

On the Topology of Air Navigation Route Systems

Xiaoqian Sun, MSc, PhD

Beihang University, 100191 Beijing, China

Sebastian Wandelt*, MSc, PhD

Beihang University, 100191 Beijing, China

(*) Contact address: wandelt@buaa.edu.cn

Florian Linke, Dipl.-Ing.

German Aerospace Center, 21079 Hamburg, Germany

Air Traffic Management (ATM) is the dynamic, integrated management of traffic in airspace. As aircraft fly through the sky, they follow pre-planned routes, much like highways on the ground. In this research, we, for the first time, analyse the air navigation route system of fifteen different countries from a consistent worldwide airspace database and compare these airspace structures using complex network theory. We investigate the following five metrics: degree, distance strength, weighted betweenness centrality, weighted closeness centrality, and edge length distribution. For each metric, we perform regression analysis in order to identify abstract, complex network patterns holding for each of the countries. We find that air navigation route networks for all fifteen countries are rather heterogeneous. Furthermore, we discover that the degree distribution for all countries is better fitted by tetration, instead of an exponential function, as believed in previous work on single countries. Analysis of weighted betweenness centrality shows that some countries (e.g. USA) are robust against random or targeted node failures; while other countries (e.g. South Africa) are rather vulnerable. The hierarchical clustering based on the regression coefficients shows that the countries with similar geographical features are clustered together.

Keywords: *Traffic engineering, Transport management, Transport planning*

List of notations

The set of notations used in the paper is as follows:

- k : Degree of a node
- $P(\geq k)$: Cumulative degree distribution
- $P(\geq b)$: Cumulative betweenness distribution
- $P(\geq c)$: Cumulative closeness distribution
- H : Shannon diversity index

1. Introduction

Air transportation systems are complex systems, since there are large numbers of heterogeneous components with complex structures and interactions between groups of different components (Janić, 2007). The requirements on air traffic are increasing, not only regarding lower costs, but also higher safety, better service quality, broader social acceptance, and more environmental friendliness (European Commission, 2011). With increased demand of air traffic, the current structure of air transportation systems will reach the limit in the next years (Bonnefoy, 2008). As one major stakeholder in air transportation systems, it is a challenging task for ATM to regulate the flow of air traffic and the use of airspace in a safe, cost-efficient, environmental-friendly way (Djokic *et al.*, 2008; Lillo *et al.*, 2011).

Airport networks have been the focus of many researchers, where airports are nodes, and an edge exists if there is a flight between two airports. However, airport network is only one perspective of air transportation systems. Another perspective is to consider how aircraft actually fly through the airspace. Aircraft have to follow so-called air routes. Air routes are "highways in the sky" and they are designed to channel the flow of air traffic in a predictable manner to ensure safety (Cordell, 2006). A route is composed of several route segments, where one route segment consists of two consecutive significant points, such as Very-high-frequency Omni-directional Range (VOR), Distance Measuring Equipment (DME), TACtical Air Navigation (TACAN), Non-Directional Beacon (NDB), or Designated Point (DPN) (Cordell, 2006). Fig. 1 shows one example of an air route from Hamburg to Frankfurt in Germany, where AHP stands for Aerodrome heliport. Aircraft from Hamburg to Frankfurt have to fly along the air route rather than a straight line. The left hand side of Fig. 1 visualizes the air route; the middle shows the ID, latitude, longitude, and types of the air navigation points; the right hand side shows the start and end points for the route segments, together with their length in km.

In this study, with the perspective of Cai *et. al.* (Cai *et al.*, 2012) as a starting point, we investigate topological properties of air navigation route network in terms of five key metrics: degree, distance strength, weighted betweenness centrality, weighted closeness centrality, and edge length distribution. In total, we analyse fifteen regional networks for countries from six continents: Three countries each for Africa, Asia, Europe, and South America; two countries for North America; and one country for Oceania. For each continent, these countries have the largest number of nodes. We extract the data from EAD to build the air navigation route networks for the fifteen selected countries. Fig. 2 shows the air navigation route network for each of the countries (Note that Congo refers to Democratic Republic of Congo).

The four main contributions of our paper are: 1) We find that air navigation route networks for all fifteen countries are rather heterogeneous and the degree distribution is better fitted by tetration than exponential function, as believed in previous work (Cai *et al.*, 2012; Gurtner *et al.*, 2014). 2) Based on the power law behaviour of distance strength distribution, we show that the fifteen countries can be categorized into two groups: Industrial countries where hub air navigation nodes have large connectivities and long-distance connections; while emerging countries do not have hub nodes and the emerging countries could consider to adapt their network structure from industrial countries when expanding their air navigation route system. 3) Our analysis for weighted betweenness centrality shows that some countries (e.g. USA) are robust against random or targeted node failures; while other countries (e.g. South Africa) are rather vulnerable. 4) The hierarchical clustering based on the regression coefficients shows that the countries with similar geographical features are clustered together.

This paper is organized as follows. In Section 2, we provide the literature review on complex network theory in air transportation systems. Section 3 presents the fifteen countries and how we extract the data from EAD to model air navigation route systems. In Section 4, we present and discuss the results of network analysis for air navigation route systems. The paper is concluded in Section 5.

2. Literature Review

Complex network theory has been recently applied to the study of air transportation systems. Most research focused on airport networks, where each node represents one airport, and an edge exists if there is a flight between two airports (Holmes and Scott, 2004; Barrat *et al.*, 2004; DeLaurentis *et al.*, 2008; Reggiani *et al.*, 2010; Cardillo and et. al., 2013). Guimera *et. al.* analysed the worldwide air transportation

network with nodes as cities and found that the most connected cities are not necessarily the most central ones (Guimera *et al.*, 2005). Bagler investigated the airport network of India to explore its various properties and compare them with their topological counterparts (Bagler, 2008). Wang et. al. analysed the air transportation characteristics for New York metroplex airports (Wang *et al.*, 2008). Kotegawa et. al. used topological structures of service networks to examine trade-offs between efficiency metrics established from passenger, airline, and air navigation service provider perspectives (Kotegawa *et al.*, 2010, 2014). Arranz et. al. explored the network-wide effects of different prioritisation strategies in air transportation systems (Arranz *et al.*, 2013). Fleurquin et. al. studied flight delay propagation in the US airport network (Fleurquin *et al.*, 2013a) and developed an agent-based model to reproduce the delay-spreading dynamics under severe weather conditions (Fleurquin *et al.*, 2013b). Lehner investigated the structure-function networks for European air transport, with airport network represents the structure and passenger flow network characterizes the function (Lehner, 2013).

Related works (Cai *et al.*, 2012; Vitali and et. al., 2012; Gurtner *et al.*, 2014; Sun and Wandelt, 2014; Sun *et al.*, 2014) either consider air navigation route network at an individual country level (China or Italy) or for a whole continent (Europe). Especially, their analysis is based on different datasets, thus, these networks are generally not comparable and it is difficult to explain similarities and differences. In this paper, we use data from a single, consistent database (European Aeronautical Information Service Database, EAD) provided by EUROCONTROL (<http://www.ead.eurocontrol.int/eadcms/eadsite/index.php.html>). We extract the data from EAD to build the air navigation route network. Given a consistent view on the global air navigation network, it is, for the first time, feasible to study the structures of different sub-networks.

3. Modelling of air navigation route systems

3.1. Air navigation route networks

Topological analysis is usually one of the first steps in understanding the behavior of complex networks (Newman, 2004; Kurant and Thiran, 2006). We briefly introduce basic concepts from complex network theory, relevant for the analysis of air navigation route networks. A network is a set of nodes and edges. Edges can be either directed or undirected: While directed edges can only be traversed from their source to the target, undirected edges can be accessed in the opposite direction as well. In weighted networks each edge is labelled with a weight (in unweighted networks the weight is often assumed to be one). A path is an ordered sequence of nodes and edges, linking a source node and a target node. A shortest path in a weighted network between two nodes is a path such that the sum of the weights of its edges is minimized (Pfeifer, 2013).

In the air navigation route network, nodes are air navigation points and edges are route segments connecting air navigation nodes. In this research, we consider the air navigation route network as an undirected network, with the assumption that the route network is approximately symmetric (Lillo *et al.*, 2011). We use the great circle distance as edge weight. The great circle distance is computed from latitude/longitude pairs for source and target node, using the Haversine formula (Sinnott, 1984).

Table 2 summarizes five metrics for the undirected, weighted air navigation route network: degree, distance strength, weighted betweenness centrality, weighted closeness centrality, and edge length distribution. We use these five metrics to analyse topological properties of the air navigation route system in the following part of the paper.

3.2. Database

There are only few datasets available for analyzing the air transportation route network. For the European airspace, it is within the responsibility of Eurocontrol to manage the air traffic routes over Europe, in strong coordination with the domestic traffic agencies. The German Aerospace Center, where our research was performed, has access to the route data of Eurocontrol. We obtain the data (effective on 18 October 2012) from the centralized reference European Aeronautical Information Service Database (EAD) provided by EUROCONTROL. In EUROCONTROL, the Aeronautical Information Services (AIS) data is managed and distributed using the Aeronautical Information Exchange Model (AIXM), a semi-structured language designed to enable the management and distribution of AIS data in digital format. The two components of AIXM are: AIXM Conceptual Model and AIXM XML Schema. The AIXM Conceptual

Model describes the features and their properties within the aeronautical domain. The AIXM XML Schema is an exchange model for aeronautical data: It is an implementation of the conceptual model as an XML schema.

In this paper, we build air navigation route sub-networks for different geographic regions based on the EAD. Note that there are no air navigation routes across the Atlantic ocean in EAD. These routes connecting Europe and North America are called North Atlantic Tracks (NAT). The NAT are updated daily to take into account forecast weather conditions (See ([Rodionova et al., 2012](#)) for a further study about the NAT). In this paper, we investigate the static structures of the air navigation route system. Since we build regional air navigation route networks at a country level, the NAT are irrelevant in our study. In total, we generate fifteen regional networks for countries from six continents. For each continent, these countries have the largest number of nodes. These fifteen air navigation route networks are shown in Fig. 2.

For each country, we show seven attributes in Table 1: geographic area (km^2), population, population per geographic area, number of nodes, percentage of the nodes in the worldwide air navigation route network, number of edges, and percentage of the edges in the worldwide air navigation route network. The fifteen countries account for 55% of all the nodes (38,473) and 53% of all the edges (64,167) in the worldwide air navigation route network extracted from the EAD dataset. The bold numbers in the table are the maximum and the minimum values for each attribute. Note that the fifteen countries cover 63% of the airports in the world (<https://www.cia.gov/library/publications/the-world-factbook/rankorder/2053rank.html>).

USA has the highest number of nodes and edges while Morocco has the smallest one; Germany has the highest ratio of nodes per geographic area and China has the lowest one. Russia has the biggest geographic area and Italy has the smallest one; China has the largest population and Australia has the smallest one. In general, countries with bigger geographic area have more nodes. Note that the top three countries with highest ratio of nodes per area (Germany ranks first, followed by Italy and Japan) are also the top three countries with highest ratio of population per area (Japan ranks first, followed by Germany and Italy). These three countries have the smallest geographic area. It is interesting to investigate further topological properties of the air navigation route network and its connections to social-economic factors in the future ([Barthélemy, 2011](#)).

4. Analysis of air navigation route networks

In this section, we compare the structures of air navigation route networks for the fifteen countries. We analyse network topologies of the air navigation route system in terms of five key metrics: degree, distance strength, weighted betweenness centrality, weighted closeness centrality, and edge length distribution. These metrics are described in Table 2.

4.1. Degree

In the air navigation route network, each air navigation point is a node, and the degree of a node is the number of connections with other air navigation nodes. We depict the degree distribution for the fifteen countries in semi-log scale in the left-hand side of Fig. 3. The cumulative distribution $P(\geq k)$ is the probability that a randomly chosen air navigation node has a degree at least k . We can observe that most air navigation nodes have only a few connections with other nodes, while a few air navigation nodes have large number of connections. The nodes with high degree are often called hub nodes. The degree distributions show that air navigation route networks are sparsely connected for all fifteen countries. These findings are consistent with previous work on single countries, e.g. the Chinese air navigation route network ([Cai et al., 2012](#)) and the Italian air navigation route network ([Gurtner et al., 2014](#)).

4.1.1. Regression of degree distribution

However, in previous work ([Cai et al., 2012; Gurtner et al., 2014](#)), the authors stated that the degree distribution is approximated by an exponential function: $P(\geq k) \sim e^{\theta k}$. We also perform regression on the degree distribution with an exponential function for the fifteen countries. The average goodness of fit R^2 for the exponential function is 0.89. We find out that the degree distribution can be better fitted

by tetration $P(\geq k) = k^{\alpha k}$, with much higher R^2 (0.98). The tetration is an iterated exponentiation defined by Goodstein (Goodstein, 1947). Note that the tetration $k^{\alpha k}$ can also be written as $e^{\alpha k \ln(k)}$. Tetration can better approximate the cumulative distribution for nodes with smaller degree (less than 4). We think that in future research the degree distribution should rather be represented by tetration, instead of a simple exponential distribution.

In the current study, we use R^2 as the goodness of fit for the least square regression of cumulative distributions. It was recently shown that maximum likelihood estimate method could have a better quality of distribution fitting (Clauset *et al.*, 2009). It would be interesting to verify our results using maximum likelihood estimate method in the future.

4.1.2. Shannon diversity index of degree

In a dataset, a diversity index measures how many different types this dataset has, taking into account the evenness among the types. One popular measure is Shannon diversity index (Shannon, 1948): $H = -\sum_{i=1}^k p_i \log(p_i)$, where p_i denotes the proportion in group k . Lower values of Shannon index indicate more diversity; while higher values indicate less diversity (more evenness). Fig. 4 shows the Shannon index for air navigation nodes in the fifteen countries. We can observe that USA has the lowest Shannon index; while three European countries (France, Germany and Italy) have the highest Shannon index. This indicates that the connectivity of air navigation nodes in USA is more heterogeneous, comparing with the ones in these three countries.

4.1.3. Air navigation nodes with top degrees

For each of the fifteen countries, we select three air navigation nodes with highest degree (see Table 5). For each node, we show its ID, type, location of the node (city), the closest airport, and its distance to the closest airport. Around 70% of the top three air navigation nodes are within 10 km to their closest airports. We also observe that all the top three air navigation nodes in Argentina and Mexico (South America) are directly located at airports, while in France, Germany, and Italy (Europe) the top three air navigation nodes are relatively far away from airports. Note that there are a few hub air navigation nodes surrounded by several airports. For instance, there are many airports along Strait of Georgia around YVR (Vancouver, Canada). We find a similar situation for POU (Pingzhou, China), for WRB (Warburg, Germany), and for all the top three nodes in USA.

4.2. Distance strength

Distance strength is a metric proposed by Barrat *et. al.* to include geographic attributes in weighted complex networks (Barrat *et al.*, 2005). The distance strength of a node is defined as the cumulative distances of all the connections from (or to) the considered node (Barrat *et al.*, 2005). This metric is often referred to weighted degree in other context. We present the distance strength versus degree for the fifteen countries in log-log scale in the right-hand side of Fig. 3. The results show that topology and geography are strongly correlated. We can observe that distance strength distribution for all the fifteen countries exhibits a power law behaviour: $P(\geq s^d) \sim k^\beta$. The regression on distance strength distribution with a power law function is summarized in the third column in Table 3.

The fifteen countries can be split into two groups based on the exponent β . The countries in Group 1 (USA, Japan, France, Canada, Italy, Russia, and Germany) have an exponent β clearly larger than one: There exists a super-linear correlation between degree and distance strength. For Group 2 (Mexico, China, South Africa, Australia, Argentina, Congo, and Brazil) the exponent β is roughly one: There exists only a linear correlation between degree and distance strength. Note that Morocco is not quite representative (low $R^2 = 0.81$). This is because Morocco only has few nodes with higher degree. The distance strengths of these few nodes are outliers regarding the power law function. The existence of the outliers indicates large variation of the data and thus low goodness of fit for the power law function. For example, there is only one node with degree 10 (FES, near Saiss Airport) and its distance strength is 1,707 km, while there is also only one node with degree 12 (TNG, near Tangier Ibn Battouta Airport) and its distance strength is 571 km. These values are outliers according to the power law function.

It is interesting to note that the seven countries in Group 1 are industrial countries from North America, Europe, and Asia. The air navigation route networks in these countries are very heterogeneous and air navigation nodes with larger number of connections also have farther-reaching route segments. In other words, the distance per route segment scales up with the degree of its nodes. Especially, most countries in Group 1 have air navigation nodes with highest degree far away from airports. The countries from Group 2 are mostly emerging countries. Following the argumentation of Barrat et al. ([Barrat et al., 2005](#)), industrial countries in Group 1 have highly heterogeneous traffic networks, where hub nodes have large connectivities and long-distance connections. With increased economic development and traffic demand, emerging countries from Group 2 might consider to adapt the network structure from industrial countries in Group 1 when they expand their air navigation route systems.

4.3. Weighted betweenness centrality

Betweenness centrality indicates how often a node lies on the shortest paths between all other nodes. In this paper, we use the great circle distance between two air navigation nodes as edge weight. The cumulative distribution of weighted betweenness centrality for the fifteen countries in semi-log scale is presented in Fig. 6. The cumulative distribution of weighted betweenness centrality can be fitted by an exponential function: $P(\geq b) \sim e^{\gamma b}$. The results are summarized in the fourth column in Table 3. We can observe that among the fifteen countries, USA shows the fastest decaying behavior, followed by Russia and Australia; while Morocco has a relatively flat decaying behavior, followed by South Africa and Congo. A country with a bigger $|\gamma|$ has a higher percentage of nodes with smaller weighted betweenness centrality.

For example, USA has the biggest $|\gamma| = 468.5$, this indicates that most nodes have rather small weighted betweenness centrality in USA (ranging from 0 to 0.06). This can also explain why USA has the fastest decaying behaviour. Smaller weighted betweenness centrality indicates that most nodes are not part of shortest paths within the air navigation route network. Any removal of the nodes might not easily collapse the whole network. The network structure of USA is rather robust against node failures. On the contrary, the three countries from Africa (Morocco, South Africa, and Congo) have relatively smaller absolute value $|\gamma| \in [22, 28]$. This indicates that some nodes in these countries have rather high weighted betweenness centrality (up to 0.37). Higher weighted betweenness centrality means that these nodes are on many shortest paths within the air navigation route network. Any removal of these hub nodes might easily collapse the whole network. Thus, their network structure is vulnerable against targeted node failures. For example, the node HBV (Brits, South Africa) has the highest weighted betweenness centrality (0.37) and also the highest degree (9); the failure of this node could collapse a large part of the air navigation route system in South Africa.

4.4. Weighted closeness centrality

Closeness centrality is based on the inverse distance of each node to every other node in the network. With the great circle distance between two air navigation nodes as edge weight, we present the cumulative distribution of weighted closeness centrality for the fifteen countries in semi-log scale in Fig. 6. We observe that the decaying behaviour of the weighted closeness centrality distribution roughly matches the size of a country's geographic area. As expected, a country with larger geographic area decays faster in the distribution. We perform regression on weighted closeness centrality distribution for the fifteen countries. We find that the distribution follows a cubic polynomial function: $P(\geq c) = 1 - \delta c^3$. The results are summarized in the fifth column in Table 3.

In general, the node with the highest weighted closeness centrality usually locates in the geographic center of a country, if nodes are evenly distributed. For instance, the node with the highest weighted closeness centrality in Russia is ULGUN (Ozero Vachlor, Fedorovskiy), and the node with the highest weighted closeness centrality in Germany is MASEK (Schlitz, Bad Hersfeld). Both nodes are located in the geographic center. The same holds for Argentina, Brazil, Congo, France, Italy, Japan, and Mexico.

For the other six countries, we find that the node with the highest weighted closeness centrality is not located in the geographic center of the country. We discuss each of the six countries as follows. For Australia the node with the highest closeness centrality is BHI (Broken

Hill, New South Wales). The node is situated around 1,000 km south-east of the geographical center. We think that this can be explained by the fact that around 84% of Australia's population lives within 50 km of the coastline, with the eastern part (New South Wales, Victoria, Queensland) covering more than 70% of the overall population. Therefore, 1) the traffic demand is much higher than in other regions of Australia and 2) the air traffic network is much better developed. For Canada the node with the highest weighted closeness centrality is YDN (Dauphin, Manitoba), located around 300 km from the southern border to the USA. The whole network is shifted towards the south, since 72% of the population is living in a 150 km wide strip along the south border to USA. Again, the traffic demand in this area is much higher than in the central and northern part of Canada. In China, the center of the network is shifted towards the east; the node with the highest weighted closeness centrality is ZHO (Zhoushou, Henan). The east coast of China, with its large seaports, was one of the first developed regions in China, and is the market that has been the focus of most international manufacturing activity to date. We find a similar shift towards the coastal regions in Morocco. The node with the highest closeness centrality is SLK (between Berrechid and Settat), in the western part of the country, because major cities are along the western coastline. For South Africa the node with the highest closeness centrality is HGV (50 km south of Johannesburg). The Tshwane metropolitan area (including Johannesburg and Pretoria) is still the major center for economical activities in South Africa, followed by Cape Town on the west-coast. Finally, for USA the node with the highest closeness centrality is STL (St. Louis, Missouri). Here, the distance between geographical center and location of STL is not as far as in the previous countries, yet still recognizable with approximately 500 km. The explanation we find is that a large part of the west is only covered by few nodes, for instance, Nevada desert. Thus, the node with the highest weighted closeness centrality is shifted slightly towards the eastern part of the country.

We conclude that the node with the highest weighted closeness centrality is not necessarily in the geographic center of a country. Rather, it is close to the center of gravity with regard to the population(Grasland and Madelin, 2001). This is because the node density of the network is strongly correlated with the regional population density in a country. The traffic demand is higher in the regions with higher population density, more nodes are needed for the air traffic network.

4.5. Edge length distribution

We show the mean and standard deviation of the edge length for the fifteen countries in the left-hand side of Fig. 7. Canada and USA have the shortest route segment (around 0.38 km) and the longest one (1,653.05 km for Canada and 1,152.14 km for USA). Germany has the smallest mean route segment (40.93 km) and the smallest standard deviation (32.85 km); while Canada has the largest mean route segment (218.03 km) and the largest standard deviation (312.92 km). Canada is the second largest country regarding geographic area in the world, its long route segment and large variation is because 72% of the population is concentrated in a 150 km wide strip along the south border to USA. The vast majority of Canada is inhabited and the populated areas are far away from each other. Therefore, only a few air navigation nodes with long-distance connections are located in the north part of Canada. Although Russia has the largest geographic area, we find that air navigation nodes with short-distance connections are rather evenly distributed among the whole country.

We present edge length distribution for the fifteen countries in semi-log scale in the right-hand side of Fig. 7. We can observe that the edge length distribution follows an exponential function: $P(\geq l) \sim e^{\lambda l}$. We show the regression for the edge length distribution in the sixth column in Table 3.

4.6. Clustering analysis

Based on the coefficients of the regressions ($\alpha, \beta, \gamma, \delta$, and λ in Table 3), we compute the hierarchical clustering for the fifteen countries and the results are presented in Fig. 8. Hierarchical clustering is one way to build a hierarchy of clusters, based a measure of dissimilarity between sets of observations. Euclidean distance is one of the most common used measures: The distance between two points $a = (a_1, a_2, \dots, a_n)$ and $b = (b_1, b_2, \dots, b_n)$ is defined as $\|a - b\|_2 = \sqrt{\sum_i^n (a_i - b_i)^2}$.

We find out that the most similar countries are France and Italy. In most cases, the countries with similar geographical features are grouped together. For example, Canada and Russia belong to one cluster. Note that Argentina and China are in one cluster, we think this is because the network structures of China are rather unbalanced and shifted towards east. The shifted structures make China more similar to the shape in Argentina. The cluster of Japan and Morocco might be because they both have long and narrow geographical shapes and the network structures are strongly geographically constrained.

Note that the grouping of different countries depends on the threshold of the Euclidean distance in the clustering. For the purpose of better explanation, we can imagine a hypothetically horizontal line in Fig. 8. Different thresholds would lead to different country clusters. Therefore, the clustering of different country groups is dynamic.

5. Conclusion

In our work, we analysed air navigation route system for fifteen countries as a complex network. Our major findings are summarized in Table 3. We find out that air navigation route networks for all fifteen countries are heterogeneous: most air navigation nodes have only a few connections with other nodes and a few hub air navigation nodes have large number of connections. The degree distribution is better fitted by tetration than exponential function, as believed in previous work. The distance strength exhibits power law behaviour. Hub nodes in industrial countries have large connectivities and long-distance connections. With increased economic development and traffic demand, our analysis suggests that emerging countries could consider to adapt the network structure from industrial countries when they expand their air navigation route system. The weighted betweenness centrality shows an exponential behaviour. The weighted closeness centrality follows a cubic polynomial function for all fifteen countries. The edge length distribution for all fifteen countries has an exponential behaviour. The hierarchical clustering based on the regression coefficients shows that the countries with similar geographical features are clustered together. Our findings help to understand the current structure and emerged new topologies of the air navigation route system.

In the current study, we focus on topological properties of air navigation route network. In future research, we will take into account traffic data. The network analysis of the air navigation route system can also be used to identify potential bottlenecks of air traffic in future, for example, an air navigation node with high degree or betweenness centrality is most likely to be congested (Cook *et al.*, 2012, 2013). It is crucial to validate this proposition with empirical data in future. Further research could also focus on the dynamics in the air navigation route system, such as how delay is propagated in the network (Pyrgiotis *et al.*, 2013).

Acknowledgment

The authors would like to thank EUROCONTROL for providing the data in this study. The authors also thank Stephan Lehner for the discussions.

REFERENCES

- Arranz A, Etxebarria I, Regidor C and Escribano B (2013) Intelligent modelling of the air transport network. In *Third SESAR Innovation Days*.
- Bagler G (2008) Analysis of the airport network of india as a complex weighted network. *Physica A: Statistical Mechanics and its Applications* **387(12)**: 2972–2980.
- Barrat A, Barthélemy M and Vespignani A (2005) The effects of spatial constraints on the evolution of weighted complex networks. *Journal of Statistical Mechanics: Theory and Experiment* **2005(05)**: P05003.
- Barrat A, Barthélémy M, Pastor-Satorras R and Vespignani A (2004) The architecture of complex weighted networks. *PNAS* **101(11)**: 3747–3752.
- Barthélemy M (2011) Spatial networks. *Physics Reports* **499(1)**: 1–101, <http://arxiv.org/abs/1010.0302>.
- Bastian M, Heymann S and Jacomy M (2009) Gephi: An open source software for exploring and manipulating networks.

-
- Bonnefoy P (2008) *Scalability of the air transportation system and development of multi-airport systems: A worldwide perspective*. PhD thesis, Massachusetts Institute of Technology.
- Cai K, Zhang J, Du W and Cao X (2012) Analysis of the chinese air route network as a complex network. *Chinese Physics B* **21**(2): 28903, [10.1088/1674-1056/21/2/028903](https://doi.org/10.1088/1674-1056/21/2/028903).
- Cardillo A and et. al. (2013) Emergence of network features from multiplexity. *Scientific Reports* **3**(1344): 1–6.
- Clauset A, Shalizi C and Newman M (2009) Power-law distributions in empirical data. *SIAM Review* **51**(4): 661–703.
- Cook A, Tanner G, Cristóbal S and Zanin M (2013) New perspectives for air transport performance. In *Third SESAR Innovation Days*.
- Cook A, Tanner G, Cristbal S and Zanin M (2012) Passenger-oriented enhanced metrics. In *Second SESAR Innovation Days*.
- Cordell B (2006) Enroute routes. http://www.aixm.aero/gallery/content/public/design_review.
- DeLaurentis D, Han EP and Kotegawa T (2008) Network-theoretic approach for analyzing connectivity in air transportation networks. *Journal of Aircraft* **45**(5): 1669–1679, [10.2514/1.35244](https://doi.org/10.2514/1.35244).
- Djokic J, Lorenz B and Fricke H (2008) Atc complexity as workload and safety driver. In *3rd International Conferences on Research in Air Transportation (ICRAT)*.
- European Commission (2011) Flightpath 2050 europe's vision for aviation.
<http://ec.europa.eu/transport/modes/air/doc/flightpath2050.pdf>.
- Fleurquin P, Ramasco J and Eguiluz V (2013a) Data-driven modeling of systemic delay propagation under severe meteorological conditions. In *10th USA/Europe ATM Seminar*, <http://arxiv.org/ftp/arxiv/papers/1308/1308.0438.pdf>.
- Fleurquin P, Ramasco JJ and Eguiluz VM (2013b) Systemic delay propagation in the us airport network. *Scientific reports* **3**.
- Freeman LC (1978) Centrality in social networks conceptual clarification. *Social Networks* **1**(3): 215 – 239, [http://dx.doi.org/10.1016/0378-8738\(78\)90021-7](https://doi.org/10.1016/0378-8738(78)90021-7).
- Goodstein RL (1947) Transfinite ordinals in recursive number theory. *J. Symb. Log.* **12**(4): 123–129.
- Grasland C and Madelin M (2001) The unequal distribution of population and wealth in the world. *Population & Society* **36**(8): 1–4.
- Guimera R, Mossa S, Turtschi A and Amaral LAN (2005) The worldwide air transportation network: Anomalous centrality, community structure, and cities global roles. *PNAS* **102**(22): 7794–7799, [10.1073/pnas.0407994102](https://doi.org/10.1073/pnas.0407994102).
- Gurtner G, Vitali S, Cipolla M, Lillo F, Mantegna RN, Miccich S and Pozzi S (2014) Multi-scale analysis of the european airspace using network community detection. *PLoS ONE* **9**(5): e94414, [10.1371/journal.pone.0094414](https://doi.org/10.1371/journal.pone.0094414).
- Holmes B and Scott J (2004) *Transportation network topologies. Technical report*, NASA Langley Research Center.
- Janić M (2007) *The sustainability of air transportation: A quantitative analysis and assessment*. Ashgate.
- Kotegawa T, Fry D, DeLaurentis D and Puchaty E (2014) Impact of service network topology on air transportation efficiency. *Transportation Research Part C: Emerging Technologies* **40**(0): 231 – 250, [http://dx.doi.org/10.1016/j.trc.2013.11.016](https://doi.org/10.1016/j.trc.2013.11.016).
- Kotegawa T, Fry D, Puchaty E and DeLaurentis D (2010) Impact of us airline network topology on air transportation efficiency. In *4th ICRAT*.
- Kurant M and Thiran P (2006) Extraction and analysis of traffic and topologies of transportation networks. *Phys. Rev. E* **74**: 036114, [10.1103/PhysRevE.74.036114](https://doi.org/10.1103/PhysRevE.74.036114).
- Lehner S (2013) Separate yet interdependent networks: The structure and function of european air transport. In *Complex Networks IV*, vol. 476, Springer, pp. 109–120.
- Lillo F, Micciche S, Mantegna RN, Beato V and Pozzi S (2011) Elsa project: Toward a complex network approach to atm delays analysis. In *First SESAR Innovation Days*.
- Newman MEJ (2004) Analysis of weighted networks. *Phys. Rev. E* **70**(5): 056131.
- Pfeifer R (2013) Graphs and networks. <http://www.ifi.uzh.ch/ailab/teaching/formalmethods2013>.

- Pyrgiotis N, Malone KM and Odoni A (2013) Modelling delay propagation within an airport network. *Transportation Research Part C: Emerging Technologies* **27(0)**: 60–75.
- Reggiani A, Nijkamp P and Cento A (2010) Connectivity and concentration in airline networks: A complexity analysis of lufthansa's network. *European Journal of Information Systems* **19(4)**: 449–461.
- Rodionova O, Sbihi M, Delahaye D and Mongeau M (2012) Optimization of aircraft trajectories in north atlantic oceanic airspace. In *5th International Conferences on Research in Air Transportation (ICRAT)*.
- Shannon CE (1948) A mathematical theory of communication. *Bell System Technical Journal* **27(3)**: 379–423, [10.1002/j.1538-7305.1948.tb01338.x](https://doi.org/10.1002/j.1538-7305.1948.tb01338.x).
- Sinnott RW (1984) Virtues of the Haversine. *Sky and Telescope* **68(2)**: 159+.
- South A (2011) rworldmap: a new r package for mapping global data. *R Journal* **3(1)**: 35–43.
- Sun X and Wandelt S (2014) Network similarity analysis of air navigation route systems. *Transportation Research Part E: Logistics and Transportation Review* **70(0)**: 416–434, [http://dx.doi.org/10.1016/j.tre.2014.08.005](https://dx.doi.org/10.1016/j.tre.2014.08.005).
- Sun X, Wandelt S and Linke F (2014) Temporal evolution analysis of the european air transportation system: air navigation route network and airport network. *Transportmetrica B: Transport Dynamics* **0(0)**: 1–16, [10.1080/21680566.2014.960504](https://dx.doi.org/10.1080/21680566.2014.960504), [http://dx.doi.org/10.1080/21680566.2014.960504](https://dx.doi.org/10.1080/21680566.2014.960504).
- Vitali S and et. al. (2012) Statistical regularities in atm: Network properties, trajectory deviations and delays. In *Second SESAR Innovation Days*.
- Wang L, Donohue G, Hoffman K, Sherry L and Osegueda-Lohr R (2008) Analysis of air transportation for the new york metroplex: Summer 2007. In *3rd ICRAT*.

List of figures

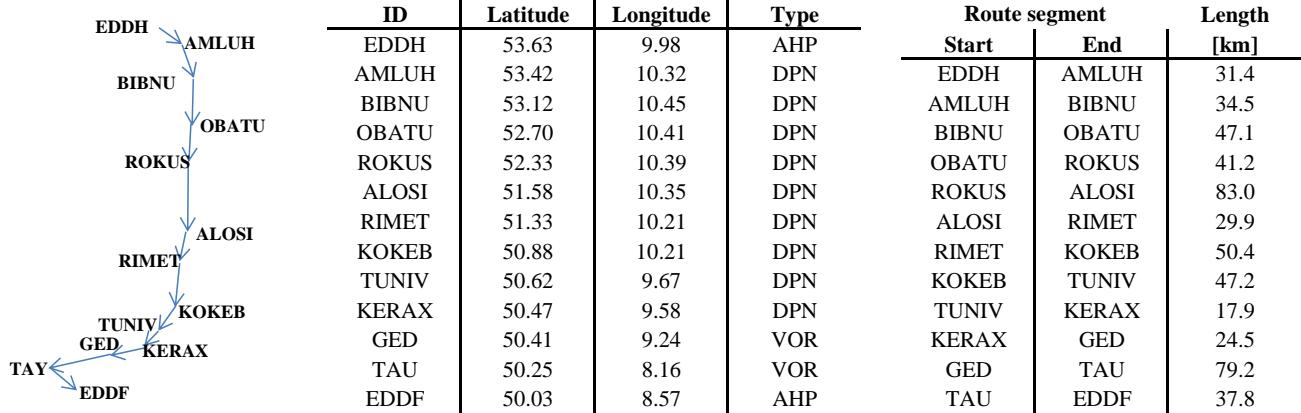


Figure 1. Left: A visualization of a single air navigation route from Hamburg (EDDH) to Frankfurt (EDDF); Center: Information on the navigation points on that route. Right: Information on the route segments.

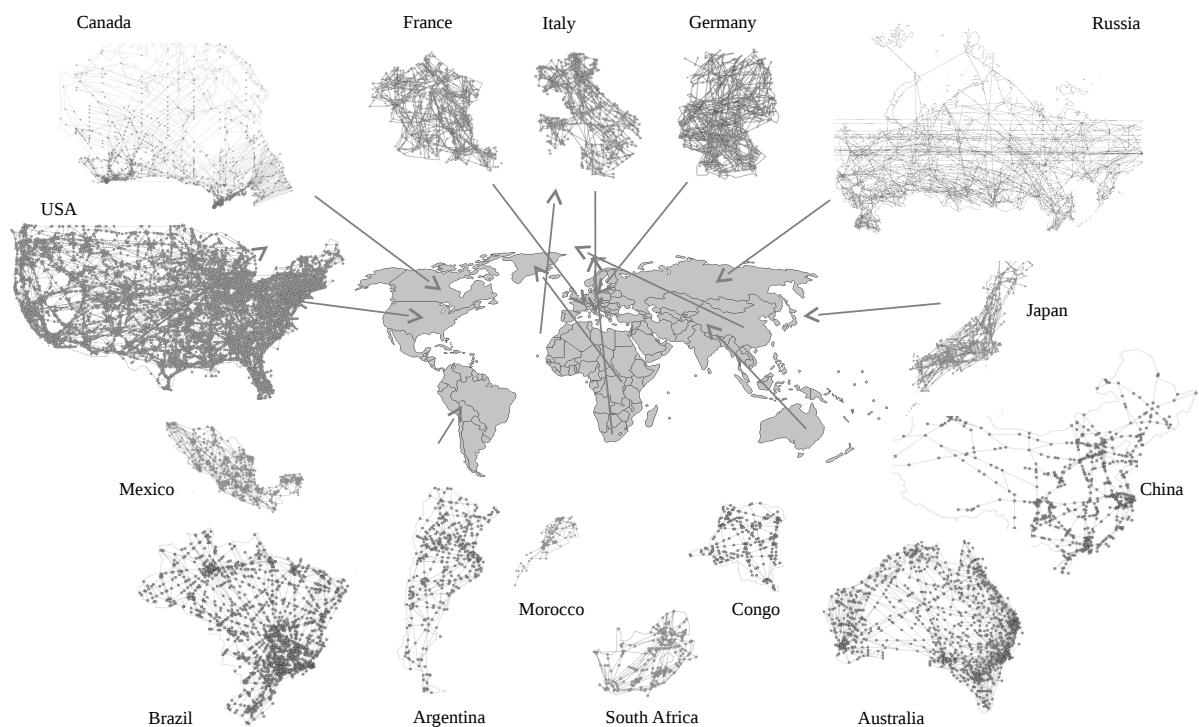


Figure 2. Air navigation route networks for the fifteen countries. The figure has been created with the network analysis and visualization software Gephi (Bastian et al., 2009), together with the free Worldmap visualization package rworldmap (South, 2011). The route network data for these figures were obtained through the European Aeronautical Information Service Database (EAD) and provided by Eurocontrol. The data is effective on 18 October 2012.

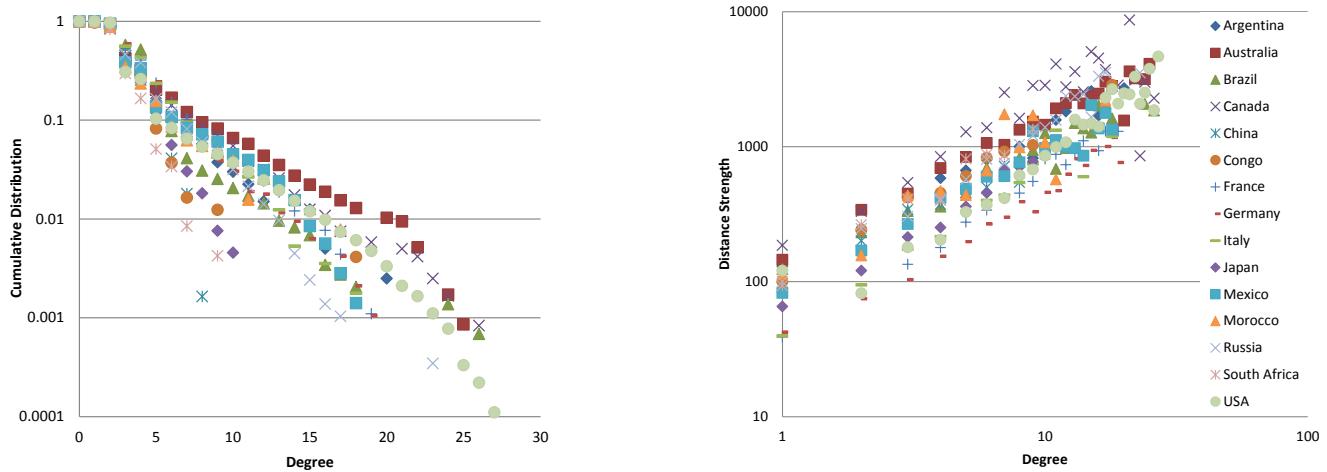


Figure 3. Degree distribution (left) and distance strength (right) of fifteen air navigation route networks

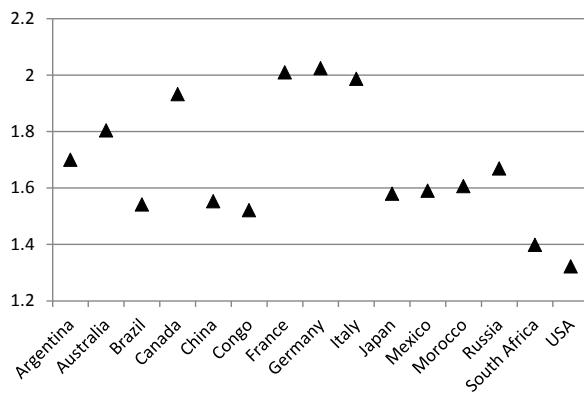


Figure 4. Shannon diversity index for air navigation nodes in the fifteen countries

Figure 5. Air navigation nodes with highest degree in the fifteen countries

List of tables

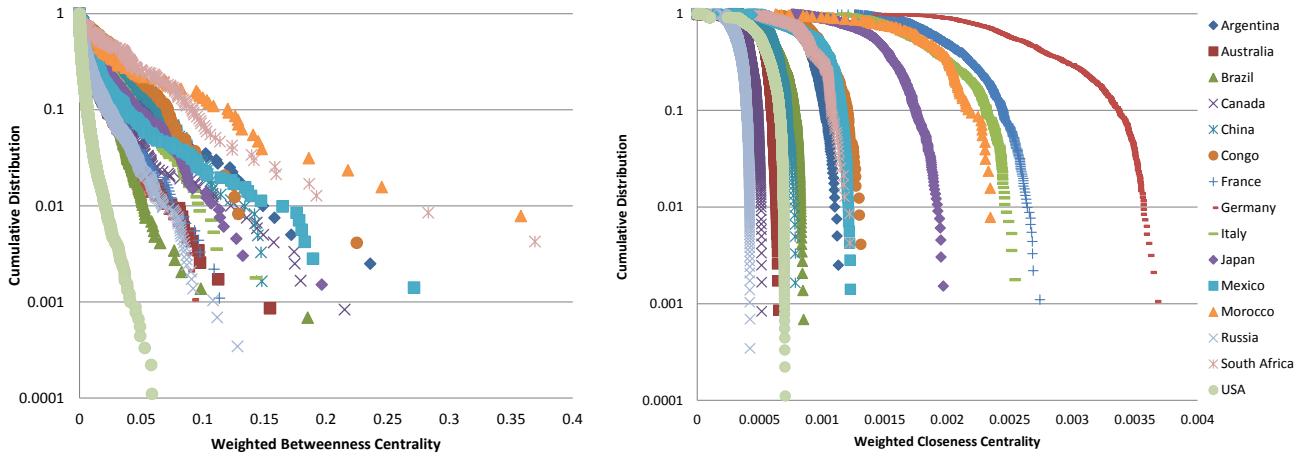


Figure 6. Weighted betweenness centrality (left) and weighted closeness centrality (right) of fifteen air navigation route networks

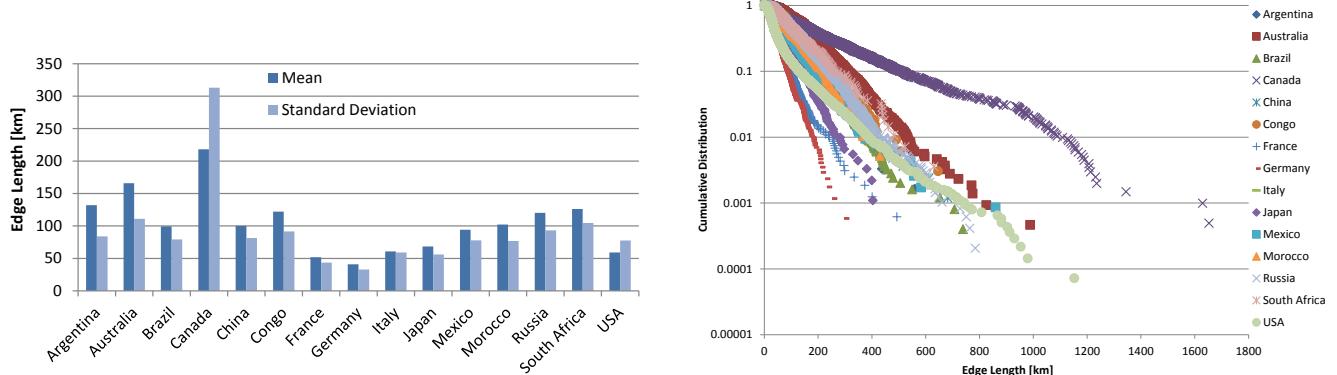


Figure 7. Edge length mean (left) and distribution (right) of fifteen air navigation route networks

Table 1. Geography and network overview for the fifteen countries

Continent	Country	Area	Population	Pop./Area	Nodes	Nodes%	Edges	Edges%
Africa	Congo	2,345,410	70,916,439	30.24	243	0.63%	328	0.51%
Africa	Morocco	446,550	31,627,428	70.83	128	0.33%	193	0.30%
Africa	South Africa	1,219,912	49,000,000	40.17	236	0.61%	281	0.44%
Asia	China	9,596,960	1,330,044,000	138.59	610	1.59%	837	1.30%
Asia	Japan	377,835	127,288,000	336.89	660	1.72%	912	1.42%
Asia	Russia	17,100,000	140,702,000	8.23	2,904	7.55%	4,868	7.59%
Europe	France	547,030	64,768,389	118.40	912	2.37%	1,619	2.52%
Europe	Germany	357,021	81,802,257	229.12	952	2.47%	1,718	2.68%
Europe	Italy	301,230	60,340,328	200.31	565	1.47%	1,009	1.57%
North America	Canada	9,984,670	33,679,000	3.37	1,202	3.12%	2,026	3.16%
North America	USA	9,629,091	310,232,863	32.22	9,050	23.52%	13,847	21.58%
Oceania	Australia	7,686,850	21,515,754	2.80	1,167	3.03%	2,154	3.36%
South America	Argentina	2,766,890	41,343,201	14.94	400	1.04%	607	0.95%
South America	Brazil	8,511,965	201,103,330	23.63	1,461	3.80%	2,493	3.89%
South America	Mexico	1,972,550	112,468,855	57.02	710	1.85%	1,164	1.81%

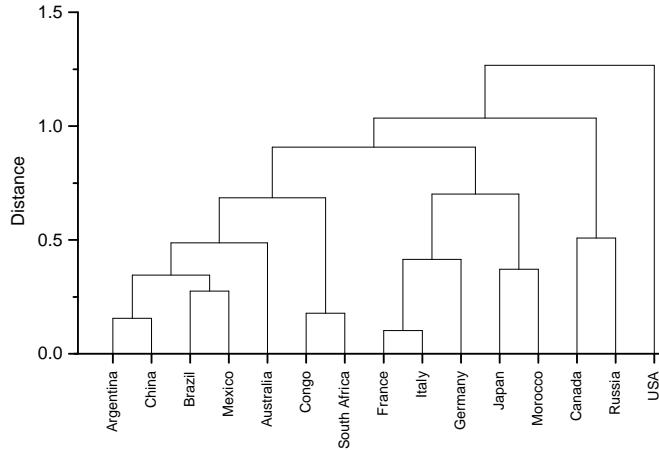


Figure 8. Hierarchical clustering for the fifteen countries

Table 2. Five metrics for the undirected and weighted air navigation route network

Metrics	Equation	Interpretation
Degree	$k_i = \sum_j a_{ij}$	where a_{ij} is the connection between air navigation node i and air navigation node j : $a_{ij} = 1$ when there is a connection existing; $a_{ij} = 0$ otherwise. This metric refers to the number of connections with other air navigation nodes.
Distance strength	$s_i^d = \sum d_{ij}$	where d_{ij} is the great circle distance between air navigation node i and air navigation node j . This metric is proposed by Barrat et al. (Barrat et al., 2005) and s_i^d refers to the cumulative distances of all connections with air navigation node i .
Weighted betweenness centrality	$b_i^w = \sum_{s \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}$	where σ_{st} is the number of shortest paths going from air navigation node s to air navigation node t ; $\sigma_{st}(i)$ is the number of shortest paths going from air navigation node s to air navigation node t and passing through air navigation node i . This metric is proposed by Freeman (Freeman, 1978) and it indicates the number of shortest paths going through an air navigation node, with the great circle distance as edge weight.
Weighted closeness centrality	$c_i^w = \frac{\sum_{j \in N, j \neq i} \sigma_{ij}^w}{(n-1)}$	where N is the set of all nodes in the network, n is the number of nodes, σ_{ij}^w is the shortest path between air navigation node i and air navigation node j , with the great circle distance as edge weight. This metric is proposed by Freeman (Freeman, 1978) and it is based on the inverse of the great circle distance for one air navigation node to every other one in the network.
Edge length distribution	—	Edge length in the air navigation route network is the great circle distance between two air navigation nodes.

Table 3. Summary of regressions on five metrics in fifteen air navigation route networks, where the minimum and maximum values for the five regressed parameters are highlighted.

Metric	Degree distribution		Distance strength VS. degree		Weighted betweenness distribution		Weighted closeness distribution		Edge length distribution	
Function type	Tetration $P(\geq k) = k^{\alpha k}$		Power law $P(\geq s^d) \sim k^\beta$		Exponential $P(\geq b) \sim e^{\gamma b}$		Cubic polynomial $P(\geq c) = 1 - \delta c^3$		Exponential $P(\geq l) \sim e^{\lambda l}$	
Country	α	R^2	β	R^2	γ	R^2	δ	R^2	λ	R^2
Argentina	-0.2317	0.98	1.015	0.99	-41.00	0.95	9.0E+08	0.98	-0.0110	0.98
Australia	-0.1654	0.97	1.024	0.98	-143.60	0.97	4.7E+09	0.98	-0.0071	0.97
Brazil	-0.1854	0.97	0.947	0.94	-73.92	0.99	1.9E+09	0.98	-0.0126	0.99
Canada	-0.2088	0.99	1.326	0.90	-140.80	0.96	9.0E+09	0.97	-0.0049	0.99
China	-0.2462	0.97	1.038	0.95	-33.82	0.99	2.6E+09	0.98	-0.0124	0.99
Congo	-0.2980	0.97	1.003	0.97	-28.52	0.99	6.4E+08	0.94	-0.0105	0.96
France	-0.1676	0.99	1.337	0.96	-74.49	0.98	7.2E+07	0.97	-0.0240	0.98
Germany	-0.1614	0.99	1.187	0.96	-89.15	0.99	2.5E+07	0.95	-0.0325	0.98
Italy	-0.1724	0.99	1.293	0.98	-55.69	0.98	7.6E+07	0.94	-0.0236	0.98
Japan	-0.2567	0.98	1.432	0.98	-51.14	0.98	1.9E+08	0.97	-0.0190	0.99
Mexico	-0.1978	0.97	1.071	0.98	-79.73	0.99	5.7E+08	0.99	-0.0139	0.97
Morocco	-0.2307	0.98	1.331	0.81	-23.72	0.96	8.6E+07	0.98	-0.0118	0.99
Russia	-0.1846	0.98	1.277	0.98	-154.30	0.97	1.6E+10	0.99	-0.0100	0.99
South Africa	-0.3247	0.98	1.032	0.96	-22.00	0.99	7.4E+08	0.97	-0.0095	0.98
USA	-0.2431	0.96	1.451	0.96	-468.50	0.98	3.2E+09	0.99	-0.0253	0.99