the form of a logarithmic function of the network size N , which is expressed as:

$$l \approx \frac{\log N}{\log(\log(N))} \quad (2.27)$$

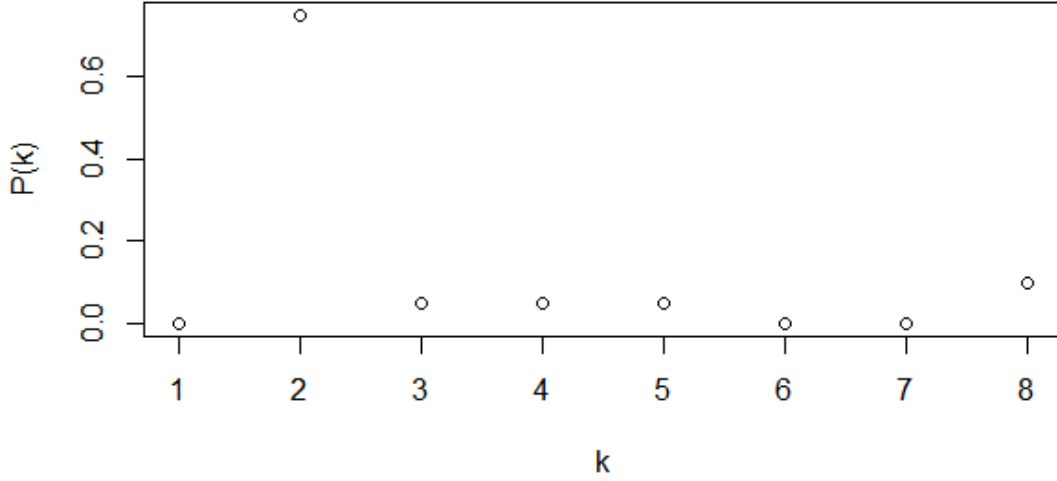


Figure 2.11 Degree distribution of the scale-free network in figure 2.10

As the size of a scale-free network increases, its clustering coefficient tends to decrease and approach zero. This is typically observed in small scale-free networks where the clustering coefficients have large values [50].

2.7 Community Detection in Complex Networks

The method of community detection by maximizing modularity is based on comparing the edge density of a random graph to that of suspected communities within the network being analyzed. The idea is that nodes in a random graph form links randomly, so they are not expected to form communities. The calculation of modularity allows us to measure the quality of each community and compare different community structures. Modularity, which ranges from -0.5 to 1, is determined by subtracting the expected number of edges in a random network of the same size from the number of actual edges in a community. If the result is positive, it suggests the presence of a potential community. The higher the

modularity score, the better the community structure. Thus, maximizing modularity in a network leads to the identification of optimal communities [51].

The process of maximizing modularity in a network can be achieved by using various algorithms. One widely used approach is the Louvain algorithm [52], which is efficient and identifies separate, non-overlapping communities. However, a limitation of this method is that it may result in finding disconnected communities. To address this flaw, a new version of the Louvain algorithm, the Leiden algorithm, has been developed and was utilized in this paper to uncover the community structure in the EAN. The Leiden algorithm ensures that the communities found are well connected and that nodes are optimally assigned to their respective communities [53]. The modularity function, represented by equation:

$$H = \frac{1}{2} \sum_c (e_c - \gamma \frac{K_c^2}{2m}) \quad (2.28)$$

is a measure of the difference between the actual number of edges within a community (e_c) and the expected number of edges in a random network of the same size. The value is calculated by summing the difference between these two values for all communities, with a higher score indicating a better community structure. The value of K_c is the sum of the degrees of the nodes in community c , while m is the total number of edges in the network. The resolution parameter γ is used to set a threshold for the density of connections within the communities.

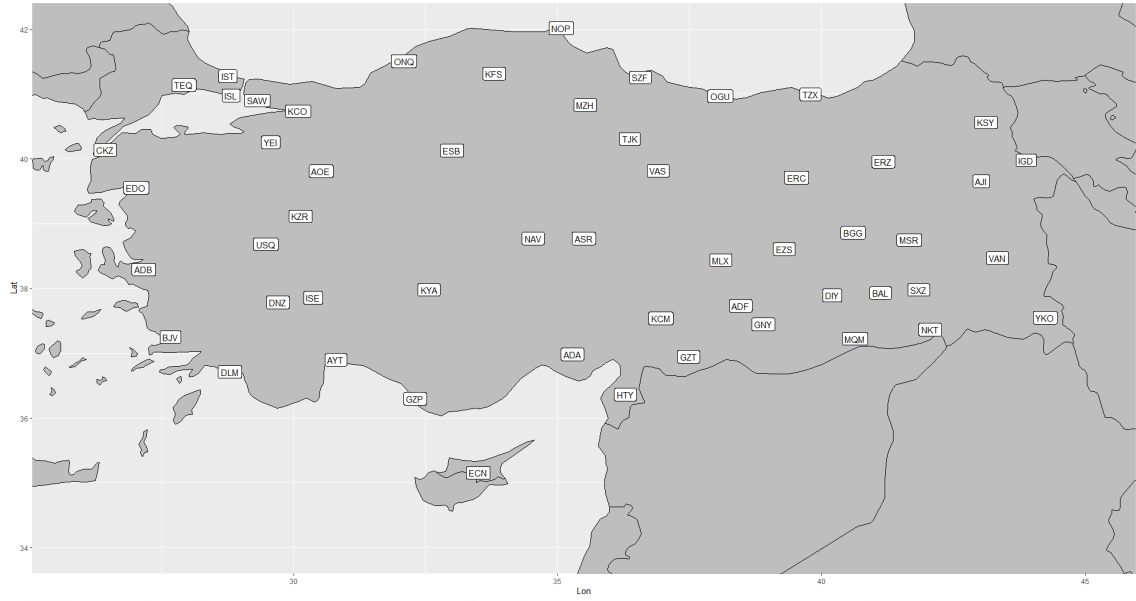
ANALYSIS OF THE TURKISH AIRPORT NETWORK

3.1 Turkish Airport Network

The Turkish air transport sector has experienced rapid growth since the government deregulated the civil aviation industry in 2003, which subsequently intensified competition among airline companies. As a result of increased competition, fares for air travel have decreased, which has stimulated the demand for air transport services in Turkey. Furthermore, the deregulation has led to a continuous expansion of Turkey's domestic route network [54].

The airport network of Turkey comprises 54 airports as shown in Map 3.1, which are regarded as the vertices of the network, and aircraft movements between the vertices serve as edges. Owing to the Covid-19 pandemic, a small number of airports were not operational during the sampling period. Since flights are typically scheduled on a week basis, three successive weekly flight data were selected for analysis during the months of July, August, September, and October to track the network's evolution. A summary of the samples is provided in Table 3.1.

Table 3.2 and 3.3 illustrate the airports that rank the highest in terms of domestic and international flight frequencies in the three samples taken. The top three airports, namely Istanbul Sabiha Gökçen (SAW), Istanbul (IST), and İzmir Adnan Menderes (ADB), are consistent across the rankings, but there are variations in the flight volumes. Specifically, SAW has the highest number of domestic flights, while IST has the most international flights. This can be attributed to the fact that Turkish Airlines, which has numerous international destinations, operates primarily out of IST. The traffic between Turkish airports and airports in Germany, Ukraine, Russia, and the United Kingdom are the busiest in terms of aircraft movements.



Map 3.1 Operational airports in Turkey in 2020 shown with their IATA codes

Table 3.1 Sampling summary

	1st Sample	2nd Sample	3rd Sample
Sampling Duration	29.07.2020 - 05.08.2020 (7 days)	26.08.2020 - 01.09.2020 (7 days)	25.09.2020 - 01.10.2020 (7 days)
Domestic Network Size	47 airports	46 airports	47 airports
Overall Network Size	232 airports	228 airports	245 airports
Domestic Network Routes	150 unique routes	147 unique routes	145 unique routes
Overall Network Routes	614 unique routes	629 unique routes	617 unique routes
Domestic Network Flights	4320 flights	4617 flights	4194 flights
Overall Network Flights	7832 flights	10161 flights	9662 flights

Table 3.2 Airports with the highest aircraft movements

1st Sample		2nd Sample		3rd Sample	
Airport	Aircraft movements	Airport	Aircraft movements	Airport	Aircraft movements
SAW	1954	SAW	2270	SAW	2192
IST	1482	IST	1378	IST	1180
ADB	656	ADB	741	ADB	673
ESB	606	ESB	669	ESB	608
AYT	493	AYT	545	AYT	532
ADA	383	ADA	389	ADA	342
TZX	334	TZX	371	TZX	306
BJV	279	BJV	328	BJV	288
DLM	221	DLM	242	DLM	232
GZT	184	GZT	195	GZT	194

Table 3.3 Airports with the highest domestic aircraft movements

1st Sample		2nd Sample		3rd Sample	
Airport	Aircraft movements	Airport	Aircraft movements	Airport	Aircraft movements
IST	1215	IST	1700	IST	1848
SAW	594	AYT	1520	AYT	1508
AYT	554	SAW	788	SAW	852
ADB	327	DLM	378	DLM	398
DUS	230	ADB	358	SVO	264
ISL	214	KBP	248	ISL	238
FRA	163	SVO	242	KBP	236
CGN	163	DUS	234	ADB	226
STR	153	STN	196	VKO	224
ESB	150	ISL	196	BJV	182

Table 3.4 and 3.5 illustrate the routes that have experienced the highest volume of aircraft traffic. The route between Sabiha Gökçen and Antalya airports is the busiest domestic route in all three samples. The prominence of routes to and from SAW is attributed to the high-frequency services offered by Pegasus Airlines,

which is based at this airport. In terms of international routes, AYT is the top airport in the 2nd and 3rd samples, largely because of the calming of pandemic air travel limitations between airports in Russia. The busiest airport pair in the 1st sample is the route between Sabiha Gökçen and Stansted but it falls to fifth place in the 2nd and 3rd samples.

Table 3.4 Top 5 busiest domestic routes in terms of aircraft traffic

Sample 1			
	Airport pair	Aircraft movements	Share of traffic (%)
1	SAW - AYT	206	4.77
2	SAW - ADB	181	4.19
3	SAW - BJV	143	3.31
4	SAW - TZX	143	3.31
5	SAW - DLM	132	3.06
Sample 2			
	Airport pair	Aircraft movements	Share of traffic (%)
1	SAW - AYT	243	5.26
2	SAW - ADB	219	4.74
3	SAW - BJV	190	4.12
4	SAW - TZX	184	3.99
5	SAW - DLM	157	3.40
Sample 3			
	Airport pair	Aircraft movements	Share of traffic (%)
1	SAW - AYT	267	6.37
2	SAW - ADB	206	4.91
3	SAW - TZX	178	4.24
4	SAW - BJV	173	4.12
5	SAW - DLM	152	3.62

Table 3.5 Top 5 busiest international routes in terms of aircraft traffic

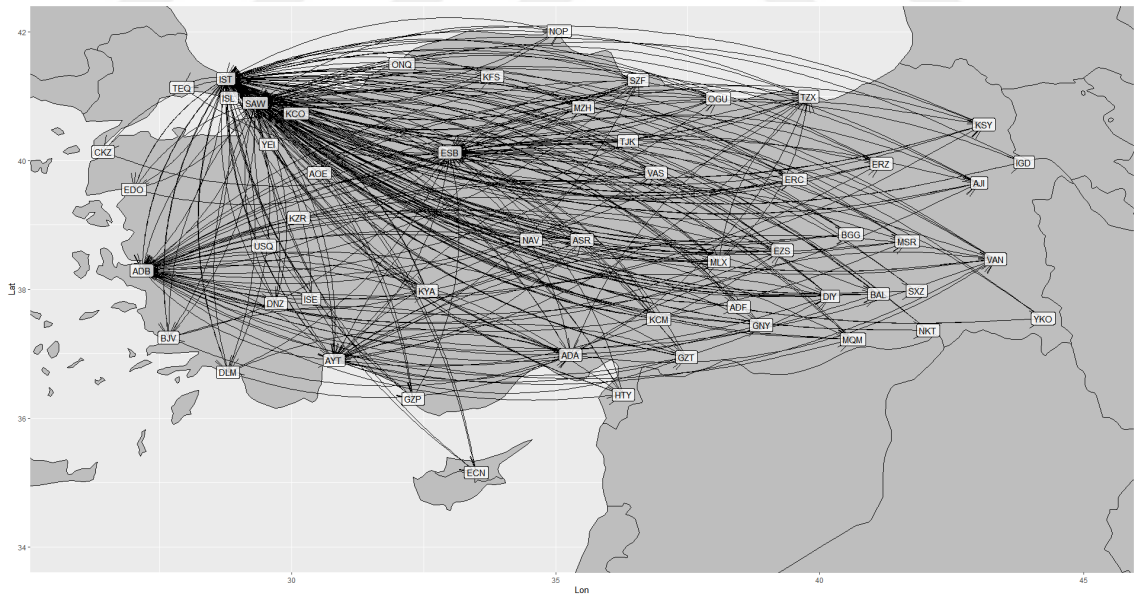
Sample 1			
	Airport pair	Aircraft movements	Share of traffic (%)
1	SAW - STN	67	1.91
2	AYT - KBP	60	1.71
3	IST - CDG	48	1.37
4	SAW - DUS	46	1.31
5	SAW - CGN	42	1.20
Sample 2			
	Airport pair	Aircraft movements	Share of traffic (%)
1	AYT - KBP	170	3.07
2	AYT - SVO	146	2.63
3	AYT - DME	98	1.77
4	AYT - VKO	94	1.70
5	SAW - STN	94	1.70
Sample 3			
	Airport pair	Aircraft movements	Share of traffic (%)
1	AYT <-> KBP	152	2.78
2	AYT <-> SVO	146	2.67
3	AYT <-> VKO	112	2.05
4	AYT <-> ROV	94	1.72
5	SAW <-> STN	80	1.46

3.2 Topological Analysis of the Domestic Network

In all three samples, the number of active airports in Turkey's domestic network, as illustrated in Map 3.2, remains relatively constant with the number of vertices ranging from 46-47 and the number of edges ranging from 145-150 (Table 3.1). However, the number of flights shows fluctuations, with the highest number of

flights occurring in the second sample, possibly because of seasonality in airline passenger demand. The mean value of the edge density of the three samples is 0.138, indicating that approximately 14% of all possible routes were served by airlines.

Throughout this study, various centrality measures were used to identify the significant and influential airports in Turkey's domestic network. The reciprocity values of all three samples were nearly equal to 1, indicating that the network edges can be considered as bidirectional. The average degree of the three samples was approximately 6, implying that, on average, each airport in Turkey had six domestic routes. The diameter of the networks ranged from 3 to 4, meaning that the farthest airports were reachable by taking three to four flights. This diameter was found to be high due to ISL airport, which had domestic arrivals but not departures. However, when this airport was removed from the network, the diameter of all samples decreased to 3. The farthest airports in the first and second samples were SAW and ISE, respectively, whereas in the third sample, they were CKZ and ASR.



Map 3.2 Domestic routes of sample 3

We initiated the analysis by examining the degree centrality. Table 3.6 presents the top 10 vertices in terms of degree centrality for the unweighted networks. This table displays the number of unique routes originating from each vertex. Istanbul

Airport ranked first in all three samples, Sabiha Gökçen was the second and Ankara Esenboğa was the third. Although Sabiha Gökçen had higher aircraft traffic, Istanbul Airport had a wider range of destinations. The degree rankings were similar across all three samples. In addition, the degree-rank distribution exhibited a similar pattern across all three samples and followed a power-law $P(k) \sim k^{-\gamma}$, with γ values ranging from 0.947 to 0.981 (see Figure 3.1). The top five airports dominated the flight traffic, while other airports had 10 or less degrees. The cumulative degree distribution of the three networks follows power laws, as well, where γ values range from 1.071 to 1.106 (see Figure 3.2). The degree structure of the network remained relatively stable throughout the sampling process.

Table 3.6 Degree rankings of domestic network

1st Sample		2nd Sample		3rd Sample	
Node	k	Node	k	Node	k
IST	42	IST	42	IST	42
SAW	35	SAW	35	ESB	35
ESB	33	ESB	35	SAW	35
ADB	25	ADB	25	ADB	24
AYT	15	AYT	13	AYT	13
ADA	12	ADA	10	ADA	9
ASR	6	TZX	6	TZX	7
TZX	6	VAN	6	MLX	6
ECN	6	ECN	5	VAN	6
VAN	6	ASR	5	DIY	5

In the analysis of the domestic network, an additional measure of centrality was taken into account, which is eigenvector centrality. This measure assigns higher priority to vertices that are connected to other high degree vertices. Table 3.7 displays the top 10 normalized eigenvector centralities for all 3 samples, with IST appearing at the top in all 3 samples. ESB is ranked the second in sample 1 and 2,

surpassing SAW, indicating that ESB offer more routes to high degree airports than compared to SAW.

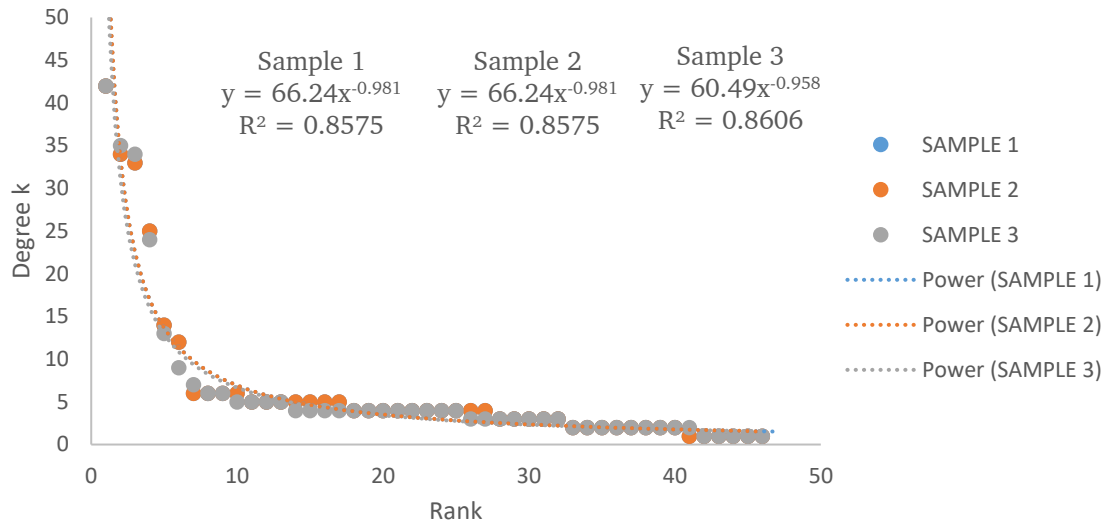


Figure 3.1 Degree distributions of samples

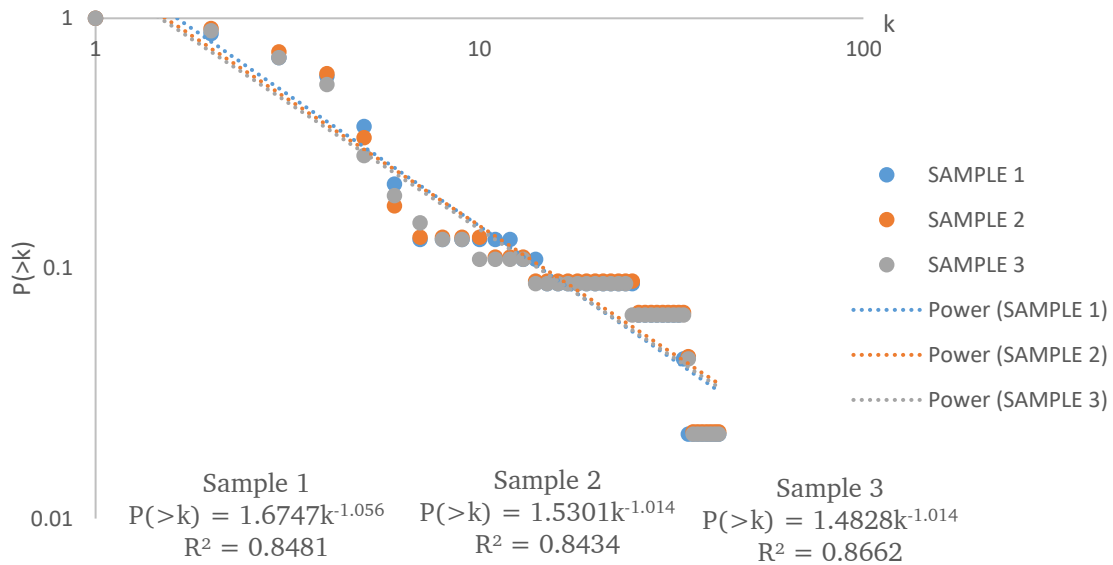


Figure 3.2 Cumulative degree distributions of samples shown on log-log scale

Next, we examine the betweenness centrality of airports in Turkey's domestic airport network, which measures the frequency of an airport being on the shortest path between any two airports in the network. The results are shown in Table 3.8, where Istanbul Airport has the highest betweenness centrality, followed by Ankara Esenboğa Airport and Istanbul Sabiha Gökçen Airport. These airports also have high degrees, making them hub airports in the network, and approximately 90%

of the domestic aircraft movements are centered around them. This contrasts with the findings of Guimera et al. [9] for the global air network, in which high betweenness airports are not necessarily highly connected. Betweenness of the networks was also observed to conform to power laws (Figure 3.3).

Table 3.7 Eigenvector rankings of domestic network

Sample 1		Sample 2		Sample 3	
Airport	Eigenvector	Airport	Eigenvector	Airport	Eigenvector
IST	1	IST	1	IST	1
ESB	0.944685383	ESB	0.965148286	ESB	0.961190669
SAW	0.933715149	SAW	0.922075517	SAW	0.920500523
ADB	0.839203072	ADB	0.828205055	ADB	0.812505114
AYT	0.612622318	AYT	0.565222983	AYT	0.573133254
ADA	0.520997305	ADA	0.480229307	ADA	0.459497861
TZX	0.390453715	TZX	0.385197926	TZX	0.417922866
VAN	0.390453715	VAN	0.385197926	VAN	0.386584125
ECN	0.390453715	GZT	0.346343074	MLX	0.38318391
DIY	0.348520923	DIY	0.346343074	GZT	0.34900404

Table 3.8 Normalized betweenness rankings of domestic network

SAMPLE 1		SAMPLE 2		SAMPLE 3	
Airport	Normalized Betweenness	Airport	Normalized Betweenness	Airport	Normalized Betweenness
IST	0.427133655	IST	0.406582492	IST	0.405169082
SAW	0.20394525	ESB	0.2195306	ESB	0.2447684
ESB	0.187984276	SAW	0.217188552	SAW	0.219661836
ADB	0.059642891	ADB	0.05976629	ADB	0.053303022
ADA	0.048129203	AYT	0.007981779	AYT	0.007650848
AYT	0.023733068	ADA	0.003789859	ADA	0.002562281
ASR	0.000866724	ASR	0.000490196	TZX	0.000388368
AJI	0.000142086	ECN	0.000237671	AJI	0.000227337
HTY	0.000142086	BJV	0.000237671	HTY	0.000227337
BAL	0.000142086	DLM	0.000237671	BAL	0.000227337

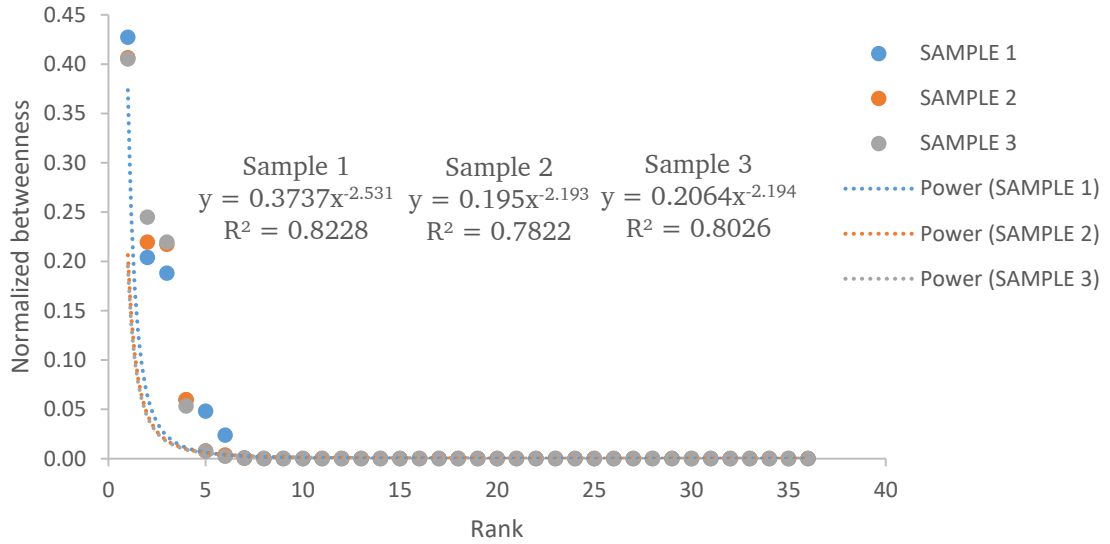


Figure 3.3 Normalized betweenness-rank distributions of samples

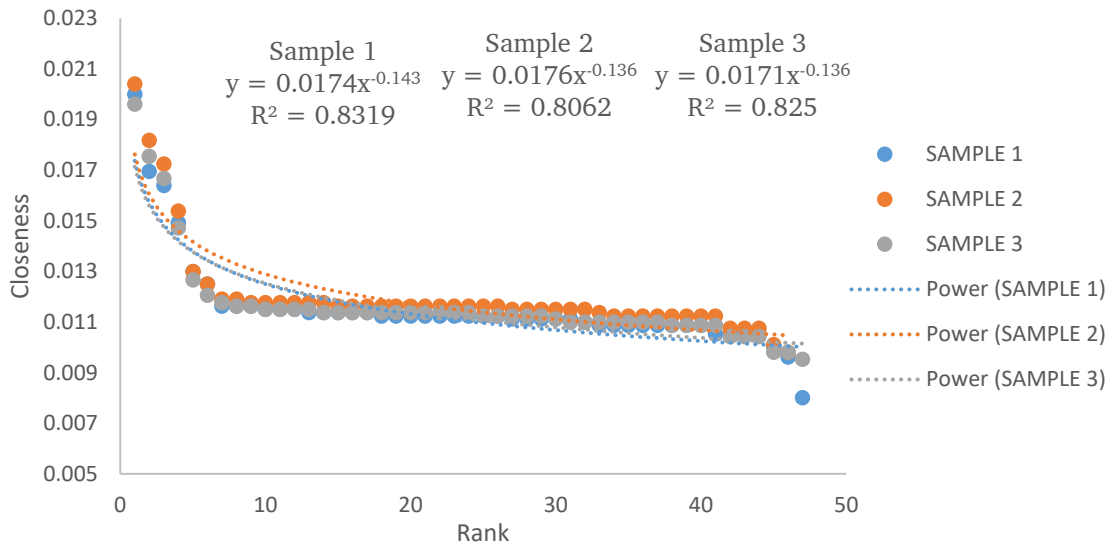
The last centrality measure we investigate is closeness, which provides insight into closeness of each airport to other airports in the network in terms of number of flights. This network measure is determined by taking the reciprocal value of the average shortest path between a particular node and all other nodes in the network. A higher closeness centrality indicates that the airport can be reached with fewer connecting flights. Table 3.9 presents the airports ranked in terms of their closeness centrality, and we find that Istanbul Airport is the leader airport in the three networks, Esenboğa the second and Sabiha Gökçen the third. The top three airports, in essence, have the shortest average topological distance to every other airport in the network. Finally, closeness centrality distribution of the samples were found to follow power laws like the other centrality measures (Figure 3.4).

3.3 Connectivity Performance

The diameter of TAN is equal to 4 in 2 samples because Istanbul Atatürk Airport was operating international cargo operations only at those times and domestic aircraft traffic was not very frequent. Thus, to accurately analyze the network for connectivity, this airport was removed from the network. Consequently, the diameters of sample 2 and sample 3 immediately drop down to 3, hence the average diameter of TAN becomes 3. Therefore, a passenger must take, on average, 3 flights in order to travel between any two airports within TAN.

Table 3.9 Closeness rankings of domestic network

Sample 1		Sample 2		Sample 3	
Airport	Closeness	Airport	Closeness	Airport	Closeness
IST	0.02	IST	0.020408163	IST	0.0196078
ESB	0.016949153	ESB	0.018181818	ESB	0.0175439
SAW	0.016393443	SAW	0.017241379	SAW	0.0166667
ADB	0.014925373	ADB	0.015384615	ADB	0.0147059
AYT	0.012987013	AYT	0.012987013	AYT	0.0126582
ADA	0.0125	ADA	0.0125	ADA	0.0120482
TZX	0.011627907	TZX	0.011904762	TZX	0.0117647
ECN	0.011627907	VAN	0.011904762	MLX	0.0116279
VAN	0.011627907	ECN	0.011764706	VAN	0.0116279
BJV	0.011494253	DIY	0.011764706	DIY	0.0114943

**Figure 3.4** Closeness distribution of three samples (all follow power laws)

The characteristic path length is a critical metric to assess the connectivity of an airport network. It is defined as the mean value of flight numbers needed to connect any random airport pairs of the network. In TAN, the average path length ranged from 1.86-1.90, indicating that passengers could travel to any airport in the network with no more than two connecting flights on average. However, to travel between the farthest two airports in the first and second samples (SAW-ISE & CKZ-ASR), passengers needed to make at least three connecting flights.

The clustering coefficient or transitivity is a metric that indicates the degree of clustering in a network, which is the proportion of existing triangles to all possible triangles. A higher clustering coefficient suggests greater interconnectedness within the network. In the analyzed samples, the global clustering coefficients ranged from 0.20-0.22, which is significantly lower than the clustering coefficient of 0.47 observed in Turkey's domestic network in 2014 [26]. This suggests that the interconnectivity of the network has decreased significantly over the past six years, possibly due to the fast addition of new airports without a corresponding increase in the number of flights. The ongoing Covid-19 pandemic may also have contributed to this decrease in interconnectivity by preventing the opening of new routes.

Finally, we discuss local connectivity issues such as cliques and local clustering coefficient in Turkey's domestic airport network. Fully connected subnetworks consisting of 7 airports were observed in all samples (IST, ESB, ADB, AYT, ADA, VAN, TZX). These cliques represent airports that are highly interconnected within the domestic network. In addition to cliques, the local clustering coefficient is a measure of the likelihood that two of an airport's neighbors are connected to each other. The highest degree airport, IST, had a local clustering coefficient of 8.1%, while the second highest degree airport, SAW, had a local clustering coefficient of 11.2%. Interestingly, airports that offer 5 to 6 connections had very high clustering coefficients ranging from 0.8 to 0.9, suggesting that airports with relatively lower degrees tend to connect to already well-connected airports.

The degree of assortativity in a network measures how nodes with similar attributes tend to be linked to each other. If nodes with the same attribute are more likely to be connected, then the network is considered assortative, whereas if they are less likely to be connected, the network is considered disassortative. In the case of Turkey's domestic airport network, the assortativity was negative in all three samples, with an average value of -0.78. This means that poorly connected airports tend to attach to major hubs in the network. In TAN, the top five high degree airports connect on average to 30 airports, while the remainder only offer 3-4 unique connections. This shows that the network follows a hub-and-spoke structure, where the air traffic coming from poorly connected spoke airports flows

through major hubs. The pandemic has affected the aviation industry, and it is likely that the number of connections offered by these airports will increase significantly after the pandemic ends.

Another degree correlation that we examine is the relationship between the clustering coefficients and degrees of vertices. We observe that the clustering coefficient decreases linearly as the degree of vertices increases, as shown in Figure 3.5. In general, the less connected airports in the network tend to have higher clustering coefficients, which suggests that they are mainly connected to major hubs. This pattern is observed in most cases, with only a few exceptions. Some of these less connected airports, such as BGG, EDO, ERC, GZP, KCM, KYA, and VAS, may serve as stopover locations for low-cost carriers that offer point-to-point services. On the other hand, as the degree of vertices in the network increases beyond the average value of 6, there is a significant decrease in the clustering coefficient. We observe a similar pattern in all three samples. This structure is in line with the evolution of the network in the past two decades, during which the two biggest carriers of Turkey utilized Istanbul as their hub location, leading to the formation of a domestic air traffic network centered around this city. Airports with degrees below the network average tend to have high clustering coefficients, indicating a higher tendency to form clusters. New airports added to the network tend to connect more frequently with major hubs than less connected airports, with these hubs acting as gateways between different clusters in the network.

The Pearson correlation coefficient between the average degree and betweenness centrality for the network samples was found to be 0.912988, indicating that airports with high degrees also tend to connect different regions of the network, acting as superhubs. Another noteworthy observation is that, although there is no explicit mathematical relationship between degree and closeness centrality, there exists an almost perfect linear relationship between the two measures, as demonstrated in Figure 3.6. This relationship was consistent across all network samples.

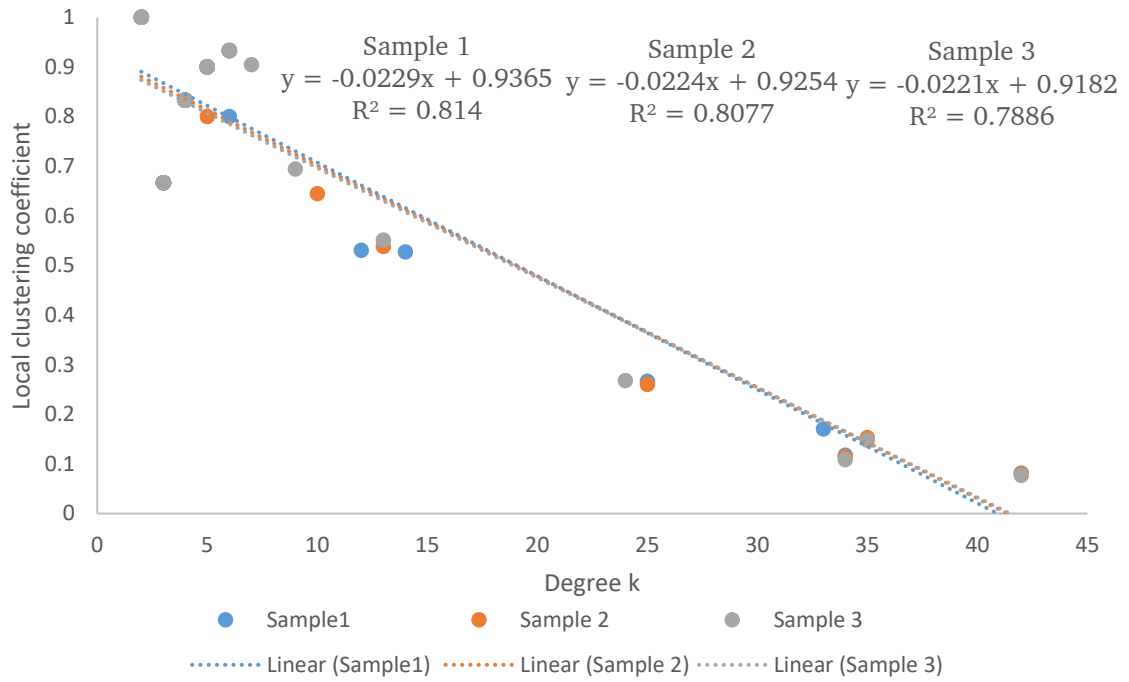


Figure 3.5 Relation between the degree and clustering coefficient of the nodes

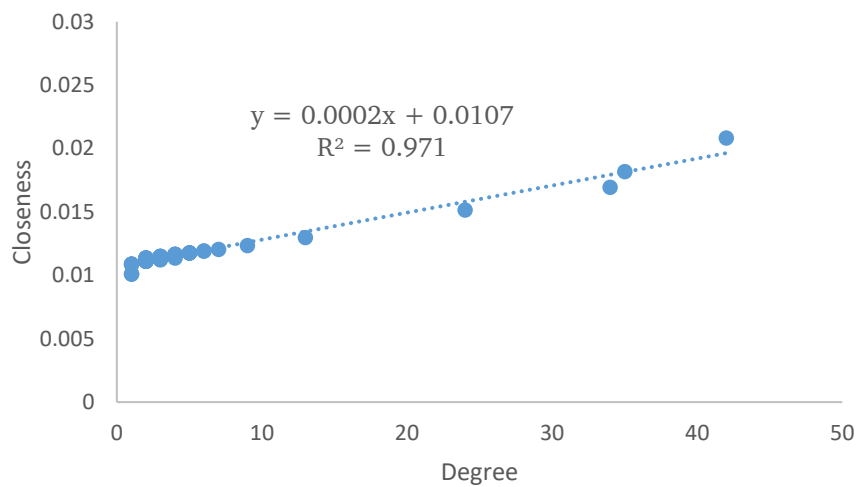


Figure 3.6 Relation between the degree and closeness centrality of the nodes of sample 3

3.4 Comparison of the Domestic Network with Simulated Models

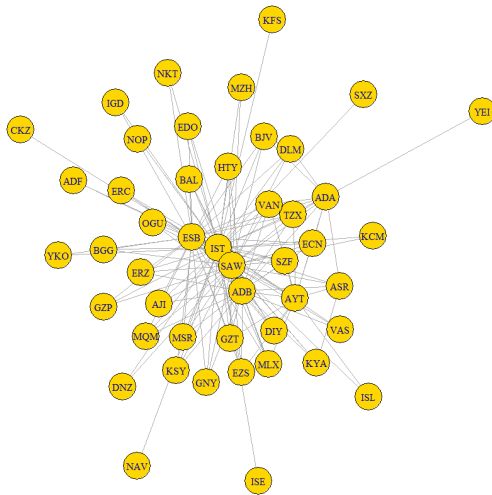
To assess how the characteristics of the domestic air network compare with different network models, simulations of three different models were conducted.

These models were chosen to have similar network sizes, average degrees, and edge densities to the domestic network. The three models used were a random network, a small-world network, and a scale-free network. Each of these models had 47 vertices and edges varying between 135-147, and their characteristics were compared with those of the domestic network. Table 3.10 provides the details on the parameters of each model and the domestic network. A visual comparison is shown in Figure 3.7.

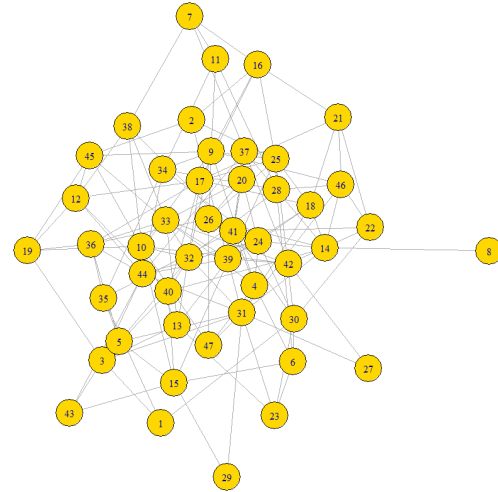
Table 3.10 Parameters of domestic network and simulated models

Network Metrics	Sample Averages	Random network	Small-world network	Scale-free network
number of vertices	47	47	47	47
number of edges	147	147	141	135
average degree	6	6.255319	6	5.744681
edge density	0.138308	0.1359852	0.1304348	0.1359852
reciprocity	0.978251	1	1	1
diameter	3	4.1549	4.4447	2.4934
average path length	1.883707	2.256487	2.474611	1.878423
transitivity (clustering coefficient)	0.207291	0.134323	0.2839487	0.1469492
Degree-degree correlation (assortativity)	-0.776052	-0.04642041	-0.02127617	-0.7542199
Degree-betweenness pearson correlation	0.912988	0.8811249	0.7613222	0.9850329

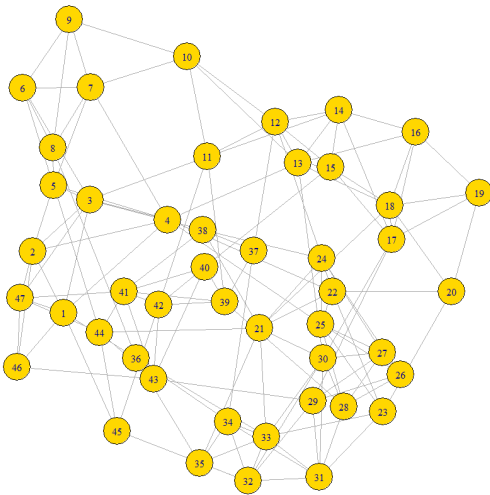
The domestic airport network of Turkey was contrasted with a random network model of equal size and edge count, using the Erdős-Renyi model. The findings revealed that the diameter of the random network was a bit greater than the domestic network, with an average of 4.1549. The average path length of the random network model, which was compared to the Turkish domestic airport network, was slightly larger, suggesting that it has lower connectivity performance than the domestic network. Specifically, the mean value of the diameter of the random network was 4.1549, which makes it somewhat higher than the diameter of the domestic network. In addition, the random network had a significantly



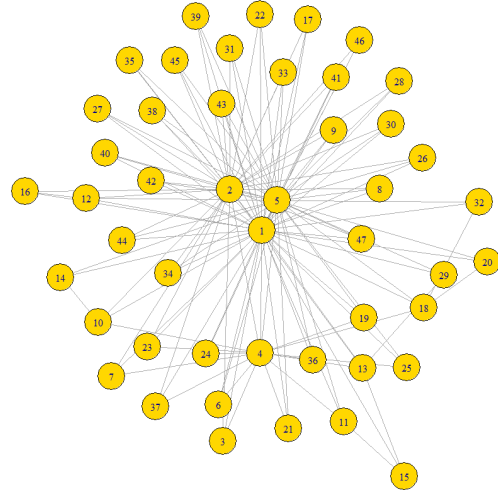
a) Domestic network of sample 1



b) Random network



c) Small-world network



d) Scale-free network

Figure 3.7 Visual comparison of the domestic network with simulated models

lower average clustering coefficient of 0.134323, indicating reduced interconnectivity than the domestic network. The random network had an assortativity value close to zero, indicating that there was no tendency for vertices to attach preferentially to others with similar degrees. This is because, in the random network, all vertices are equally likely to be connected to one another.

The next simulation is for a small-world network using the Watts-Strogatz model. After running the simulation 10,000 times, the results show that the average diameter of the small-world network is 4.4447, which is higher compared to the

domestic network. The average path length is higher, as well, indicating that the small-world network has comparatively reduced connectivity performance. However, the interconnectivity within the small-world network is moderately better as the clustering coefficient is larger than that of the domestic network. Like the random network, the small-world network has close to zero assortativity, indicating a lack of preferential attachment.

We used the Barabasi-Albert model to simulate a scale-free network with the same number of vertices and edges as the domestic network. The simulations were run 10000 times, and the resulting average diameter of the scale-free network was found to be 2.4934, which was smaller than that of the domestic network. The average path lengths of both networks were found to be very similar. However, scale-free network's clustering coefficient was lower compared to the domestic network. The degree of assortativity in the scale-free network was -0.7542199, like the domestic network, which proves the presence of preferential attachment in making connections between the nodes. This was also observed in the domestic network. Overall, among the simulated networks, the scale-free model appears as the most similar to domestic TAN.

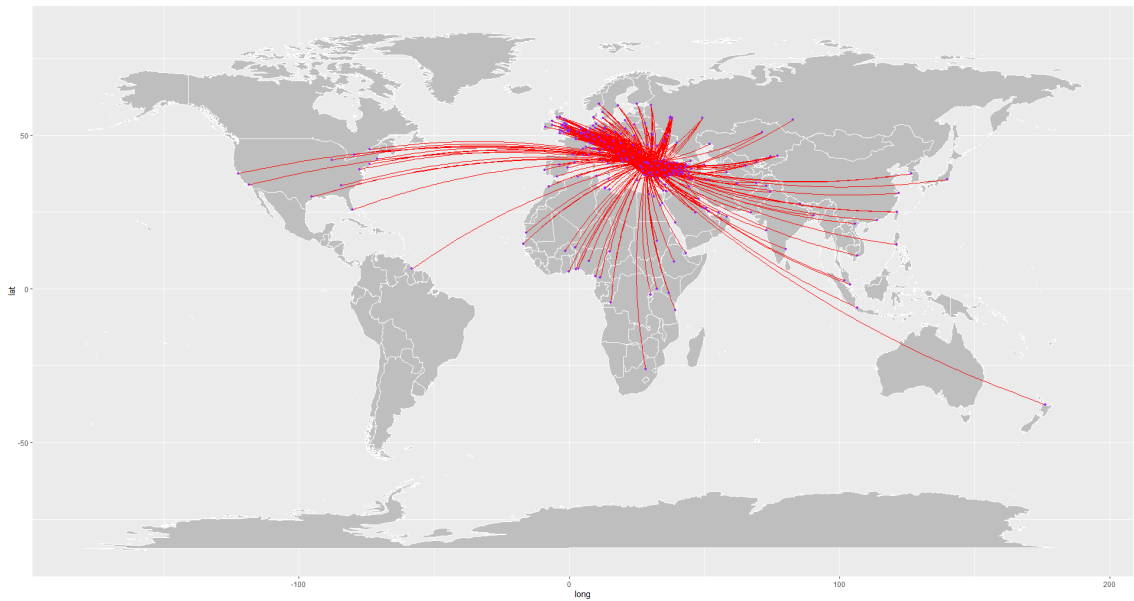
3.5 Topology of the Overall Network

In this section, we are analyzing the flight networks of Turkish airports, including international routes as shown in Map 3.3. We note that the reciprocities of the three samples are roughly equal to 1 (Table 3.11) , indicating that the network is bidirectional and can be treated as undirected. The diameter of the network is 5 in the first two samples and 6 in the third, meaning that passengers can travel between any two points in the network in a maximum of 5 or 6 flights. The nodes having the largest topological path length in all the samples are Zonguldak and Ostend Bruges airports.

We also calculated the average path length, which ranges from 2.29 to 2.41 with an overall average of 2.35. This relatively small value suggests that the network has a small-world nature, meaning that it has a high level of connectivity. However, we note that the degree of assortativity of the overall networks is significantly different from the domestic network, with an average value of -0.53.

This indicates that while the network has a hub-and-spoke characteristic, there is a higher number of hubs than the domestic network and these hubs also connect to each other increasing hub-to-hub traffic.

In contrast to the domestic network, the overall network has the highest number of nodes in the third sampling period (245 airports). This increase in size could be attributed to the easing of COVID-19 border restrictions in various countries. On the other hand, like the domestic network, aircraft traffic and distinct routes reaches its highest point in the second sample, likely due to a seasonal increase in air travel demand for holiday making routes. The edge densities of all three samples are similar, on average, only 2.3% of all possible routes were served.



Map 3.3 All routes flown in sample 3

The average degree of the samples ranges from 5.04 to 5.52, with a mean value of 5.29. The behavior of the degree distributions indicates that there are a small number of nodes with many connections, while majority of the nodes are poorly connected. This pattern can be described by power laws having where γ values range from 0.952 to 1.039. The cumulative degree distributions also follow power laws. Istanbul Airport has the highest number of routes, Sabiha Gökçen the second and Antalya Airport the third. Despite only operating cargo flights after Istanbul Airport opened, Istanbul Atatürk Airport is interestingly a high connected airport

because it served as a hub for Turkish Cargo. Turkish Cargo has been improving its fleet and routes, making it a major air cargo operator globally.

Table 3.11 Overall network measures

Network Metrics	First Sample	Second Sample	Third Sample	Averages
nodes	232	228	245	235
edges	614	629	617	620
Aircraft movements	7832	10161	9662	9218
D	5	5	6	5
L	2.406889	2.293261	2.354031	2.351394
reciprocity	0.9581205	0.991929	0.989236	0.979761933
edge density	0.02291387	0.024306	0.020642	0.022620863
<k>	5.293103	5.517544	5.036735	5.282460667
Degree of assortativity	-0.5317066	-0.5258	-0.5425	-0.533336233

4.1 Sample Information

In this chapter, we evaluate the performance of airports in terms of centrality and connectivity and investigate the hub and community structure of the European Airport Network (EAN). Our analysis focuses on the topology of the network in 2019, as this was the latest year before the Covid-19 pandemic significantly impacted the industry. To model the network, we use a weighted and undirected graph comprising 423 airports as nodes and 4477 routes as edges.

In the EAN, only 5% of all possible routes were flown, yet it still managed to handle a total of 1,715,483,190 passengers. The average number of passengers handled per airport was 4,055,435, while the median value was 898,946, indicating that the strength distribution of the network is skewed to the right and decays exponentially. On average, 191,585 passengers were handled on each route, but the histogram of route traffic peaks at around 107,000 (Figure 4.1 (a)). The average route distance was 1,326 km, with 50% of routes being shorter than 1,173 km, indicating that the EAN is mostly a short-haul flight network (Figure 4.1 (b)). The distance and the weight distributions of the network were also skewed to the right decaying logarithmically.

One common method of ranking airports and routes is by the number of passengers handled. Amsterdam Schiphol airport (AMS) is the top-ranking airport in terms of passenger traffic, handling over 51 million passengers within Europe. Among the top 10 routes (as shown in Table 4.1), all but one route (Athens - Larnaca) are domestic and each handles over 2 million passengers in both directions.

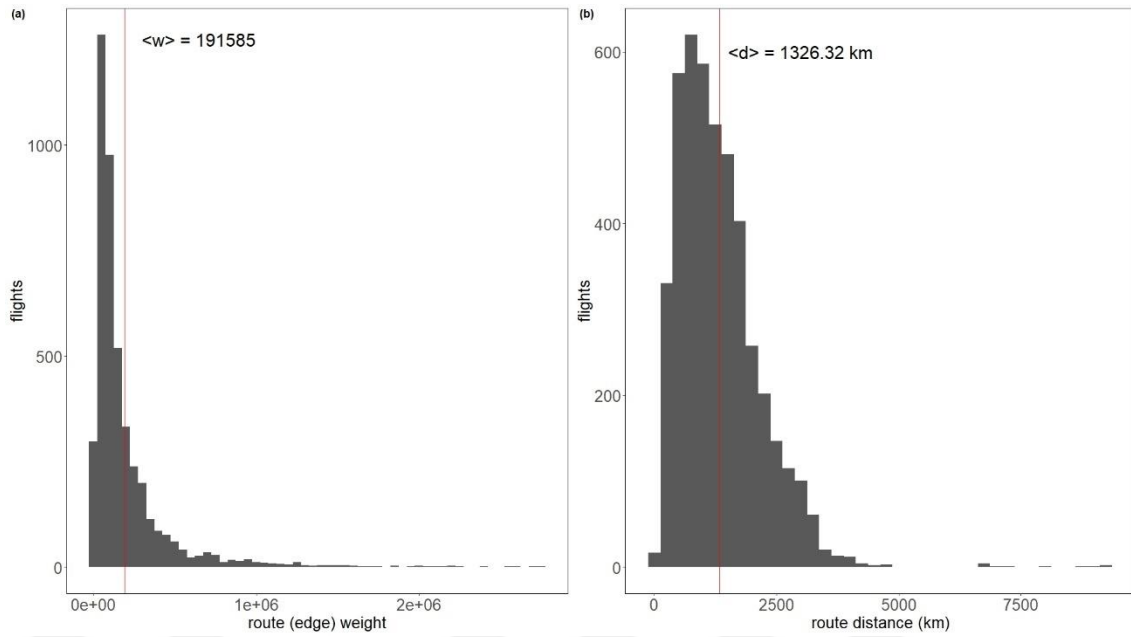


Figure 4.1 Distributions of edge weights and route distances in the EAN. The vertical axes represent the counts of the number of flights for both figures. In (a), the average number of passengers handled on the routes is 191,585, but the histogram of route traffic peaks around 107,000, (b) indicates that the average route distance is 1,326 km, but 50% of the routes are shorter than 1,173 km

4.2 Connectivity and Centrality Analysis

The level of ease with which a passenger can travel between airports in an air transport network is referred to as connectivity. The simplest way to measure connectivity is through the characteristic path length, which is 2.55 in the EAN. This implies that, on average, a passenger needs to take 2-3 flights to reach any random airport in the network. The average clustering coefficient is another measure of interconnectivity, and it is 0.47, indicating that 47% of an airport's neighbors in the EAN are connected to each other. The EAN's weighted clustering coefficient is 0.52, indicating that edges with larger weights are more likely to form triangular routes.

The centrality of airports is an important factor in understanding the structure of air transport networks. One way to measure centrality is by using the degree metric, which counts the number of direct connections an airport has with other airports in the network. In the EAN, the average degree of airports is 21.17, with a power-law degree distribution, indicating that only a few airports have a large

number of connections (Figure 4.2 (a)). Amsterdam Schiphol (AMS) is the airport with the highest degree centrality, and it is also the top airport in terms of eigenvector and weighted betweenness centralities, making it a key player in the network's connectivity. London Stansted (STN) has the second-highest degree centrality, but its ranking in other centrality measures is lower due to its connections with mostly low-degree airports (Table 4.2). The distribution of airport degrees shows that 25% of airports have 3 or fewer connections, 50% have 9 or fewer connections, and only the top 25% of airports have more than 27 connections.

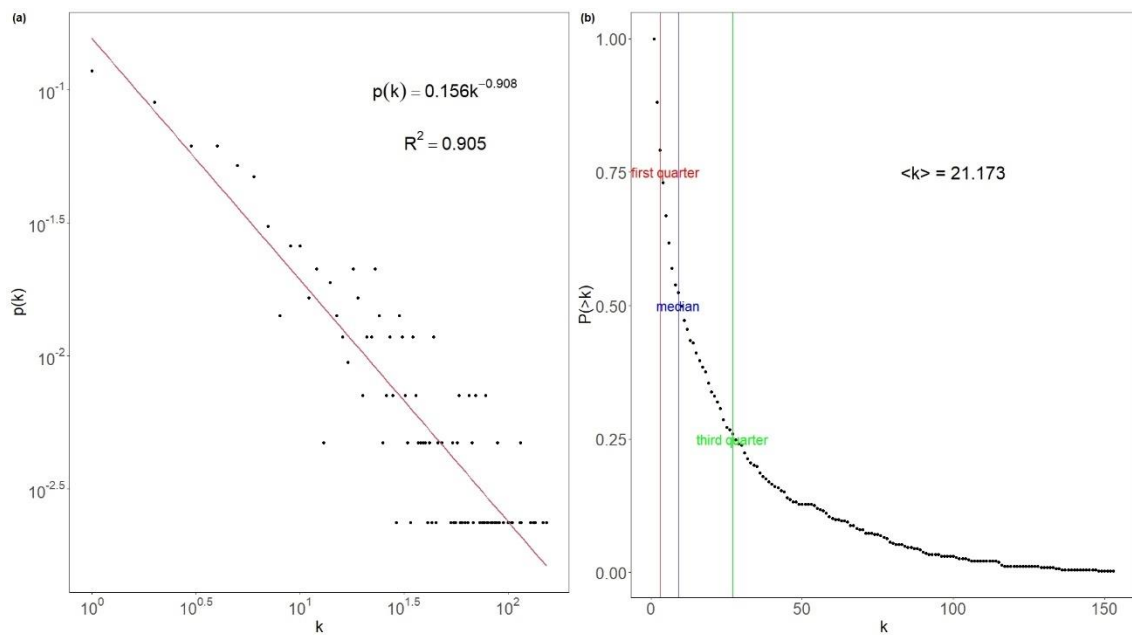


Figure 4.2 (a) Power-law degree distribution where only a small number of airports have a large number of connections. The cumulative degree distribution in (b) illustrates that 25% of airports in the network have 3 or fewer connections, 50% have 9 or fewer connections, and only the top 25% of airports have more than 27 connections.

Istanbul Atatürk Airport (IST) has the highest binary betweenness centrality as it lies on the shortest paths of airport pairs most frequently. However, it drops to the 85th position in the weighted betweenness ranking due to its shutdown and transfer of operations to the new airport of Istanbul. Similarly, London Gatwick airport (LGW) also drops to the 22nd position in the weighted betweenness ranking. Some airports such as Friedrichshafen (FDH) and Tromsø (TOS) have

high betweenness centralities despite having fewer routes and passengers, as they lie on the shortest paths of airport pairs frequently. However, when considering the distance and route weights, such anomalies are not observed. The weighted betweenness ranking (Table 4.2) is different from the binary betweenness ranking, which is based purely on the binary topology. The Kendall rank correlation coefficient between the two rankings was computed as 0.35 with a p-value $< 2.2\text{e-}16$, indicating a substantial difference between the two rankings.

Table 4.1 Ranking of airports and routes in terms of passenger traffic

Airport	Strength	Route		Weight
Amsterdam Schiphol	51,023,479	Istanbul Sabiha Gökçen	Izmir Adnan Menderes	2,751,298
Frankfurt am Main	45,039,039	Athens	Thessaloniki Macedonia	2,709,784
Adolfo Suárez Madrid-Barajas	44,566,835	Istanbul Sabiha Gökçen	Antalya	2,617,038
Barcelona	44,297,706	Barcelona	Adolfo Suárez Madrid-Barajas	2,572,844
Athens	42,076,448	Athens	Santorini	2,396,292
Paris Charles de Gaulle	41,154,368	Frankfurt am Main	Berlin-Tegel	2,251,779
Munich	38,135,299	Athens	Larnaca	2,221,212
London Heathrow	37,906,564	Toulouse-Blagnac	Paris-Orly	2,218,211
London Gatwick	37,307,711	Istanbul Sabiha Gökçen	Ankara Esenboga	2,184,197
Rome Fiumicino	31,884,165	Barcelona	Palma De Mallorca	2,173,938

4.3 Network Correlations

The analysis of correlations between centrality and connectivity measures in the network provides insights into the structure and topology of the network. One interesting finding is that there is a strong correlation between degree and

Table 4.2 Degree (k), eigenvector (k_E), betweenness (C_B) and weighted betweenness (C_{BW}) rankings of airports. Airports are shown with their IATA codes. Full names of the airports can be found in appendix A

Airport	k	Airport	k_E	Airport	C_B	Airport	C_{BW}
AMS	153	AMS	1	ISL	0.11	AMS	0.35
STN	147	BCN	0.98	LGW	0.10	CPH	0.32
FRA	135	MAD	0.96	FDH	0.09	ATH	0.30
LGW	133	LHR	0.93	TOS	0.	LHR	0.28
MUC	128	FRA	0.93	PMI	0.07	CDG	0.21
DUB	116	CDG	0.90	SEN	0.07	FCO	0.18
BCN	115	MUC	0.82	VAR	0.06	IST	0.18
CDG	115	ATH	0.78	GRO	0.05	OSL	0.16
MAN	114	FCO	0.68	LPA	0.04	FRA	0.16
MAD	105	LGW	0.63	CPH	0.04	BCN	0.15

strength (Figure 4.3), indicating that airports with a higher number of routes also handle more passengers. Additionally, the power-law relationship between degree and strength (Figure 4.4 (a)) suggests that a small increase in the number of an airport's routes leads to a larger increase in its passenger traffic. Moreover, the correlation between eigenvector centrality and strength is even stronger, suggesting that airports that offer connections to highly connected airports tend to handle relatively more passengers than those connected to lower-degree airports. Interestingly, there is a weak negative correlation between the clustering coefficients and degree centrality, indicating that highly connected airports are less likely to have their neighbors form links with each other.

According to Figure 4.3, there is no significant correlation between the unweighted betweenness centralities and degree, strength, and eigenvector centralities of the nodes. This suggests that nodes that act as transfer points may not necessarily have a large number of routes or handle a significant amount of passenger traffic based solely on their topological characteristics. Surprisingly, among the top 30 airports with the highest betweenness centralities in the EAN, 12 of them have less than 50 routes and handle less than 1 million passengers.

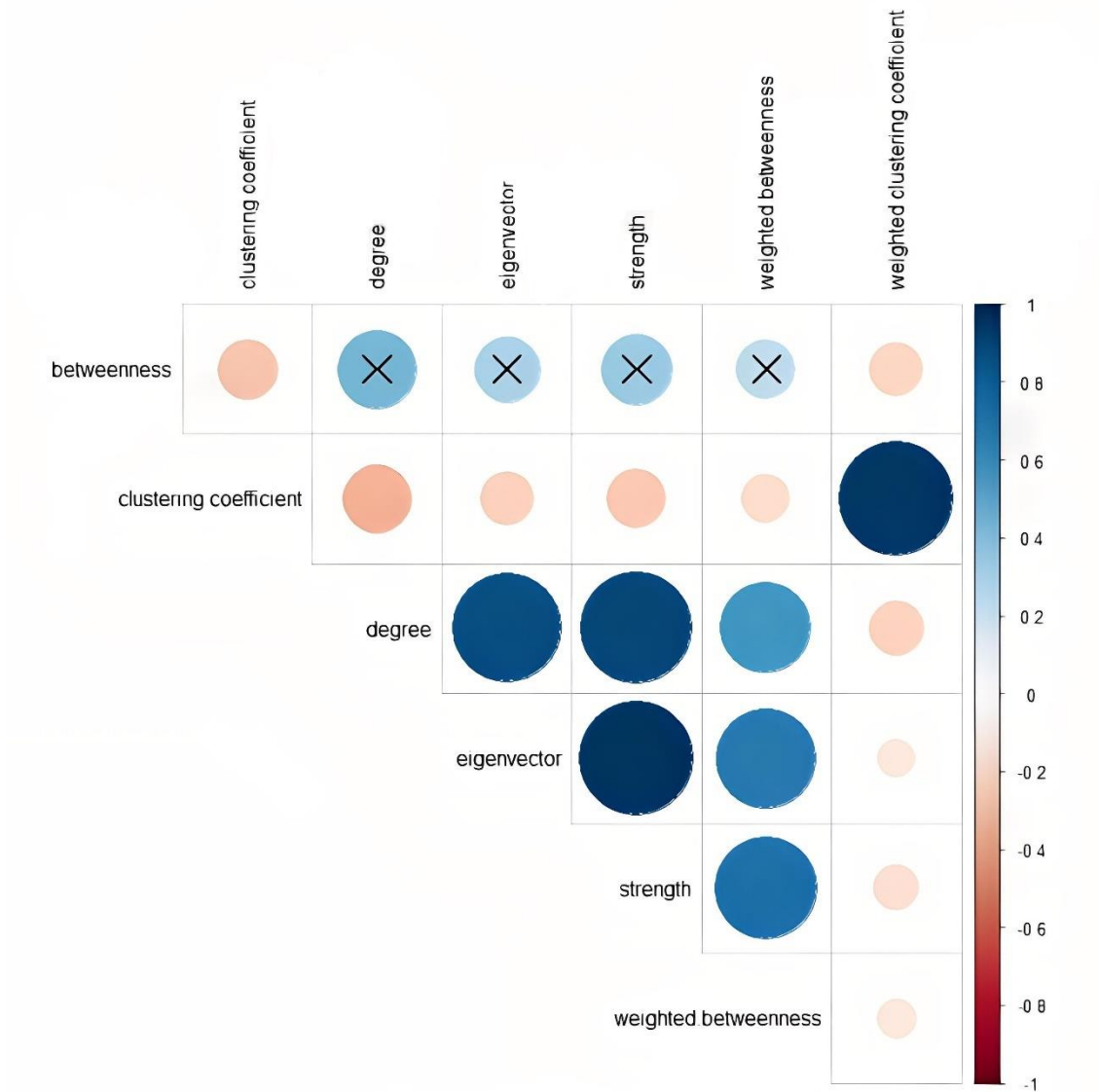


Figure 4.3 The Pearson correlation matrix is a graphical representation of the correlation between different measures of the network. The size of the circles indicates the strength of the correlation coefficient, with larger circles representing higher absolute values. The matrix includes statistically significant ($p < 0.05$) correlations, marked with circles, as well as statistically insignificant correlations, marked with a cross. correlation matrix of the network.

The assortativity coefficient of the network is -0.20, which indicates a hub-and-spoke structure where a few highly connected airports serve as hubs for many less connected airports. However, the coefficient is closer to 0 than -1, which means that there is also a significant amount of point-to-point traffic in the network. This is not surprising given the success of low-cost carriers that prioritize short-haul

point-to-point connections in Europe [55], while full-service carriers focus on long-haul intercontinental flights from their hub airports.

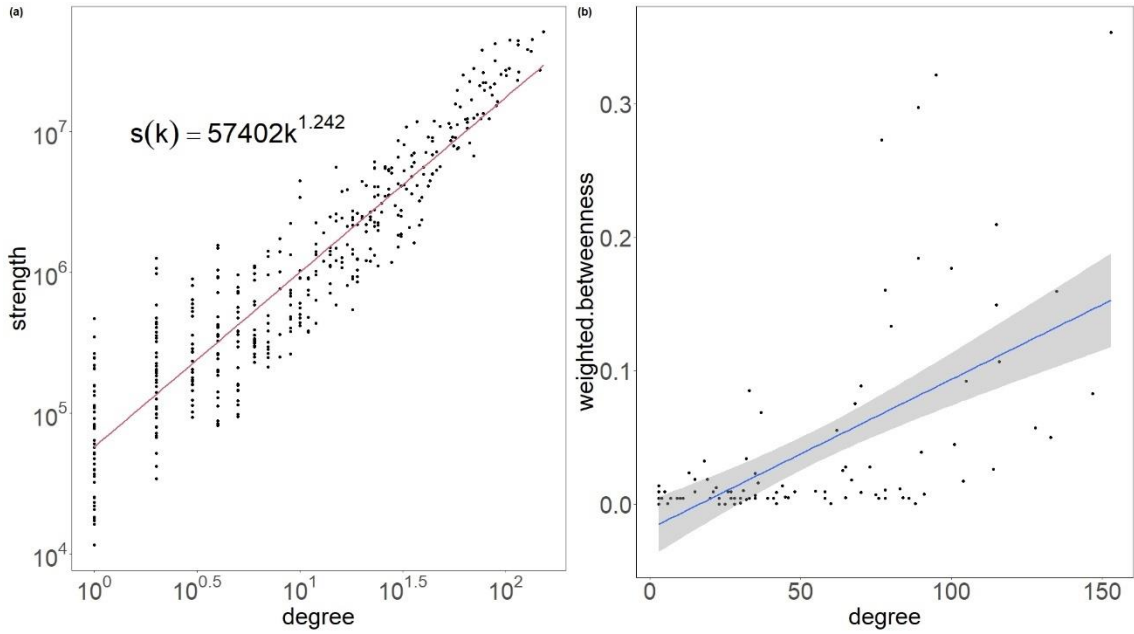


Figure 4.4 (a) Degree vs strength on log-log scale and (b) degree vs weighted betweenness.

Figure 4.4 (b) indicates that there is a moderate correlation (0.61) between weighted betweenness and degree centrality, but a stronger correlation (0.77) between weighted betweenness and strength centrality. This suggests that passengers tend to choose airports with high traffic volume and proximity for flights that require transfers. Additionally, some airports with low binary betweenness, such as Amsterdam Schiphol (AMS), London Heathrow (LHR), and Istanbul (IST), have high weighted betweenness (Table 4.2), indicating that the latter is a more relevant measure of airport hub performance. Figure 4.5 shows that most of the top 30 weighted betweenness airports have high degrees and handle a significant amount of passenger traffic, making them major hubs of the EAN. Only Bodø (BOO) and Tromsø (TOS) airports have lower degrees among the top 30. It is also noteworthy that the top 30 weighted betweenness airports are distributed across 16 different countries, suggesting a well-dispersed hub network throughout the continent.

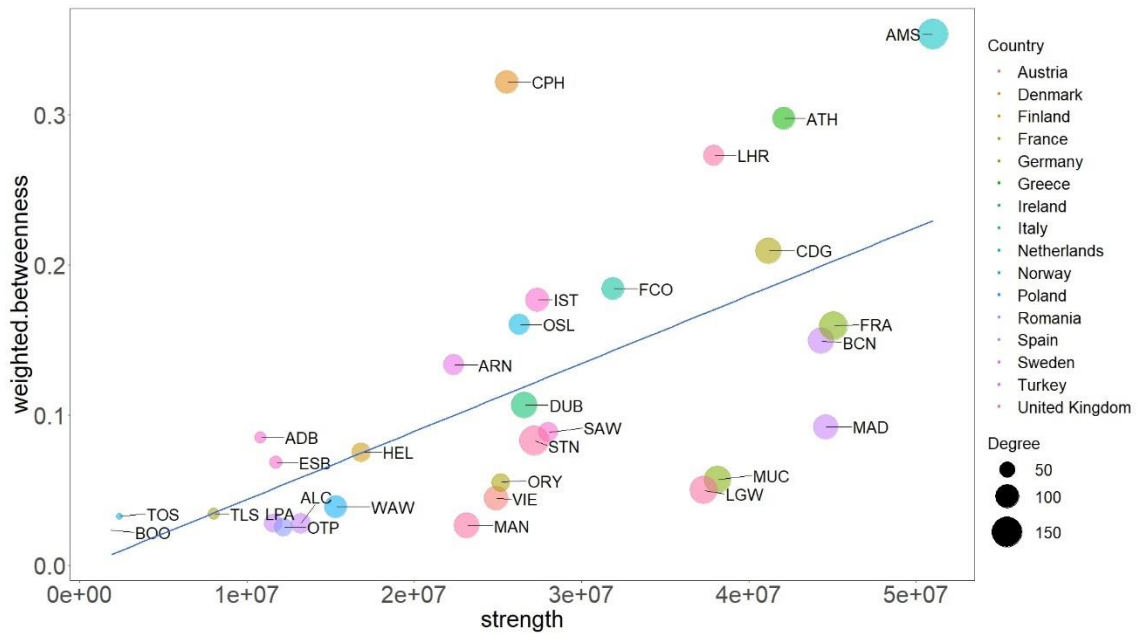


Figure 4.5 The top 30 airports in the European Air Network ranked by their weighted betweenness, with the size of the circle representing the airport's degree, and the color indicating the country in which it is located. The y-axis represents the airport's strength. There is a clear positive correlation between the weighted betweenness and strength of airports, meaning that the busiest airports tend to also have the highest weighted betweenness values.

4.4 Identification of Community Structure

Identifying community structure in a network can have practical applications such as reducing the workload of air traffic control or developing regional market strategies for airlines and airports. One way to do this is by using modularity maximization, an unsupervised learning method that takes the minimum internal connection density (resolution parameter) of the communities as an input and then maximizes the modularity function to determine the number of communities. As the resolution parameter is adjusted, the number and quality of communities may change.

The Leiden algorithm was used to identify community structure in the EAN, and this was done by maximizing modularity. In this algorithm, the minimum internal connection density (resolution parameter) of the communities was set as an input, and the algorithm then tried to maximize the modularity function while determining the number of communities for the given resolution parameter. The

passenger numbers were used as edge weights, and gamma values between 0.0 and 2.0 with 0.1 incremental steps were used as resolution parameters. The algorithm was run 10,000 times for each gamma value with a total of 200,000 iterations. At a gamma value of 1, the average maximum modularity score obtained was 0.247 with a standard deviation of 0.004 (Figure 4.6 (a)). At this value of gamma, the average number of communities found by the algorithm was 7.402 with a standard deviation of 0.501 (Figure 4.6 (b)). Therefore, EAN was investigated for communities at gamma equal to 1, which yielded on average 7-8 non-overlapping communities.

At the maximum modularity, the Leiden algorithm identifies 7 airport clusters in the EAN (Table 4.3). The smallest community, Community 7, is composed of 4 disconnected airports in Iceland, where Reykjavik is the main domestic airport and the other 3 offer routes only to this airport. The largest community (Map 4.1 (a)) is dominated by airports in Italy (33), followed by Poland (11) and Romania (11). Most of the other members of this community are located in Eastern Europe.

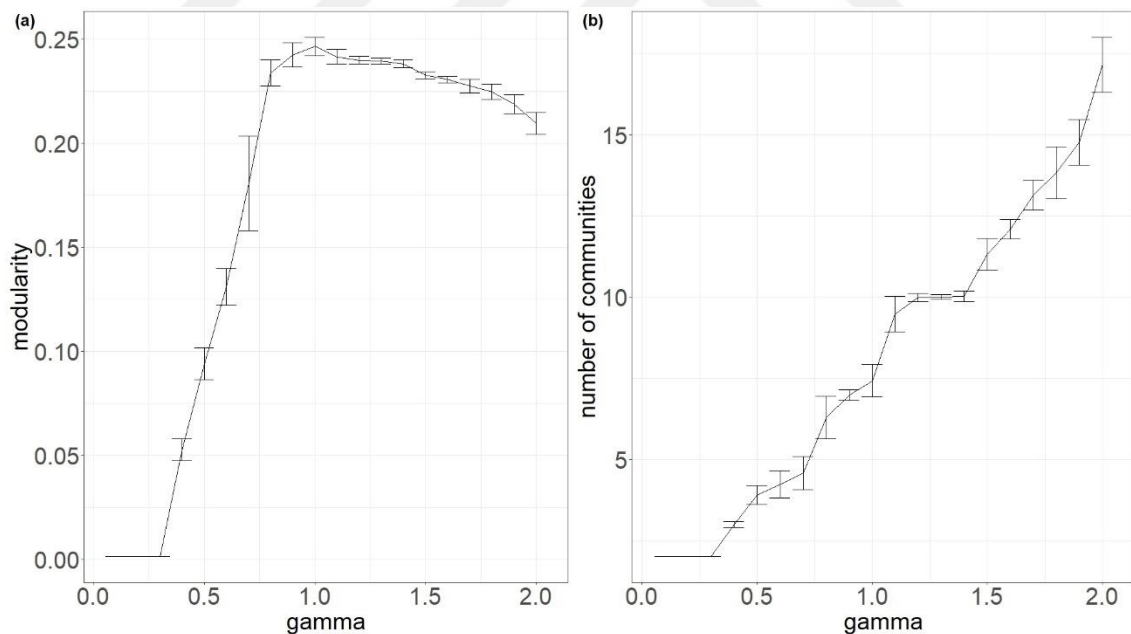


Figure 4.6 (a) modularity and (b) number of communities versus gamma.

The two low-cost airports in London, Luton and Stansted, are also part of this community and serve as connecting hubs for the UK and other member countries. The Scandinavian countries Norway (37), Sweden (19), Finland (19), and Denmark (6) form their own giant cluster (Map 4.1 (b), with Oslo Lufthavn,

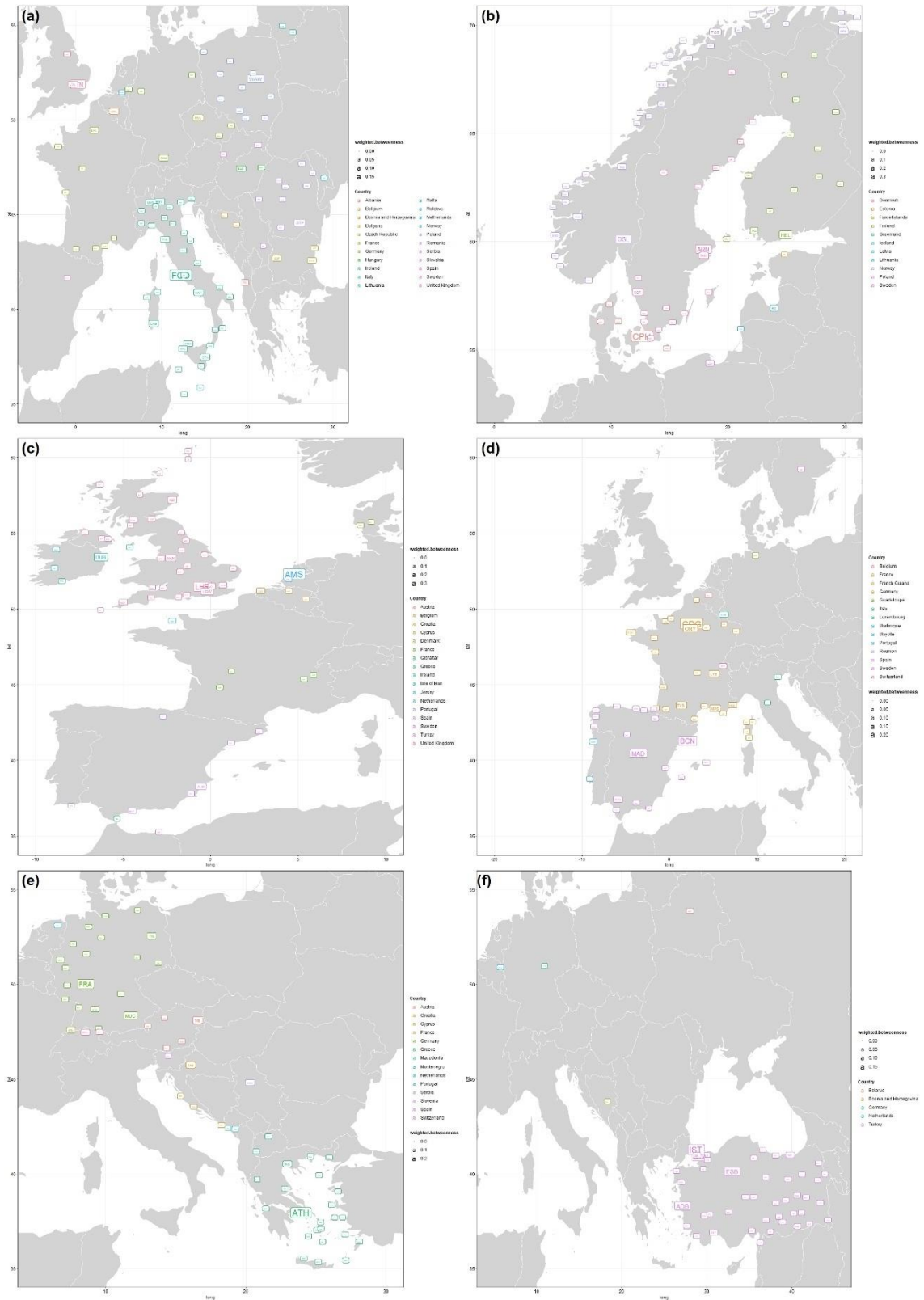
Stockholm-Arlanda, Helsinki Vantaa, and Copenhagen Kastrup serving as the main hubs of this cluster.

Table 4.3 7 communities of the EAN identified by modularity maximization

	Community						
Network measure	1	2	3	4	5	6	7
Number of nodes	91	86	70	66	59	47	4
Number of edges	600	208	379	327	281	181	3
Average path length	2.07	2.74	2.16	2.17	2.07	1.94	1.5
Average clustering coefficient	0.34	0.64	0.31	0.64	0.71	0.46	0
Diameter	5	6	5	5	4	3	2
Assortativity coefficient	-0.31	-0.4	-0.32	-0.38	-0.47	-0.76	-1
Highest degree airport	STN	OSL	AMS	MAD	FRA	IST	RKV
Highest weighted betweenness	FCO	OSL	LGW	MAD	ATH	SAW	RKV

The third community comprises 32 UK airports and a few Western European airports, with London Heathrow, London Gatwick, Dublin, and Amsterdam Schiphol serving as the major hubs (Map 4.1 (c)). The fourth largest community is mainly made up of French (28) and Spanish (23) airports, with Paris Charles de Gaulle, Paris-Orly, Adolfo Suárez Madrid–Barajas, and Barcelona serving as the connecting hubs (Map 4.1 (d)). The fifth community is primarily composed of Greek (21) and German (17) airports, with Frankfurt am Main, Munich, and Athens serving as the most central airports in this community (Map 4.1 (e)). The sixth community is composed of 43 Turkish airports and four other airports, with Istanbul, Sabiha Gökçen, Ankara Esenboğa, and Izmir Adnan Menderes serving as the hubs of passenger traffic in this community (Map 4.1 (e)). The fifth and sixth communities are the most connected ones, with smaller average path lengths and diameters, and larger clustering coefficients compared to the other communities and the overall network. The formation of the clusters is mainly influenced by

geographical proximity. Additionally, the sixth community is identified as a hub-and-spoke network with an assortativity of -0.76.



Map 4.1 Community structure of the EAN. Airport labels are scaled for weighted betweenness centrality.

5.1 Discussion of the Turkish Airport Network

The following is the summary of the key points discussed in the text about the Turkish airport network (TAN) and recommendations on how the network can be improved:

- Highly central hub airports are typically situated in cities with high population and economic activity and serve as major connection points for other airports in the network. This particular structural feature reduces the number of routes needed to create a fully connected network compared to a random point-to-point network.
- The findings of the current research were compared with those of previous studies in the literature, and the network was examined from a holistic perspective. The air transport network in Turkey was found to exhibit small-world properties, which is consistent with previous studies (as shown in Table 5.1). The average path length of Turkey's network is comparable to that of other countries, indicating that hub airports play a crucial role in linking different parts of the network, thereby reducing the network's resilience.
- According to the analysis, the Turkish air transport network has a small clustering coefficient compared to other networks, suggesting that the growth rate of nodes in the network outpaces the growth rate of edges. This means that the interconnectivity of the network could be improved by connecting lower-degree nodes with each other, for example, through regional services. Therefore, it may be beneficial to introduce such services to enhance the network's interconnectivity.

Table 5.1 Comparison of airport networks of Turkey and other countries/regions

Reference	Country	No.of nodes (n)	No.of edges (m)	<k>	(L)	(C)	Network structure
[35]	India	79	455	11,52	2,26	0,66	SW
[15]	World	3883	27051	13,93	4,4	0,62	SF SW
[56]	Italy	50	310	12,4	1,98~2,14	0,07 ~0.1	SF SW Fractal
[57]	USA	272	6566	48,28	1,84~1,93	0,73~0,78	SW
[58]	China	144	1018	14,14	2,23	0,69	SW
[17]	Australia	131	596	9,1	2,9	0,5	SW
[59]	Argentina	39	168	4,3	1,86	0,41	SF SW
This study	Turkey	54	147	6,31	1,88	0,21	SF SW

- After conducting simulations of three network models (random, small-world, and scale-free), the scale-free network model was found to be the most similar to the real Turkish air transport network, as it exhibited scale-free properties and followed a power law pattern in terms of centrality metrics and correlation calculations. This indicates that the network is not randomly connected and follows a preferential attachment mechanism, where hub airports tend to be connected to a greater number of other airports. This finding provides valuable insights into the underlying structure of the network and may help improve its overall efficiency and robustness.
- The high disassortativity in TAN suggests that it follows a hub-and-spoke structure with a small number of large hubs and an abundance of peripheral airports that preferentially connect to them. Therefore, instead of expanding the airport network by constructing new airports, it would be more beneficial for Turkey to improve the connectivity of existing airports.
- The airports with the highest connectivity in TAN, namely IST, SAW, ESB, ADB, and AYT, also have the highest centrality in terms of betweenness and closeness. However, it is noteworthy that SAW has the highest number of flights among these airports.
- The analysis of TAN revealed that the network's strong hubs are located in the western part of Turkey, while the eastern region lacks such central airports.

This highlights the need for regional air transportation services in the eastern region, and one recommendation is to establish a regional low-cost carrier based in one of the eastern cities of Turkey. This carrier could connect the region's airports to each other and to the main hubs in the west, improving the overall connectivity of the network.

- The primary constraint of this study is that the data were collected following the initial Covid-19 opening-out period. To mitigate the impact of any pandemic and to analyze the network throughout the year instead of various sample periods, additional research may be necessary. Future studies may also include international flights and various metrics, such as passenger numbers, cargo weights, and aircraft movements.

5.2 Discussion of the European Airport Network

5.2.1 Previous Studies on the EAN

Table 5.2 provides a comparison between the EAN and previous studies that have examined European airports as network nodes. Paleari, Redondi, and Malighetti [60] discovered that the EAN is a small-world network with a degree distribution that follows a Pareto law instead of a power law. The authors associated this discovery with high congestion rates at European airports. Additionally, the EAN has a lower clustering coefficient of 0.38 in comparison to the current study's finding of 0.47, which may be due to an increase in point-to-point flights over the past ten years.

Lillo et al. [61] reported a longer average path length of 3.12-3.29, which is due to this study including substantially more nodes that belong to the Eurocontrol area. However, the authors discovered an exponentially decaying degree distribution, confirming the heterogeneity of airport degrees. The average degrees found in these two studies (12-13.6) are smaller than those found in the current research paper (21.2), indicating a significant increase in the number of routes offered by airports, which may be attributed to the emergence of low-cost airline services.

Table 5.2 Summary of the previous studies and this study on the EAN

Network metrics	This study	Paleari, et al. (2010)	Lillo et al. (2011)	Cardillo et al. (2013)	Sun et al. (2015)	Wandelt et al. (2019)
Sampling duration	1 year	1 day	1 week	1 day	217 days	2002-2013
Number of nodes	423	467	1153-1375	450	-	447
Edge weights	passengers	seats	flights	flights	seats	passengers
Number of edges	4477	5544	7642-8498	multilayered airline networks	-	6159
Average degree	21.173	12.10	12.0-13.6	-	24.21	22.35
Average path length	2.55	2.8	3.12-3.29	2.2-2.8	-	2.86
Assortativity coefficient	-0.20	-	-	-	-	-0.12
Average clustering coefficient	0.47	0.38	-	0.2-0.4	0.3931	0.47
Degree distribution	power law	Pareto law	exponential	power law	Exponential	truncated power law

Cardillo et al. [20] found that the degree distribution follows a power law similar to the one found in the current study. Furthermore, the authors confirmed that the networks of low-cost carriers are responsible for the increase in the clustering coefficient by investigating the contribution of individual airline route networks to the aggregate network. This is important because the clustering

coefficient measures triangle formation in the network and is evidence that these airlines prefer point-to-point connections.

Sun, Wandelt, and Linke [21] found the highest average degree of 24.21 in their study, which covered several weeks in 2011, 2012, and 2013. Wandelt, Sun, and Zhang [23] also observed an increase in the degree centrality and attributed this increase to the successful competition of low-cost carriers in Europe.

5.2.2 EAN and Simulated Networks

In Table 5.3, we analyzed and compared the EAN network measures with those of randomly generated networks, small-world networks, and scale-free networks, each of which were simulated 10,000 times. For the scale-free network, we used the same power-law exponent as that of the EAN. Our findings show that the average path length and clustering coefficient of the EAN are similar to that of a small-world network. However, the EAN has a relatively large diameter of 6. The assortativity coefficients of the random and small-world networks are close to zero because the nodes in these networks are connected randomly and do not exhibit preferential attachment. Conversely, the assortativity of the scale-free network and the EAN is comparable because they both allow for some degree of preferential attachment, which leads to hub and cluster formation. The scale-free network has a shorter average path length and diameter than the EAN, indicating better connectivity performance.

5.2.3 Configuration of the EAN

The research findings suggest that the EAN is a scale-free network, meaning that it has a heterogeneous distribution of degree and strength. This implies that there are a few highly connected hubs in the network that offer routes not only to smaller spokes but also to other large hubs resulting in a lower assortativity coefficient compared to a typical scale-free network model. The degree and betweenness metrics computed for the EAN provide insight into the spatial concentration of flights in the network. However, in order for an airport to be a good hub in an air transport network, it also requires temporal concentration of flights. This means that the airport must have a wave-system in place [62], which allows incoming flights to connect with outgoing flights to maximize the number

of connecting possibilities for passengers. Unfortunately, the annual nature of the data used in this study doesn't provide information about whether airports in the EAN have such wave-systems in place to offer attractive connecting flights to passengers.

Table 5.3 Comparison of the EAN with network models that have the same number of nodes and edge density (0.05)

Network topological metrics	Random network	Small-world network	Scale-free network
Number of nodes	423	423	423
Number of edges	4477	4653	4477
Average degree	21.168	22	21.168
Average path length	2.279	2.709	2.023
Diameter	3.243	4.025	3.59
Assortativity coefficient	-0.005	-0.001	-0.285
Average clustering coefficient	0.0502	0.536	0.509
Degree distribution	binomial	binomial	power law

5.2.4 The Development of Air Traffic Market and Strategies of Airline Businesses

Sallan and Lordan [8] conducted a study which discovered that low-cost airlines like Ryanair and Southwest, that typically operate point-to-point networks, have the lowest spatial concentration. Low-cost airlines emerged in Europe in the early 2000s after the liberalization of air traffic in the European Common Aviation Area (ECAA) [63]. The rise of low-cost airlines in Europe over the last decade has contributed significantly to the increase in assortativity and clustering coefficients. On the other hand, airline groups with hybrid business models (full-service carriers with low-cost subsidiaries such as Air France-KLM, IAG, etc.) have the highest values of spatial concentration, which means that these airlines have their operations centered on hubs like Amsterdam and Barcelona. This study identifies the hubs with high spatial concentration in Europe and can aid hybrid airlines in determining which airports to operate from in the EAN. Additionally, marketing executives of airports with high spatial concentration could encourage hybrid carriers to operate at their airports using the results of this study. The implementation of a single market in European aviation, which allows airlines to

fly between any airport pair of member states, has made the point-to-point route network highly intricate. Low-cost airlines, which often operate from secondary airports of large cities [55], require an operational environment that meets the demands of their business model, such as shorter turn-around times and reduced landing fees. Therefore, these airports should adapt to the specific needs of low-cost airlines.

Employing a point-to-point network can result in a trade-off between efficiency and robustness, as it reduces susceptibility to node removal and makes connecting flights impractical, particularly for intra-community airport travel. Additionally, this type of operation typically limits competition since new routes are often introduced by a single airline [63]. Low-cost carriers may opt to use airports with lower spatial concentration, as they are more likely to introduce new routes and expand their market share. In turn, airport executives could establish mutually beneficial agreements with low-cost carriers to improve financial outcomes for both parties.

5.3 Conclusion

The third chapter of this study utilized complex network theory to analyze the Turkish Airport Network (TAN) using FlightRadar flight data immediately after the first wave of Covid-19. The findings reveal that TAN is highly organized and nodes have a preference for connecting to each other, indicating a hub and spoke network structure. Topology calculations such as degree distribution, average path length, and clustering coefficient support this result. However, it should be noted that these findings are limited to the specific context of the first opening period of Covid-19 and cannot be generalized beyond this case study.

The fourth chapter of this study aimed to analyze the European airport network by considering passenger traffic as the weight of edges between airports. The most connected nodes in the EAN are not necessarily the most central ones topologically. However, when taking into account the distance per passenger traffic weighted betweenness centralities of airports, the results reveal significant differences that are more aligned with their connectivity. This is because some routes have higher passenger traffic, and some airports are more geographically

central, thus making them natural passenger hubs. The study finds that binary betweenness centrality is inadequate for measuring the hub performance of airports. Future research could apply the weighted betweenness measure to global or continental flight networks to investigate if the results hold for long-haul routes. Additionally, the effective distance could be modified to include powers of geographical distance, similar to a gravity law.



REFERENCES

- [1] W. Weaver, "Science and Complexity," *American Scientist*, vol. 36, no. 4, pp. 536, 1948.
- [2] The Architecture of Complexity – Herbert A. Simon (1962), *Proceedings of the American Philosophical Society*, Vol. 106, No. 6 (Dec. 12, 1962), pp. 467-482.
- [3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex Networks: Structure and Dynamics," *Physics Reports*, vol. 424, no. 4-5, pp. 175-308, Feb. 2006.
- [4] H. Jeong, Z. Neda and A.L. Barabasi, "Measuring preferential attachment for evolving networks," *Europhysics Letters*, vol. 61, no. 4, pp. 567-572, 2003.
- [5] Barabasi, A. L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, 286(5439), 509-512. [4] H. Jeong, Z. Neda and A.L. Barabasi, "Measuring preferential attachment for evolving networks," *Europhysics Letters*, vol. 61, no. 4, pp. 567-572, 2003.
- [6] M. Zanin and F. Lillo, "Modelling the air transport with complex networks: A short review," *The European Physical Journal Special Topics*, vol. 215, no. 1, pp. 5-21, 2013.
- [7] Airbus, Global Market Forecast. "2019-2038." (2019). [Online]. Available: <https://www.airbus.com/content/dam/corporate-topics/strategy/global-market-forecast/GMF-2019-2038-Airbus-Commercial-Aircraft-book.pdf>. [Accessed: Jan. 16, 2019].
- [8] J. M. Sallan and O. Lordan, "Air Transport Networks and Complex Networks," in *Air Route Networks Through Complex Networks Theory*, Eds. Elsevier, 2020
- [9] R. Guimera, S. Mossa, A. Turtleschi, and L. A. N. Amaral, "The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles," *Proc. Natl. Acad. Sci. USA*, vol. 102, no. 22, pp. 7794-7799, May 2005. DOI: 10.1073/pnas.0407994102.
- [10] W. Li and X. Cai, "Statistical analysis of airport network of China," *Physica A: Statistical Mechanics and its Applications*, vol. 337, no. 1-2, pp. 242-248, 2004.
- [11] R. Doganis, "Flying off course: Airline economics and marketing," Routledge, 2009.
- [12] O. Lordan, J. M. Sallan, P. Simo, and D. Gonzalez-Prieto, "Robustness of the air transport network," *Transportation Research Part E: Logistics and Transportation Review*, vol. 68, pp. 155-163, Dec. 2014.
- [13] A. Cook, H. A. Blom, F. Lillo, R. N. Mantegna, S. Micciche, D. Rivas, et al., "Applying complexity science to air traffic management," 2015.

- [14] Liu Hao, Xiao-Bing Hu, Saini Yang, Ke Zhang, and Ezequiel Di Paolo, "Application of Complex Network Theory and Genetic Algorithm in Airline Route Networks," *Transportation Research*, 2011.
- [15] R. Guimera, L.A.N. Amaral, "Modeling the world-wide airport network," *Eur. Phys. J. B*, vol. 38, no. 2, pp. 381-385, 2004.
- [16] C. Li-Ping, W. Ru, S. X. Xin-Ping, Z. Jin-Song, L. Wei, and C. Xu, "Structural Properties of US Flight Network," in 2017 International Conference on Smart Transportation and Future Mobility (CSTFM), Chengdu, China, 2017, pp. 129-134, doi: 10.1109/CSTFM.2017.8066961.
- [17] M. M. Hossain and S. Alam, "A Complex Network Approach Towards Modeling and Analysis of the Australian Airport Network," *J. Air Transp. Manage.*, vol. 60, pp. 1-9, Jan. 2017, doi: 10.1016/j.jairtraman.2016.12.008.
- [18] J. Lin, "Network analysis of China's aviation system, statistical and spatial structure," *Journal of Transport Geography*, vol. 22, pp. 109-117, 2012, doi: 10.1016/j.jtrangeo.2011.12.002.
- [19] O. Lordan and J.M. Sallan, "Core and critical cities of global region airport networks," *Physica A: Statistical Mechanics and its Applications*, vol. 513, pp. 724-733, Jan. 2019, doi: 10.1016/j.physa.2018.08.123.
- [20] A. Cardillo, J. Gomez-Gardenes, M. Zanin, M. Romance, D. Papo, F. del Pozo, and S. Boccaletti, "Emergence of Network Features from Multiplexity," *Scientific Reports*, vol. 3, 2013, Art. no. 1344, doi: 10.1038/srep01344.
- [21] X. Sun, S. Wandelt, and F. Linke, "Temporal Evolution Analysis of the European Air Transportation System: Air Navigation Route Network and Airport Network," *Transportmetrica B: Transport Dynamics*, vol. 3, no. 2, pp. 153-168, 2015. [Online]. Available: <https://doi.org/10.1080/21680566.2014.960504>.
- [22] Pien, K., K. Han, W. Shang, A. Majumdar, and W. Ochieng. 2015. "Robustness Analysis of the European Air Traffic Network." *Transportmetrica A Transport Science* 11 (9): 772-792. <https://doi.org/10.1080/23249935.2015.1087233>.
- [23] S. Wandelt, X. Sun, and J. Zhang, "Evolution of Domestic Airport Networks: A Review and Comparative Analysis," *Transportmetrica B: Transport Dynamics*, vol. 7, no. 1, pp. 1-17, 2019. [Online]. Available: <https://doi.org/10.1080/21680566.2017.1301274>.
- [24] G. Gurtner, S. Vitali, M. Cipolla, F. Lillo, R. N. Mantegna, S. Miccichè, and S. Pozzi, "Multi-Scale Analysis of the European Airspace Using Network Community Detection," *PLOS ONE*, vol. 9, no. 5, p. e94414, May 2014, doi: 10.1371/journal.pone.0094414.
- [25] M. G. Song and G. T. Yeo, "Analysis of the air transport network characteristics of major airports," *The Asian Journal of Shipping and Logistics*, vol. 33, no. 3, pp. 117-125, 2017.
- [26] S. K. Depren and F. G. Yavuz, "The Network Analysis of The Domestic And International Air Transportation Structure Of Turkey," *Mugla Journal of Science and Technology*, vol. 4, no. 2, pp. 148-155, 2018.

- [27] U. Erdem, K. M. Cubukcu, and D. Tsiotas, "Mapping the topology of the air transport network in Turkey," *Environment and Planning A: Economy and Space*, vol. 52, no. 1, pp. 6-9, Jan. 2020, doi: 10.1177/0308518X19895367.
- [28] A. E. Sarilgan, "The Requirements Of Developing The Regional Airline Transportation In Turkey," *Anadolu University Journal of Social Sciences*, vol. 11, no. 1, pp. 69-88, 2011.
- [29] O. Icoz, A. Kutuk, and O. Icoz, "Social media and consumer buying decisions in tourism: The case of Turkey," *Tourism Management Perspectives*, vol. 25, pp. 13-22, 2018. [Online]. Available: <https://doi.org/10.1016/j.tmp.2017.10.006>.
- [30] Ö. Akman and C. Öztürk , "Care Behaviors of Nursing Students and the Influencing Factors: A Comparative Study", *İstanbul Gelişim Üniversitesi Sağlık Bilimleri Dergisi*, no. 14, pp. 182-200, Aug. 2021, doi:10.38079/igusabder.867994
- [31] E. Lehman, T. Leighton, and A. Meyer, "Chapter 5: Graph Theory," MIT OpenCourseWare, Fall 2010. [Online]. Available: https://ocw.mit.edu/courses/6-042j-mathematics-for-computer-science-fall-2010/f471f7b7034fabe8bbba5507df7d307f_MIT6_042JF10_chap05.pdf. [Accessed: 08-Feb-2023].
- [32] E. Lehman, T. Leighton, and A. Meyer, "Mathematics for Computer Science: Chapter 6 - Directed Graphs," MIT OpenCourseWare, Fall 2010. [Online]. Available: https://ocw.mit.edu/courses/6-042j-mathematics-for-computer-science-fall-2010/e6db7638031b754f5f68012946af4763_MIT6_042JF10_chap06.pdf. [Accessed: 09-Feb-2023].
- [33] M. Barthelemy, "Spatial Networks," *Physics Reports-Review Section of Physics Letters*, vol. 499, no. 1-3, pp. 1-101, 2011. [Online]. Available: <https://doi.org/10.1016/j.physrep.2010.11.002>.
- [34] M.E.J. Newman, "Mathematics of Networks," in *The New Palgrave Dictionary of Economics*, London: Palgrave Macmillan, 2008, pp. 1-5, doi: 10.1057/978-1-349-95121-5_2565-1.
- [35] G. Bagler, "Complex network view of performance and risks on airport networks," in *Airports: Performance, Risks, and Problems*, 2009, pp. 199-205.
- [36] A. Barrat, M. Barthelemy, R. Pastor-Satorras, and A. Vespignani, "The Architecture of Complex Weighted Networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 11, pp. 3747-3752, Mar. 2004, doi: 10.1073/pnas.0400087101.
- [37] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social Networks*, vol. 1, no. 3, pp. 215-239, 1978, doi: 10.1016/0378-8733(78)90021-7.
- [38] L. Dall'Asta, A. Barrat, M. Barthelemy, and A. Vespignani, "Vulnerability of weighted networks," *J. Stat. Mech.-Theory Exp.*, p. P04006, 2006. doi: 10.1088/1742-5468/2006/P04006.

- [39] M. E. J. Newman, "Assortative mixing in networks," *Phys. Rev. Lett.*, vol. 89, no. 20, p. 208701, Nov. 2002, doi: 10.1103/PhysRevLett.89.208701.
- [40] R. J. Mondragón, "Estimating Degree–degree Correlation and Network Cores from the Connectivity of High–degree Nodes in Complex Networks," *Scientific Reports*, vol. 10, no. 1, p. 5668, 2020. [Online]. Available: <https://doi.org/10.1038/s41598-020-62523-9>.
- [41] A. L. Barabasi, "Network Science," *Philosophical Transactions of the Royal Society A-Mathematical Physical and Engineering Sciences*, vol. 371, no. 1987, 2013. <https://doi.org/10.1098/rsta.2012.0375>.
- [42] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, no. 1, pp. 47-97, Jan. 2002, doi: 10.1103/RevModPhys.74.47.
- [43] P. Erdős and A. Rényi, "On the evolution of random graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, vol. 5, pp. 17-61, 1960.
- [44] B. Bollobás, "Random graphs," in *Cambridge Studies in Advanced Mathematics*, vol. 73. Cambridge: Cambridge University Press, 1985.
- [45] B. Bollobás, "The Evolution of Random Graphs," *American Mathematical Society*, 1984.
- [46] B. Bollobás, "Degree sequences of random graphs," *Discrete Math.*, vol. 33, 1981, pp. 1-17.
- [47] D. J. Watts and S. H. Strogatz, "Collective dynamics of ‘small-world’ networks," *Nature*, vol. 393, no. 6684, pp. 440-442, Jun. 1998, doi: 10.1038/30918.
- [48] J. Guare, *Six Degrees of Separation: A Play*, Vintage Books, New York, 1990.
- [49] A. L. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509-512, Oct. 1999, doi: 10.1126/science.286.5439.509.9
- [50] L. Ostroumova Prokhorenkova and E. Samosvat, "Global Clustering Coefficient in Scale-Free Networks," in *Algorithms and Models for the Web Graph*, A. Bonato, F. Graham, and P. Prałat, Eds. Springer, Cham, 2014, vol. 8882. [Online]. Available: https://doi.org/10.1007/978-3-319-13123-8_5.
- [51] U. Brandes, D. Delling, M. Gaertler, R. Gorke, M. Hoefer, Z. Nikoloski and D. Wagner, "On modularity clustering," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 2, pp. 172-188, Feb. 2008, doi: 10.1109/TKDE.2007.190689.
- [52] V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, Oct. 2008, Art. no. P10008, doi: 10.1088/1742-5468/2008/10/P10008.

- [53] V. A. Traag, L. Waltman, and N. J. van Eck, "From Louvain to Leiden: Guaranteeing Well-connected Communities," *Scientific Reports*, vol. 9, no. 1, p. 5233, 2019. <https://doi.org/10.1038/s41598-019-41695-z>.
- [54] T. Çetin and S. Benk, "Regulation, deregulation, and competition in the Turkish airline industry," in *The Political Economy of Regulation in Turkey*, New York, NY, Springer, 2011, pp. 193-214.
- [55] A. Magdalina and M. Bouzaima, "An Empirical Investigation of European Airline Business Models: Classification and Hybridisation," *Journal of Air Transport Management*, vol. 93, p. 102059, 2021. [Online]. Available: <https://doi.xuorg/10.1016/j.jairtraman.2021.102059>.
- [56] M. Guida and F. Maria, "Topology of the Italian airport network: A scale-free small-world network with a fractal structure?," *Chaos, Solitons & Fractals*, vol. 31, pp. 527-536, 2007.
- [57] Z. Xu and R. Harriss, "Exploring the structure of the U.S. intercity passenger air transportation network: a weighted complex network approach," *GeoJournal*, vol. 73, no. 2, pp. 87–102, 2008, doi: 10.1007/s10708-008-9166-4.
- [58] Wang, J., Mo, H., Wang, F., & Jin, F. (2011). Exploring the network structure and nodal centrality of China's air transport network: A complex network approach. *Journal of Transport Geography*, 19(4), 712-721. <https://doi.org/10.1016/j.jtrangeo.2010.08.010>.
- [59] M.P. Guillaumet, "A complex network approach to the Argentinian Airport Network," 2018, doi: 10.13140/RG.2.2.33946.54723.
- [60] S. Paleari, R. Redondi, and P. Malighetti, "A comparative study of airport connectivity in China, Europe and US: Which network provides the best service to passengers?," *Transportation Research Part E: Logistics and Transportation Review*, vol. 46, no. 2, pp. 198-210, Feb. 2010, doi: 10.1016/j.tre.2009.08.003.
- [61] F. Lillo, S. Micciche, R. N. Mantegna, V. Beato and S. Pozzi, "ELSA project: Toward a complex network approach to ATM delays analysis," 2011 SESAR Innovation Days, Toulouse, France, 2011, pp. 1-10, doi: 10.1109/SID.2011.6117307.
- [62] G. Burghouwt and J. de Wit, "Temporal Configurations of European Airline Networks," *Journal of Air Transport Management*, vol. 11, no. 3, pp. 185-198, 2005.
- [63] F. Dobruszkes, M. Givoni, and T. Vowles, "Hello Major Airports, Goodbye Regional Airports? Recent Changes in European and US Low-cost Airline Airport Choice," *Journal of Air Transport Management*, vol. 59, pp. 50-62, 2017. <https://doi.org/10.1016/j.jairtraman.2016.11.005>.

Airport of Turkey and their IATA codes:

Adana Sakirpasa Airport (ADA)

Ankara Esenboga Airport (ESB)

Antalya Airport (AYT)

Balikesir Koca Seyit Airport (BZI)

Batman Airport (BAL)

Bodrum Milas Airport (BJV)

Bursa Yenisehir Airport (YEI)

Canakkale Airport (CKZ)

Dalaman Airport (DLM)

Denizli Cardak Airport (DNZ)

Diyarbakir Airport (DIY)

Edremit Korfez Airport (EDO)

Elazig Airport (EZS)

Erzincan Airport (ERC)

Erzurum Airport (ERZ)

Gaziantep Oguzeli Airport (GZT)

Hatay Airport (HTY)

Igdir Airport (IGD)

Istanbul Ataturk Airport (IST)

Istanbul Sabiha Gokcen Airport (SAW)

Izmir Adnan Menderes Airport (ADB)

Kahramanmaras Airport (KCM)
Kars Harakani Airport (KSY)
Kastamonu Airport (KFS)
Kayseri Erkilet Airport (ASR)
Konya Airport (KYA)
Kutahya Zafer Airport (KZR)
Malatya Erhac Airport (MLX)
Mardin Airport (MQM)
Mus Airport (MSR)
Nevsehir Kapadokya Airport (NAV)
Ordu Giresun Airport (OGU)
Samsun Carsamba Airport (SZF)
Sanliurfa Guney Anadolu Airport (GNY)
Sinop Airport (SIC)
Sivas Nuri Demirag Airport (VAS)
Tekirdag Corlu Airport (TEQ)
Trabzon Airport (TZX)
Usak Airport (USQ)
Van Ferit Melen Airport (VAN)
Zonguldak Airport (ONQ)

Airports of the European Network mentioned in the text with their IATA codes:

Alicante International Airport (ALC)
Amsterdam Airport Schiphol (AMS)
Stockholm-Arlanda Airport (ARN)
Eleftherios Venizelos International Airport (ATH)

Barcelona International Airport (BCN)

Bodø Airport (BOO)

Charles de Gaulle International Airport (CDG)

Copenhagen Kastrup Airport (CPH)

Dublin Airport (DUB)

Leonardo da Vinci–Fiumicino Airport (FCO)

Friedrichshafen Airport (FDH)

Frankfurt am Main Airport (FRA)

Girona Airport (GRO)

Helsinki Vantaa Airport (HEL)

Santorini Airport (JTR)

Larnaca International Airport (LCA)

London Gatwick Airport (LGW)

London Heathrow Airport (LHR)

Gran Canaria Airport (LPA)

London Luton Airport (LTN)

Adolfo Suárez Madrid–Barajas Airport (MAD)

Manchester Airport (MAN)

Munich Airport (MUC)

Paris-Orly Airport (ORY)

Oslo Lufthavn (OSL)

Henri Coandă International Airport (OTP)

Palma De Mallorca Airport (PMI)

Reykjavik Airport (RKV)

Southend Airport (SEN)

Thessaloniki Macedonia International Airport (SKG)

London Stansted Airport (STN)

Toulouse-Blagnac Airport (TLS)

Tromsø Airport (TOS)

Berlin-Tegel Airport (TXL)

Varna Airport (VAR)

Vienna International Airport (VIE)

Warsaw Chopin Airport (WAW)



PUBLICATIONS FROM THE THESIS

Conference Papers

- [1] C. Ersoz, S. Kilic and H. O. Aldemir, "Exploring Turkey's airport network structure: a complex network approach," in Proceedings of the 6th International New York Conference on Evolving Trends in Interdisciplinary Research & Practices, New York, NY, USA, Apr. 2022, pp. 183-188.
- [2] C. Ersöz and F. Karaman, "Analysis of Europe's Air Transport Network from a Complex Network Theory Perspective," presented at the 2nd International Conference on Innovative Academic Studies, Konya, Türkiye, Jan. 2023.

Papers

- [1] C. Ersoz, S. Kilic, and H. O. Aldemir, "Analysis of Turkey's airport network structure and centrality in the opening-out period after the first wave of COVID-19: A complex network approach," Case Studies on Transport Policy, vol. 10, no. 4, pp. 2506-2518, Nov. 2022. DOI: 10.1016/j.cstp.2022.11.013.
- [2] C. Ersöz and F. Karaman, "Centrality and connectivity analysis of the European airports: a weighted complex network approach," Transportation Planning and Technology, vol. 46, no. 2, pp. 200-223, 2023. [Online]. Available: <https://doi.org/10.1080/03081060.2023.2166509>.