UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

DEPARTEMENT BEDRYFS- EN SISTEEMINGENIEURSWESE
DEPARTMENT OF INDUSTRIAL AND SYSTEMS ENGINEERING

FRONT PAGE FOR FINAL PROJECT DOCUMENT (BPJ 420) - 2017

Information with regards to the mini-dissertation

Title	Airport route development decision support tool for Cape Town Air Access
Author	Robertson, K.M.
Student number	13103408
Supervisor	Viljoen, N.M.
Date	2017/09/28
Keywords	Accessibility, Air transport network, Betweenness centrality, Complex network theory, Efficiency, Simulation, Vulnerability, Weighted clustering coefficient
Abstract	The connectedness and accessibility of Cape Town International Airport is critical for the economic growth and development of the Western Cape. This study aims to assist the decision-makers at Cape Town Air Access with their route development decisions. Complex Network Theory is used to run simulations under various scenarios to evaluate the impact of route developments on CTIA. Three metrics (betweenness centrality, weighted clustering coefficient and efficiency) are used to calculate the accessibility of CTIA when new flights are added to the network, or the capacity on existing flights is increased or decreased. This method gives the decision-makers a network view of route changes, which is different to what is currently available and thus may offer greater insight into the network.
Category	Operations Research
Sensitive Parts	

Declaration

1. I understand what plagiarism is and I am aware of the University's policy in this regard.
2. I declare that this is my own original work
3. Where other people's work has been used (either from a printed source, internet or any other source) this has been carefully acknowledged and referenced in accordance with departmental requirements
4. I have not used another student's past work to hand in as my own
5. I have not allowed and will not allow, anyone to copy my work with the intention of handing it in as his/her own work

Signature

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Airport route development decision support tool for Cape Town Air Access

Kyle M. Robertson

13103408

Submitted in partial fulfilment of the requirements for the degree of

Bachelors of Industrial and Systems Engineering

September 28, 2017

Contents

List of Figures

List of Tables

Acronyms

AAN Australian Airport Network

ACSA Airports Company South Africa

ATN Air Transport Network

CAN Chinese Airport Network

CNT Complex Network Theory

CTAA Cape Town Air Access

CTIA Cape Town International Airport

DOT Department of Transport

DST Decision Support Tool

GDP Gross Domestic Product

JFK John F. Kennedy International Airport

LAX Los Angeles International Airport

LTTS Long-Term Turnaround Strategy

ORTIA O.R. Tambo International Airport

PMI Purchasing Managers' Index

RPK Revenue Passenger Kilometres

SAA South African Airways

USA United States of America

WAN World Airport Network

Executive Summary

The connectedness and accessibility of Cape Town International Airport (CTIA) is critical for the economic growth and development of the Western Cape. Cape Town Air Access (CTAA) is responsible for bolstering the long-term growth and sustainability of Cape Town's air access by making decisions on which routes to develop and which existing routes to expand. These decisions are multi-faceted and have wide-ranging impacts on the Air Transport Network (ATN). This study aims to assist the decision-makers at CTAA by creating a model with a network view of their route development decisions.

Complex Network Theory (CNT) is used to create the model, which has been applied in many fields to study both the topology and dynamic nature of complex networks. In this study, it is used to model CTIA's flight network and in turn measure its accessibility under various scenarios. Three scenarios are tested: Increasing the capacity on key flights, decreasing the capacity on key flights and introducing new flights to the network. The key flights are identified by calculating the *vulnerability* of airports in the network, which measures the impact that removing an airport has on the accessibility of CTIA. The new flights added to the network are the busiest international airports in each of the 5 strategic countries identified by CTAA (Angola, Canada, Ethiopia, Kenya and the United States of America (USA)). Two heuristics are employed in the model to account for the number of seats stolen as a result of an increase in capacity on a direct flight. The *Proportionality heuristic* 'steals' seats from connecting flights with the same proportion as passengers currently using each route to travel to CTIA. The *Switching behaviour heuristic* uses flight-cost and travel-time data between regions to predict the number of passengers to switch to the direct flight as opposed to using the connecting flight. Three metrics (betweenness centrality, weighted clustering coefficient and efficiency) are used to calculate the accessibility of CTIA when new flights are added to the network, or the capacity on existing flights is increased or decreased. These metrics translate the word definition of *accessibility* (flexibility, time-efficiency and cost-efficiency of travel) into a measurable quantity. The research question is: *"Can a CNT approach assist decision makers in evaluating the impact of route changes on the Western Cape's air access?"*. Positive feedback from the company was received regarding the project's usefulness. The project manager described it as "Novel and worthwhile for what we are doing at CTAA", and "The methodology and the way it is structured makes a lot of sense". He also gave constructive criticism regarding the results of the analysis, which need to be more suitable for audiences such as politicians and the CTAA steering committee.

Based on the results of the simulations, Schiphol Airport in Amsterdam and Franz Josef Strauss Airport in Munich should be expanded. This challenges the current thinking at CTAA which focuses on developing and expanding smaller routes. In terms of developing new direct flights, the results suggest that Miami International, Edmonton International and John F. Kennedy International Airport (JFK) should be developed.

Chapter 1

Introduction

When Bartolomeu Dias found that Europe could trade directly with Asia by sailing around the southern tip of Africa in 1487, it was a watershed moment for globalisation. In the 20$^{\text{th}}$ century, another watershed moment for global trade occurred. The commercial passenger aircraft completely redefined global travel, as large volumes of people could travel to almost any part of the globe within a day. The world became more accessible than ever.

In modern times, accessibility is important because the growth of a country's economy and its ability to compete globally is greatly affected by its air transportation network [16, 27]. Countries are therefore striving to persuade airlines to fly directly to their airports, as this is the fastest and most convenient way to travel. The speed and convenience of a direct flight over multiple flights and lengthy layovers is much more attractive to tourists and business-people alike. The thriving province of the Western Cape has realised this.

1.1 The Western Cape

The province has grown tremendously in both tourism and business in the last few decades. It is one of the most popular tourist destinations in the world, with attractions such as the Cape Winelands, mountain and wilderness reserves, world-famous historical sites and a vast stretch of coastline. It is a key economic contributor to the country as a whole, and is crucial in terms of food security and trade access through its two ports in Saldanha and Cape Town. Despite its considerable influence in South Africa, the province is far behind Gauteng in terms of air access. O.R. Tambo International Airport (ORTIA) in Johannesburg is the main hub of the country. It handles most of the direct international flights and, as a result, gets all the benefits of being more connected to the rest of the world. This advantage is very difficult to overturn because of ORTIA's existing relationships with major airlines and its larger airport infrastructure.

1.2 Company background

CTAA exists to improve the accessibility of CTIA and therefore close the gap on Gauteng. It is a division of Wesgro, whose aim is to promote tourism, trade and investment in the Western Cape. The company works in conjunction with private sector, government and Airports Company South Africa (ACSA) to expand existing routes and maintain relationships with airlines currently flying to CTIA. They also develop new routes to CTIA, which requires significant resources because airlines need to have an incentive to fly somewhere — be it cash incentives, subsidies on airport fees or partnerships and campaigns to boost passenger volumes. After all, making a profit is the main goal, therefore airlines will not be inclined to invest in new developments if there is high risk or uncertainty. The airline industry is also a highly competitive environment, therefore companies will only pursue the most attractive opportunities.

The effort required to market the route also costs money, therefore, CTAA needs to make good decisions regarding which airlines and routes to incentivise. These decisions are multi-faceted, taking into account economic feasibility, the competitive landscape in the airline industry and political and strategic alignment for the Western Cape. The right balance of objectives must be struck to bolster the long-term growth and sustainability of Cape Town's air access. With this goal of long-term sustainability, it is critical to understand how route development decisions today could induce unanticipated changes in the World Airport Network (WAN) tomorrow. Unfortunately, building a clear view of these complex interdependencies for decision-makers to appreciate is not trivial. Currently, this is a significant gap in the CTAA decision-making process.

1.3 Research design

The problem being addressed is that CTAA needs a Decision Support Tool (DST) to help them understand the network impacts of their route development decisions so that their decisions are better informed. The data sources available provide the direct seat capacity between every airport pair in the world, the number of passengers flying business class versus economy class on these flights, and the proportion of passengers flying to CTIA on different routes. A hypothetical case of the routing data is shown in Figure 1.1.

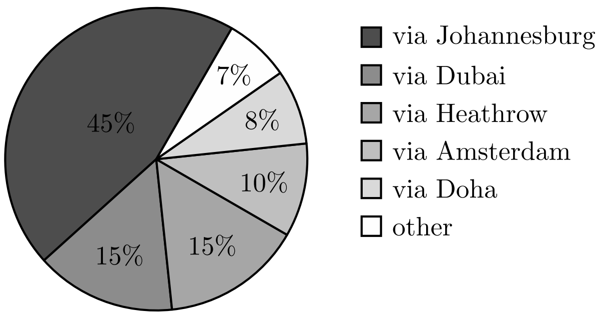


Figure 1.1: Hypothetical case of the proportion of different routes taken by passengers flying to CTIA from JFK.

This figure shows the percentage of passengers flying to CTIA from JFK on each possible route. If the direct flight is developed, capacity will be 'stolen' from the routes via Johannesburg, Dubai, Heathrow, Amsterdam, Doha and others. The number of 'stolen' seats is not exact, therefore heuristics are required to approximate them. The two types of heuristics that will be used to account for these 'stolen' seats are briefly described below (a more detailed description can be found in Section 4.3):

Proportionality heuristic: The number of seats stolen from a particular connecting route is directly proportional to the number of passengers using that route. Thus, more seat capacity will be stolen from the JFK - ORTIA - CTIA route than the JFK - Dubai - CTIA route because more people currently use ORTIA than Dubai as a connecting airport.

Switching behaviour heuristic: The first heuristic is quite simplistic in that it does not take into account real-world factors such as cost and travel time that may affect people's decision to switch to the direct flight or remain on the connecting flight. Cost and travel time are factored in by estimating the number of people to switch based on the trade-off between the increase in price and the decrease in travel time.

These heuristics will be used in conjunction with a field of science called CNT. This field is a combination of statistical mechanics and graph theory, therefore it covers a broad range of disciplines. It has been effectively used to analyse many kinds of extremely complicated networks, and so it is the natural choice when analysing a large airport network such as this. The proposition is therefore that by using CNT along with the two aforementioned heuristics, a deeper insight into the flight network will be offered than what is currently available. Using these heuristics in combination with CNT to solve the problem leads to the research question: *"Can a CNT approach assist decision makers in evaluating the impact of route changes on the Western Cape's air access?"* To address this question, the network will be modelled, different route development scenarios will be evaluated with CNT and feedback from the company will determine whether the model is useful or not.

1.4 Research methodology

The methodology will follow that of Manson [21], which is a useful framework to use when addressing a design research problem. *Awareness of the problem* is the first step, which has been fulfilled in the preceding sections of this report. Design is a creative, trial-and-error process in which many iterations of the solution are run through before the best solution is achieved [12]. The *suggestion of a solution* is therefore only a general suggestion of the solution to the problem because the final 'artefact' that actually solves the problem may be quite different. The suggested solution to improving decision-making at CTAA is to use a CNT approach. Two 'artefacts' will be built: Firstly, the network will be modelled using the `igraph` package in R, which is useful for both network analysis and network visualisation [8]. Secondly, the model will be used to test various scenarios in route expansion and development while applying the heuristics outlined in Section 1.3.

The *development* step of the process begins once a solution has been suggested. The model will be both descriptive and prescriptive. It is descriptive in the sense that the decision-makers at CTAA will gain an appreciation of the network impacts of route changes, and prescriptive as it "seeks to prescribe ways to do things more effectively" [21]. In this case, it will prescribe which routes will have the most significant impact on CTIA's accessibility. In order to measure this accessibility, four CNT metrics have been identified as applicable tools (Section 4.2 explains these in detail): *Efficiency* measures how efficient a network is at linking nodes. In an airport network, nodes represent airports and the link between airports is the flights between them. Flights can be described by a variety of data-types, from travel time, ticket prices, seat capacity or a combination of these. The type of data chosen to describe links depends on the data available and the research question. The *vulnerability* of a node determines how critical it is to both the efficiency and connectivity of a network. For example, if one were to remove a large airport such as Heathrow from the WAN, the efficiency and connectedness of the entire network would be greatly affected. Heathrow is therefore described as a vulnerable airport. A *clustering coefficient* measures how cohesive an airport's immediate neighbourhood is by determining how connected its neighbours are to each other. *Betweenness centrality* measures the degree to which an airport links other airports to each other, as well as its importance to the network.

Once these metrics have been established, the scenarios will be built and tested in the following way:

Scenario 1 - Increasing the seat capacity on strategic flights: Part of the company's mandate is to expand existing routes. Expansion occurs in various ways, such as increasing flight frequency, changing the aircraft type or making seasonal flights into year-round flights. The effect of expansion on key flights will be measured.

Scenario 2 - Decreasing the seat capacity on strategic flights: Similar to Scenario 1, but the capacity on key flights will be decreased to analyse the effect.

Scenario 3 - Adding flights to the network: The company has identified five countries which are considered strategic destinations for the Western Cape's connectivity. Key international airports in the USA, Canada, Angola, Ethiopia and Kenya will be individually added to the network.

Continuing with Manson's framework, the *evaluation* step aims to check that the model behaves as expected using verification and validation techniques. Verification ensures that the model accurately reflects the network which it aims to model and that the results can be trusted. Validation is similar to verification but goes a step further by ensuring that the model meets CTAA's requirements and needs. This will require a workshop with the decision-makers at CTAA, where the purpose and functionality of the model is communicated. This will not only be a presentation, but also an interactive session where the decision-makers can give their input and test various scenarios themselves. This will allow them to make a much more informed decision as to the usefulness of the model.

Multiple iterations of the design may be needed before the model can be fully verified and validated. The final step in the process, *conclusion*, must answer the research question based on feedback from the company. The reasons for success or failure must be established and suggestions for improvement put forward.

1.5 Document outline

The literature review in Chapter 2 assesses the external factors that affect the airline industry and their influence on the model. An introduction to CNT is given and its applicability to this problem is demonstrated. Recent advances in CNT are discussed, and the context of the problem within this research field is identified. Chapter 3 analyses the three types of data provided by the company and indicates how they will be used in the model. Chapter 4 deals with the formulation of the network, details the metrics used in the model, discusses the heuristics to be used and verifies the accuracy of the model. Chapter 5 gives the results of testing the model under various scenarios and using different heuristics. These results are analysed in detail to identify which routes are the best to expand/develop as well as to identify patterns in the behaviour of the model. Chapter 6 gives the company's response to the model's usefulness and concludes this report.

Chapter 2

Literature review

This chapter is a review of the available literature to assist with the design of a solution. Section 2.1 discusses the macro-environment of the airline industry and the impact of global events on air travel. Section 2.2 provides a simple introduction and explanation of the science of CNT. Sections 2.3 and 2.4 explain the two applicable uses of CNT and why they are relevant to this problem. Recent advances in CNT are discussed in Section 2.5. Section 2.6 provides some context to the network in which CTIA is found.

2.1 External factors affecting the airline industry

The airline industry is closely linked to the global economy. When the economy is doing well, people have more money to spend on travel, and businesses are more dynamic and expansive. Figure 2.1 shows how the opposite is also true, as the 2008 recession resulted in global passenger volume decline. The blue line tracks the manufacturing Purchasing Managers' Index (PMI), which is a widely used index to measure the state of the manufacturing industry. A diffusion index value of more than 50 indicates expansion in the industry and less than 50 indicates contraction [18]. The green line is another popular PMI, used to track global services, which make up the majority of most advanced economies [22]. The red line tracks the growth in global Revenue Passenger Kilometres (RPK), which is the total number of passengers multiplied by the number of kilometres flown.

On average, global RPKs grow by about 5.5% annually [17], however, there are fluctuations which are the result of a number of factors:

Airfares: From February 2016 to February 2017, the cost of air travel dropped by about 10% in real terms [17]. This is largely due to the recent drop in oil price from over \$100 per barrel in June 2014 to a 14 year low of \$29.30 per barrel in January 2016 [20]. The resulting decrease in airfares and a "cyclical pick-up in global economic conditions" caused a 7% growth in global RPKs in February, which is well above the long-run average [17]. Figure 2.2 shows how closely the oil price is linked to the cost of flying.

Demand seasonality: It is clear that the most popular travelling time is during the middle of the year, as shown by the large annual spikes in Figure 2.3. This is likely due to the high demand during the Northern hemisphere's Summer season. The smaller spikes at the end of each year are most likely due to Festive travel and the Southern Hemisphere's Summer season. This is visible when analysing CTIA's direct-flight passenger volumes from Europe, as shown in Figure 2.4. There is only a slight dip during the Winter months for the majority of these routes, but two are very noticeable: Firstly, Heathrow International Airport (black line) drops dramatically from over 30 000 passengers in Summer, to just over 10 000

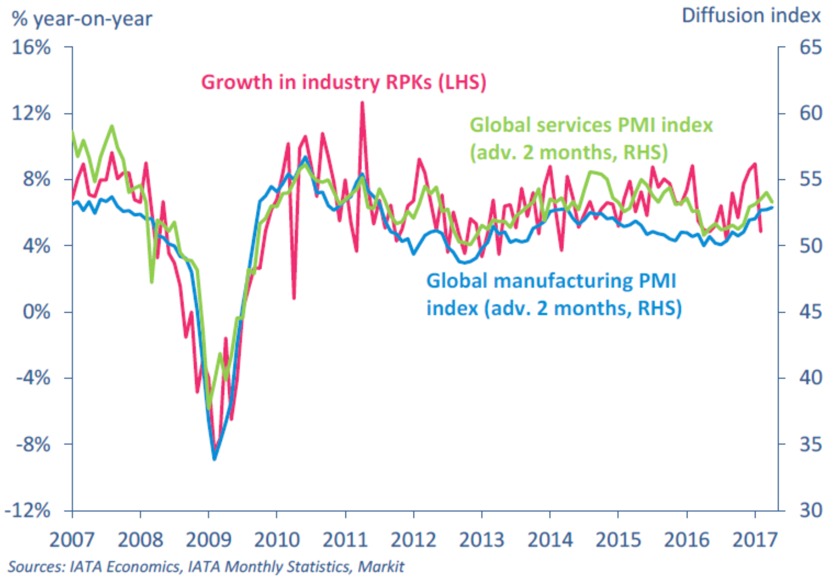


Figure 2.1: The relationship between global business vitality and air passenger volume growth from 2007 to 2017 [17]

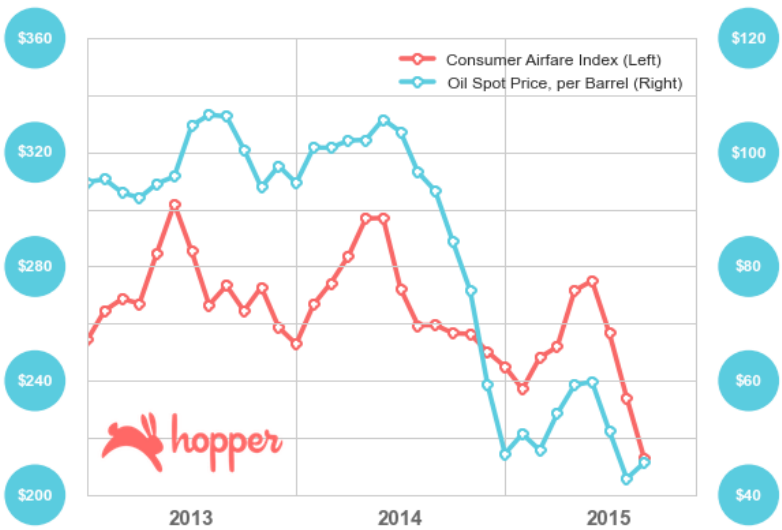


Figure 2.2: The correlation between airfares and oil price [15]

passengers in Winter. Secondly, the Franz Josef Strauss Airport in Munich (pink line) only flew to CTIA from January-March and October-December in 2015. This is an indication of the popularity of Cape Town during its Summer months, and its comparatively low demand during Winter.

Politics: The impact of politics on the aviation industry has recently become very significant. After Britain voted to leave the European Union, airlines scrambled to form contingency plans due to the uncertainty over Brexit negotiations. However, this only affects intra-European flights to and from Britian, therefore South Africa is not impacted. Emirates Airline has redeployed some of its capacity away from the USA after Donald Trump's travel ban immediately reduced booking to the USA by 35% in March [26]. In terms of local

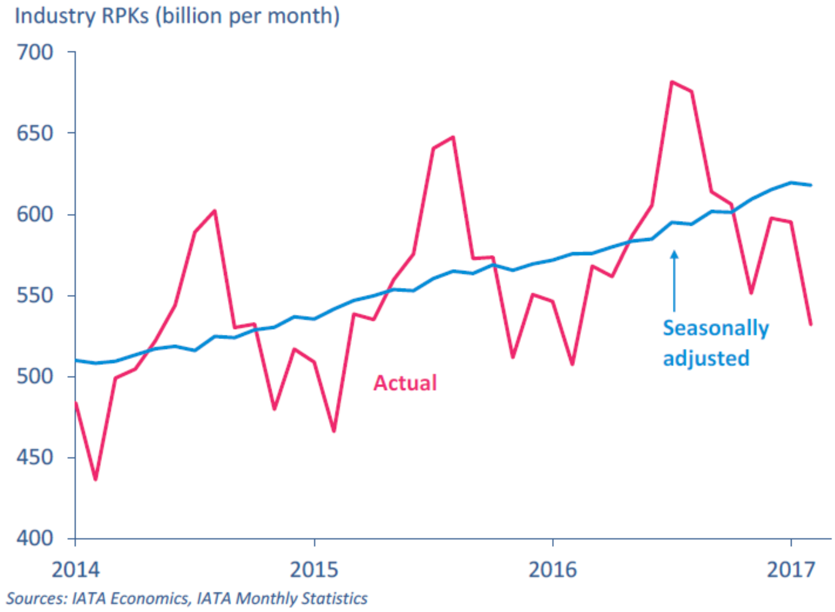


Figure 2.3: The global seasonality in air passenger volumes from 2014 to 2017 [17]

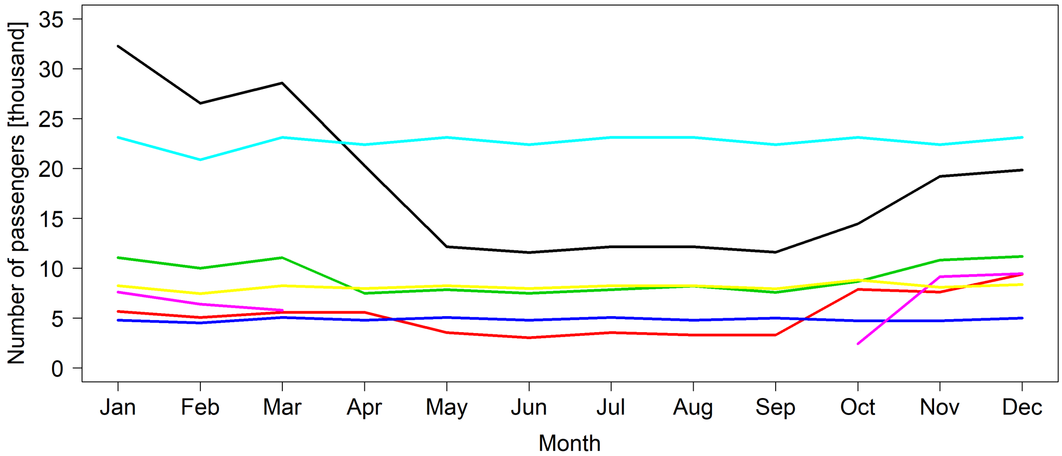


Figure 2.4: The seasonality in air passenger volumes from major European airports to CTIA in 2015.

developments, the Department of Transport (DOT) has implemented the Air Transport Strategy 2015–2020, which aims to grow the South African airline industry and is based upon the National Development Plan — Vision 2030. The country's national flag carrier, South African Airways (SAA) is in a dire financial position, with current losses estimated to be around R4.5 billion [23]. If its Long-Term Turnaround Strategy (LTTS) is successful, then it may begin to use CTIA for international flights, however, it currently only uses ORTIA for these operations [9].

Other factors: Factors such as natural disasters, terrorism and aircraft accidents can have significant impacts on the airline industry. However, since they are scarce and unpredictable, they will not be included in the model, which will aim to represent the network during 'normal' functioning.

2.2 Complex Network Theory

Many real-world systems can be described as *complex networks*. Living systems consist of a huge number of chemical interactions between proteins and genes [29]. Social networks consist of a large number of people who interact with each other in varying degrees [28]. The World Wide Web has a huge collection of web pages with links to other pages [3]. The complexity and size of the interactions in these networks are almost impossible to study without using CNT [3].

In its simplest form, CNT is the study of networks which consist of nodes and links. Nodes are the elements of the network and links are the interactions between these elements. CNT has been used extensively to model air transportation networks, where airports are represented by nodes and the flights between them by links. Literature in this field is divided into two main categories: Firstly, to study the *topology* (Section 2.3) and secondly, to study the *vulnerability* (Section 2.4) of air transportation networks. These categories are not mutually exclusive, as some papers include both.

2.3 The study of network topology

The pioneering work done by Erdös and Rényi [11] is considered the simplest of complex network models [7]. The model starts with N disconnected nodes and builds the network by randomly adding links. The paper describes the "gradual development and step-by-step unravelling of the complex structure of the graph", and discovered certain structural properties of a random graph. This implies that networks can be characterised by certain topological features, and therefore compared to other networks. Clear evidence of the topological analysis of networks is shown in [16] where the Australian Airport Network (AAN) is compared to that of China [27], India [2], Italy [13], the USA [30], and the world [14].

Apart from the random graph of Erdös and Rényi [11], two other topologies are common in literature —*small-world* and *scale-free*. A network has a *small-world* property if most of its nodes can be reached by another node using a relatively short path [7]. A common example of this property is found in social networks. Travers and Milgram [24] performed the now famous experiment in which it was discovered that two US citizens randomly chosen were linked by an average of six acquaintances. All the airport networks of Australia [16], China [27], India [2], Italy [13], the USA [30] and the world [14] were found to exhibit a *small-world* topology. *Scale-free*: Barabási and Albert [3] showed that many networks exhibit a power-law degree distribution. The degree of a node is the number of connections it has to other nodes. For example, if an airport has flights to 20 other airports, it is said to have a degree of 20. The degree distribution is a probability function of the number of connections per node, $P(k)$, where k is the degree of a node. The probability of a node having a certain degree k is described by the function: $P(k) \approx C \times k^{-\lambda}$. According to the formula, high degree nodes are less likely to occur than lower degree nodes. Barabási and Albert [3] give two reasons for this, namely growth and preferential attachment (aspects not accounted for in the Erdös and Rényi [11] model). As the network grows, the higher degree nodes will have a higher rate of connectivity than others. Thus, an initial connectivity difference between two nodes will increase as the network grows. This results in a network with a few highly connected nodes (hubs).

Studying network topology in this way provides a static description of a network, which can be useful. However, CNT is also used in dynamic modelling, which is necessary when the focus of this study is on the impact of changes in routes on the network.

2.4 Vulnerability studies

The purpose of a vulnerability study is to determine the effect of removing nodes and/or links from a network. Disruptions to the functioning of airport networks can have huge economic implications [19]. Recent examples of major disruptions include the eruption of the Icelandic Volcano 'Eyjafallajökull' in 2010, which cost an estimated 1.7 billion USD and the air controllers strike in Spain, which cost approximately 134 million USD to the companies affected [19].

Du et al. [10] used targeted node disruption to measure the connectedness of the Chinese Airport Network (CAN) after the strategic removal of important airports. The algorithm first removes all the nodes with a degree of 1, then 2, then 3 and continues until all the nodes have been removed. This network decomposition allowed the researchers to analyse the network's sensitivity after the continuous removal of layers of nodes. The study found that the CAN is not as robust as the WAN when high-degree nodes are disrupted, and has been used to gain a better understanding of the network. Another method is to strategically remove links, rather than nodes. Viljoen and Joubert [25] analysed the effect of targeted link disruption on the connectedness of the global shipping network. This is very similar to modelling airport networks, where ports are represented by nodes and the routes between ports by links. In this case, certain routes passing around the Horn of Africa were forced to change due to the threat of piracy [25], thus link disruption was an applicable methodology to use. An important conclusion drawn from this work is that even after removing a third of the global shipping network's links, no large communities were isolated. The ability to reach conclusions such as this using CNT (albeit in a different industry) is important for CTIA because isolating large markets such as North America and Europe can have a hugely negative impact on the Western Cape. CTIA has an indirect traveller market of about 700 000 passengers, of which almost a half travel via other airports [9]. Thus, using a vulnerability study can help to better understand the most important links determining CTIA's degree of accessibility.

2.5 Recent developments in multi-layer modelling

While the aforementioned approaches have been successfully applied to similar aviation problems, limitations have recently been discovered. The concept of a *multiplex* network has become an important consideration in studying real-world systems. Cardillo et al. [6] discuss the prevalence of layers within a network which define the relationships between nodes and the neighbourhoods in which they belong. The social network is a good example of nodes (people) interacting with other nodes depending on the type of relationship they have — be it friendship, professional etc. The type of information transmitted between friends is usually not the same as that between colleagues, an aspect not accounted for when using a single-layer graph [6]. Thus, a more accurate representation of a real-world system includes its layers, such as the way in which Buldyrev et al. [5] analysed the Internet and the Italian power grid as an integrated, multi-layer system. Interestingly, the interconnections between nodes of different layers significantly increased the vulnerability of the network, as a failure in one layer could propagate to the other layer resulting in a cascading failure of nodes [5]. Cardillo et al. [6] implemented a multi-layer model of the European ATN consisting of 15 layers, where each layer represents one of the 15 biggest airlines in Europe. The airlines were also categorised into two main families, namely *major* airlines (such as Lufthansa and British Airways) and *low-cost* airlines. It was found that when compared to the equivalent single-layer model, both the topological features of the network and the vulnerability differed [6]. In the problem at hand, such a fine-grained view is not necessary because flight cancellations resulting in the rescheduling of passengers are not considered. The network does not take into account individual flights, but rather overall seat capacity.

2.6 Problem context

The context of this problem is fairly unique when compared to the applications of CNT found in literature because the impact on the network is determined by changes in only one node's links. In other words, all the link changes will take place between CTIA and another airport, whereas the majority of literature changes links between many unique airport pairs. However, aspects of both types of studies can be applied to this problem. In terms of topology, the scope of the network will be all the airports reachable within 2 flights from CTIA. This results in a network with 825 nodes. The degree distribution of these 825 nodes is depicted in Figure 2.5. It appears to follow a power-law, with about 10 highly connected nodes in the network giving it this property. Table 2.1 shows the top twenty airports in this network ranked by degree. It is clear that Europe is a key region for connectivity, as 6 of the top 7 ranked airports in this network are European destinations. CTIA is only ranked 14$^{\text{th}}$ on this list, which is striking if one considers that the network is based on CTIA's first and second order neighbourhoods. It is not even well-connected in a network constrained on its own neighbourhood. This highlights the need for improved connectivity.

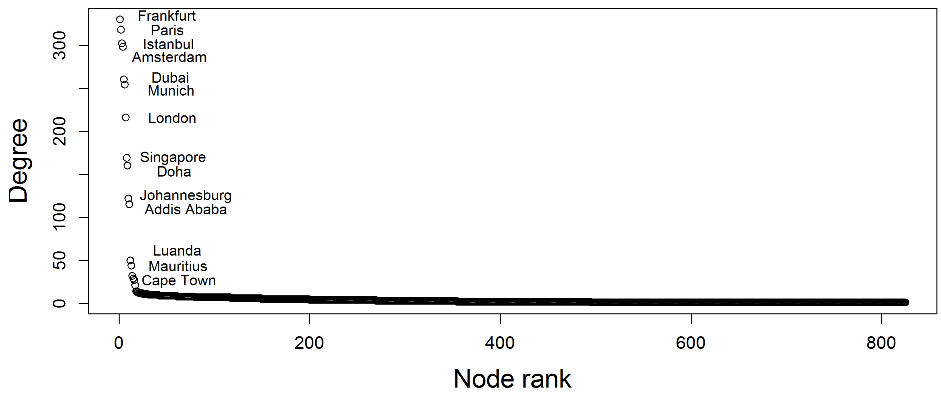


Figure 2.5: Degree distribution of annual capacity within 2 flights from CTIA in 2015, where nodes are in descending order by degree

2.7 Summary of literature review

Airport networks are dynamic, unpredictable and affected by many external influences. CNT has been proven to accurately model these networks and thus attempt to gain more insight into them. The two main applications of CNT in airport network literature are *network topology* and *vulnerability*. Most of this literature uses single-layer modelling to analyse networks, although multi-layer modelling is becoming increasingly popular. For the purpose of this report, only a single-layer model will be used. From an initial look at the network, CTIA is not a well-connected airport, an issue which this report aims to rectify. The next chapter examines the data more closely to give context into CTIA's position within its airport network.

Table 2.1: The top twenty airports in the network of two flights from CTIA ranked by degree

Rank	**Airport**	**Degree (k)**
1	Frankfurt International Airport	330
2	Paris - Charles De Gaulle Airport	318
3	Istanbul - Ataturk Airport	302
4	Amsterdam - Schiphol Airport	298
5	Dubai Airport	260
6	Munich - Franz Josef Strauss Airport	254
7	London - Heathrow Airport	216
8	Singapore - Changi Airport	169
9	Doha Airport	160
10	Johannesburg - O.R. Tambo International Airport	122
11	Addis Ababa - Bole Airport	115
12	Luanda - Quatro de Fevereiro Airport	50
13	Mauritius International Airport	44
14	Cape Town International Airport	32
15	Durban - King Shaka International Airport	29
16	Maputo International Airport	27
17	Windhoek Hosea Kutako International Airport	21
18	Nelspruit - Kruger Mpumalanga International Airport	14
19	Nairobi - Jomo Kenyatta International Airport	13
20	Beijing - Capital Airport	13

Chapter 3

Data analysis

The purpose of this chapter is to analyse the data provided by CTAA. This is important because it will allow for a better understanding of the network and provides the opportunity to anticipate the behaviour of the model. The three datasets are: *Capacity* — The direct seat capacity between every airport pair in the world on a *monthly* basis in 2015. *Routing* — The different routes used by passengers flying to CTIA from 2010 to 2016 on an *annual* basis. *Business* — The number of passengers flying business class versus economy class from 2010 to 2016 on an *annual* basis (this includes both to and from CTIA). These data sources will from now on be referred to as *CapacityData*, *RoutingData* and *BusinessData*.

Section 3.1 deals with the issue of seasonality in the data and how it is resolved. Section 3.2 compares the direct flight networks of CTIA and ORTIA. Section 3.3 examines the top destinations in terms of business-class passengers. Section 3.4 examines the most popular routes for passengers travelling from key destinations to CTIA.

3.1 Data aggregation

Since the model will combine *CapacityData*, *RoutingData* and *BusinessData*, the time intervals must be consistent. Figure 2.4 shows that passenger numbers differ depending on the time of the year. Therefore, to account for this seasonality, a semi-annual time interval will be used – From April to September (the 'Winter' season) and from October to March (the 'Summer' season). This will require that *CapacityData* be aggregated by summing the monthly seat capacity into semi-annual intervals. *RoutingData* and *BusinessData* need to be disaggregated by using *CapacityData* to determine the percentage of passengers flying in each season, and thereafter multiplying the annual values by these percentages. To illustrate this, an example is shown in Figure 3.1 with records from each database for the flight from CTIA to Heathrow. The first two tables show the monthly seat capacity in 2015. The third table calculates the sum of both periods (Summer and Winter) and the percentage of the total. These percentages are then applied to the *RoutingData* in the fourth table by multiplying the total number of passengers on that route by the percentage of each season. The same method is applied in the fifth table for *BusinessData*. This is done for both business class and economy class.

3.2 Cape Town vs Johannesburg

Analysis of *CapacityData* gives an indication of how CTIA plays a second-tier role to ORTIA. CTIA has an incoming degree of 31 and an outgoing degree of 32. In contrast, ORTIA has incoming and outgoing degrees of 91 and 93, respectively.

CapacityData

			Summer				
FROM APT CODE	TO APT CODE	January	February	March	October	November	December
CPT	LHR	32 268	26 544	28 938	14 132	19 200	19 840

			Winter				
FROM APT CODE	TO APT CODE	April	May	June	July	August	September
CPT	LHR	20 268	12 410	11 592	12 161	12 161	11 624

FROM APT CODE	TO APT CODE	Summer total	Summer (%)	Winter total	Winter (%)
CPT	LHR	140 922	63,73	80 216	36,27

RoutingData

FROM APT CODE	TO APT CODE	Total	Summer	Winter
CPT	LHR	66 785	42 560	24 225

BusinessData

FROM APT CODE	TO APT CODE	Business Class	Cabin Premium	Business Summer	Business Winter	Cabin Summer	Cabin Winter
CPT	LHR	18 376	197 156	11 710	6 666	125 639	71 517

Figure 3.1: An example of the data manipulation required to get all three datasets to semi-annular time intervals. These are records for the flight from CTIA to Heathrow.

Figure 3.2 is a Pareto chart of the direct seat capacity to CTIA in 2015. It shows that 80% of the incoming capacity (marked off by the red line) comes from only 5 of the 31 airports (shown by the bars to the left of the blue line). Almost 50% of its incoming capacity comes from ORTIA (the bar furthest on the left). Figure 3.3, which is the same Pareto chart but for ORTIA, is markedly different to Figure 3.2 in that 80% of O.R. Tambo's incoming capacity comes from 29 of the 91 airports. CTIA has only one international airport to the left of the blue line (Dubai International Airport), whereas ORTIA has 23 international airports to the left of the blue line. This is significant because it suggests that CTIA is highly dependent on ORTIA for its international connections and reiterates the need for improved access to CTIA.

3.3 Business vs economy class data

RoutingData can be used to identify which routes are likely to attract the most business-class passengers. Since approximately 80% of an airline's revenue comes from business class passengers [9], a route which has a high number of these is an attractive one because it is likely to make more money. South Africa's airport network is a hub-and-spoke design, with ORTIA serving as the primary gateway to rest of the country given its mostly business-driven demand [9]. Despite major economic growth in the Western Cape, flights to CTIA are still predominantly leisure-driven [9].

CTAA has identified 5 countries that are regarded as 'strategic global destinations', namely Angola, Canada, Ethiopia, Kenya and the USA. Since no specific airports in these 'strategic global destinations' have been provided, the assumption is that the busiest international airports are the most likely choices for route development. Where applicable, the top 10 airports in each country by international seat capacity in 2015 are shown in the Appendix, Table A.1. These 28 airports are then ranked according to the percentage of passengers in business-class seats (Table 3.1). Quatro de Fevereiro Airport in Angola is one of the few airports in this table that has a

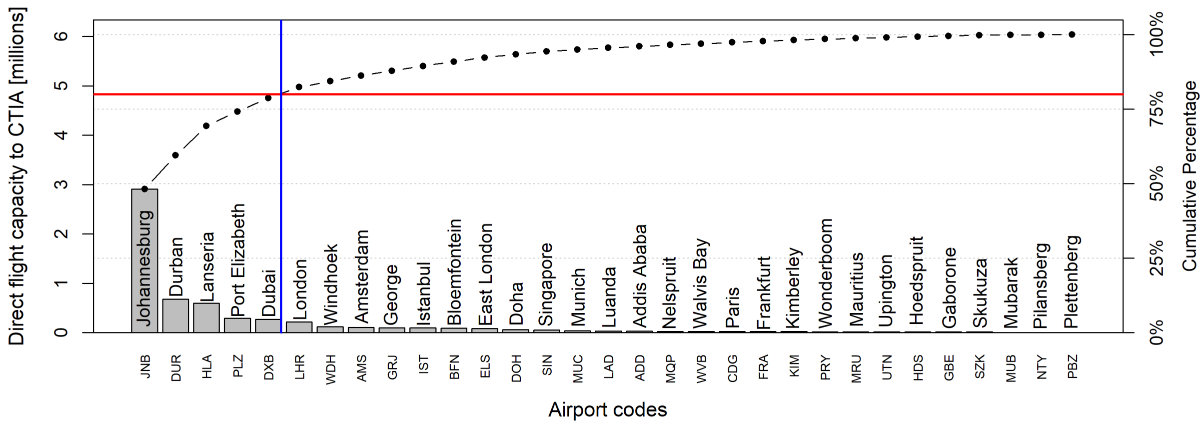


Figure 3.2: The direct seat capacity to CTIA from its 31 connected airports in 2015

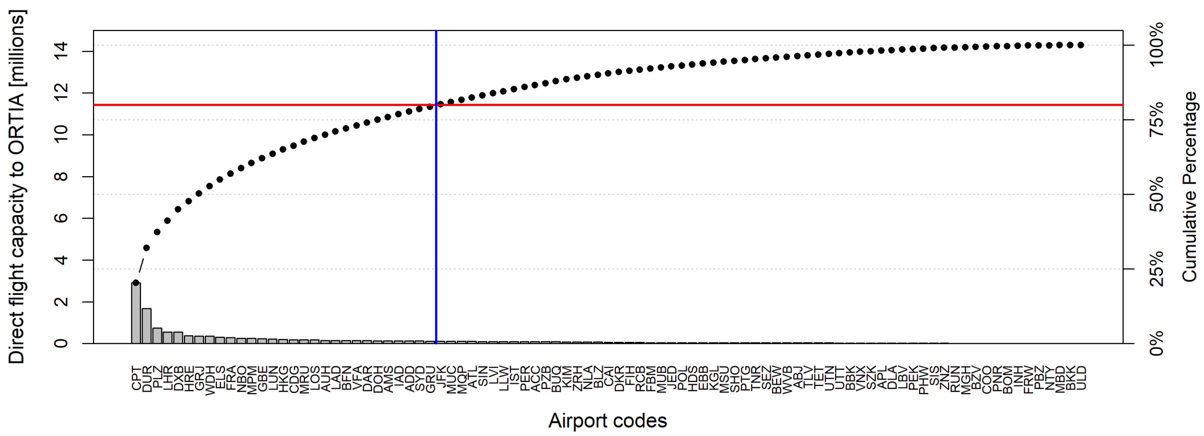


Figure 3.3: The direct seat capacity to ORTIA from its 91 connected airports in 2015

direct flight to CTIA, which is likely the reason for its number one ranking, as passengers prefer the convenience of a direct flight over connecting flights. It is quite clear that the United States dominates the top business destinations, as it has 7 airports in the top 10. This makes sense given its status as the world's largest economy. Interestingly, Canada has 8 out of 10 airports in the bottom half of the table. This is something that the Western Cape would want to improve, because the passenger volumes from Canada are relatively high (see Pearson International in Toronto for example), but the proportion of business class passengers is low. A direct flight would definitely improve this situation, especially if one of the incentives includes a trade agreement between the two countries.

Table 3.1: Ranking of the top business destinations in terms of the proportion of business class passengers to and from CTIA. The values for 'Economy class' and 'Business class' are the total number of passengers in 2016.

Rank	Airport	Country	Economy class	Business class	% in business class
1	Luanda - Quatro de Fevereiro Airport	Angola	54 287	8 450	13.47
2	Toronto Island Airport	Canada	18	2	10
3	New York - John F. Kennedy International Airport	United States	45 228	4 718	9.45
4	Houston - George Bush Intercontinental Airport	United States	5 020	522	9.42
5	Washington - Dulles International Airport	United States	14 842	1 247	7.75
6	Ottawa International Airport	Canada	810	67	7.64
7	Los Angeles International Airport	United States	13 845	1 109	7.42
8	San Francisco International Airport	United States	9 175	689	6.98
9	Hartsfield-Jackson Atlanta International Airport	United States	5 083	359	6.6
10	Miami International Airport	United States	8 014	563	6.56
11	Nairobi - Jomo Kenyatta International Airport	Kenya	38 352	2 637	6.43
12	Addis Ababa - Bole Airport	Ethiopia	20 070	1 324	6.19
13	Dallas/Ft. Worth International Airport	United States	2 868	187	6.12
14	Chicago - O'Hare International Airport	United States	8 339	493	5.58
15	Mombasa - Moi International Airport	Kenya	1 339	75	5.3
16	New York - Newark Liberty International Airport	United States	2 734	146	5.07
17	Toronto - Pearson International Airport	Canada	10 361	509	4.68
18	Lubango Airport	Angola	685	32	4.46
19	Montreal - Pierre Elliott Trudeau International Airport	Canada	4 390	128	2.83
20	Vancouver International Airport	Canada	5 618	156	2.7
21	Calgary International Airport	Canada	2 668	63	2.31
22	Halifax International Airport	Canada	605	13	2.1
23	Winnipeg Airport	Canada	746	12	1.58
24	Edmonton International Airport	Canada	1 493	19	1.26
25	Quebec - Jean Lesage International Airport	Canada	185	1	0.54
26	Dire Dawa - Aba Tenna D Yilma Airport	Ethiopia	50	0	0

3.4 Routing data

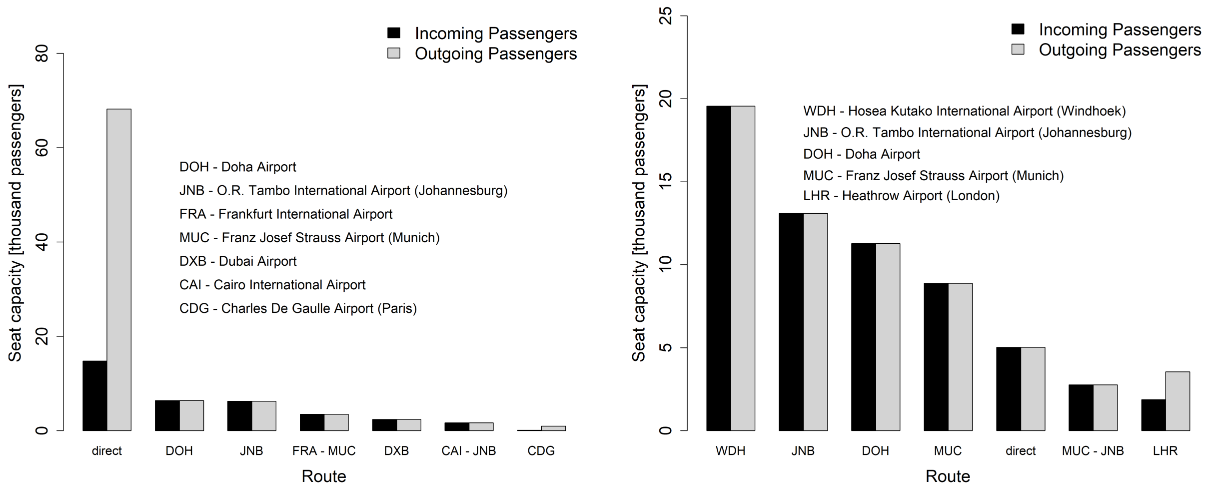
Since CTIA is dependent on ORTIA for its international connections, and the top business destinations have now been identified, the routes passengers use to travel to and from Cape Town must now be analysed. This data is the most important, because the heuristics which will be used in the model are based on *RoutingData*. An example of randomly selected entries from this data is shown in Table 3.2 for flights between Los Angeles International Airport (LAX) and CTIA.

Table 3.2: A random sample of entries showing routes between LAX and CTIA from *RoutingData* over the period 2013-2016.

Route	**No. of passengers**	**Year**
Cape Town – Amsterdam – Los Angeles	796	2013
Los Angeles – Amsterdam – Cape Town	82	2013
Los Angeles – New York (Queens) – Johannesburg – Cape Town	490	2013
Cape Town – Johannesburg – New York (Queens) – Los Angeles	373	2013
Cape Town – Amsterdam – Copenhagen – Los Angeles	291	2014
Los Angeles – Copenhagen – Amsterdam – Cape Town	0	2014
Los Angeles – Paris – Cape Town	1975	2014
Cape Town – Paris – Los Angeles	196	2014
Los Angeles – London – Johannesburg – Cape Town	563	2015
Cape Town – Johannesburg – London – Los Angeles	464	2015
Cape Town – Amsterdam – New York (Queens) – Los Angeles	15	2016
Los Angeles – New York (Queens) – Amsterdam – Cape Town	32	2016

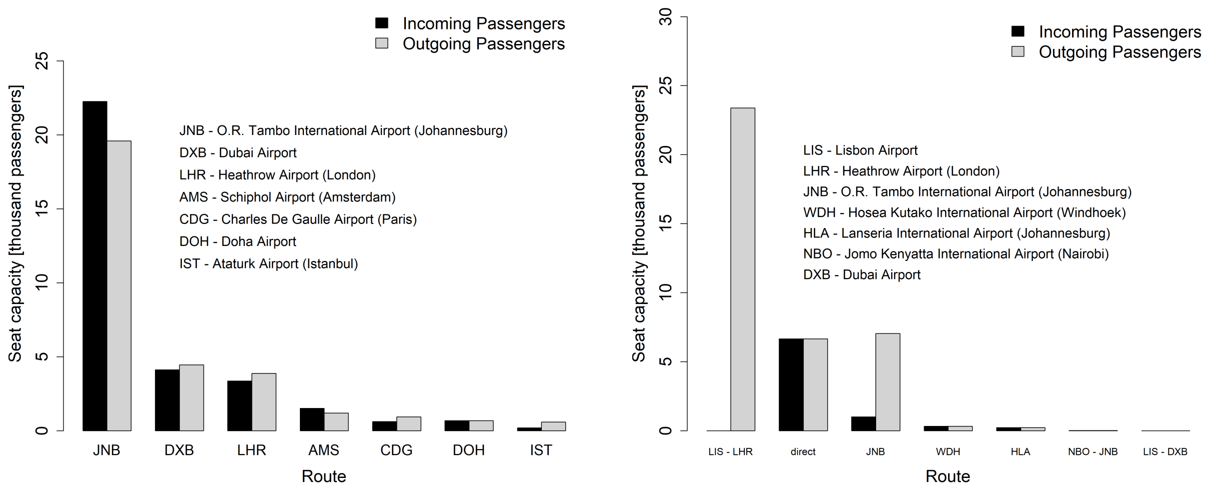
The data in Table 3.2 was chosen at random, with the purpose of showing its format, as well as to demonstrate that the number of passengers on a route from CTIA does not determine how many passengers will return via the same route. There are factors that influence whether passengers switch between routes, such as the flight scheduling and whether there is space on the aircraft. An interesting destination which shows how passenger volumes on opposing directions of a route can vary is Heathrow Airport in London. Figure 3.4a shows how the direct route has significantly more passengers going from CTIA to Heathrow compared to the other direction. This can also be seen in Figure 3.4b, where more passengers go through Heathrow to get to Frankfurt than in the opposite direction. Figure 3.4b also shows that travelling to Frankfurt direct is much less popular than going via Windhoek, which has a large German population. Similarly, Figure 3.5b shows the popularity of going through Lisbon and Heathrow to reach the mostly Portuguese-speaking Angola. Figure 3.5a reiterates the dependence of CTIA on ORTIA for its international connections, in this case the passengers travelling to and from JFK.

Other factors such as the cost of the airfare and the travel time are also important in determining whether a passenger will change his/her route. For example, if the direct route from JFK to CTIA is established, then passengers might be tempted to use this route rather than taking a connecting flight via Heathrow. This is a more convenient option because the total travel time is shorter, however, it is more expensive. There is a trade-off here between cost and travel time. Unfortunately, the data provided by CTAA does not include these two types of data between every airport, however, the solution to this lies in Figure 3.6. Cost and travel-time data has been gathered from numerous airline websites for flights between these regions of the world. This is not a 100% accurate reflection of the flights between specific airports, however, it does give the model a good estimation of the switching behaviour of passengers. This is explained further in the 'Switching behaviour heuristic' part of Section 4.3.



(a) Seat capacity on routes between CTIA and Heathrow. (b) Seat capacity on routes between CTIA and Frankfurt.

Figure 3.4: The seat capacity on routes between CTIA, Heathrow and Frankfurt. The black bars represent the incoming passengers, and the gray bars the outgoing passengers. The code under each bar is the airport through which the route passed.



(a) Seat capacity on routes between CTIA and JFK. (b) Seat capacity on routes between CTIA and Quatro de Fevereiro airport in Angola .

Figure 3.5: The seat capacity on routes between CTIA, JFK and Quatro de Fevereiro airport in Angola. The black bars represent the incoming passengers, and the gray bars the outgoing passengers. The code under each bar is the airport through which the route passed.

3.4.1 Time variability in RoutingData

RoutingData consists of records from 2010-2016. For the purpose of this model, only one entry for each route is needed, not seven years' worth, therefore the data must be altered. The first (and easiest) method is to calculate the average of the number of passengers taking certain routes

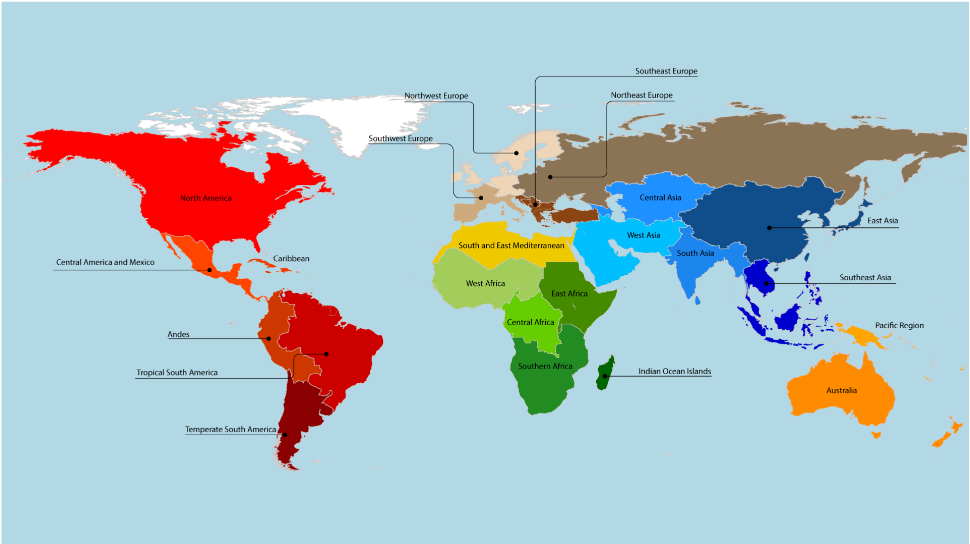


Figure 3.6: Map showing the world regions for which cost and travel-time data has been gathered. [1]

over the 7 years. This is problematic for two reasons: Firstly, some routes may not have existed in 2010 that did in 2016. For example, the Beijing – Addis Ababa – Cape Town route did not exist in 2010, but from 2011–2016, it grew to become the second most used route (15.7%) behind Beijing – Dubai – Cape Town (49.3%). Another good example is Cape Town to Nairobi (Jomo Kenyatta International Airport), which was only developed in 2016. Secondly, there are trends which cannot be ignored. For example, Dubai International Airport has grown considerably in recent times and is now one of the main hubs in the WAN. Figure 3.7 shows three of the most popular routes to get to CTIA from JFK apart from the Johannesburg route. It shows how London Heathrow (black line) was overtaken by Dubai (red line) in 2016 after several years of growth. This is a reflection of the dynamic nature of the airline industry, which the model must account for. Therefore, instead of averaging the passenger numbers for the 7 years of data, each route will be defined by a distribution which describes the probability of there being x number of passengers on that route. This also means that the model will not produce the same results every time, as passenger volumes on all routes will differ with each simulation run. This is a more accurate reflection of the real world scenario.

3.4.2 Utilisation

The fullness of the aircraft on a particular route is crucial to an airline because the fuller the aircraft, the greater the revenue and therefore the higher the margins. Figure 3.8a shows the top 5 best utilised routes direct to and from CTIA. These flights are generally 80–90% full, compared to the 60–70% utilisation experienced by the routes shown in Figure 3.8b. Despite this low utilisation, these routes are sustainable because they are domestic routes with high passenger volumes per period. In contrast, international flights have much lower passenger volumes per period. They also pose a greater risk to airlines because an empty seat on an international flight is much more costly than an empty seat on a domestic flight. Therefore, airlines need to be convinced that connecting CTIA to a proposed international destination will result in utilisation figures similar to those seen in Figure 3.8a.

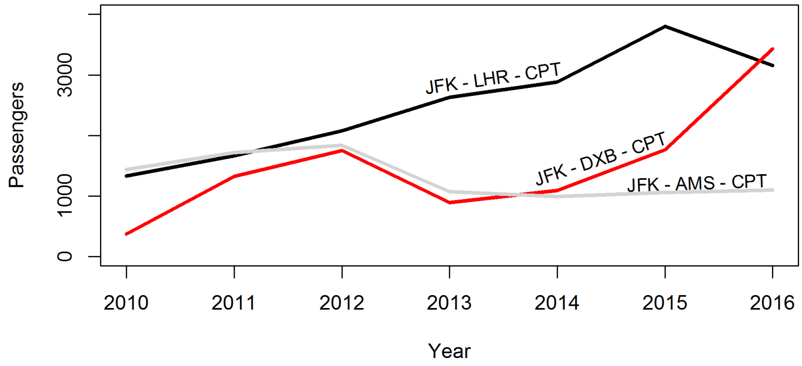
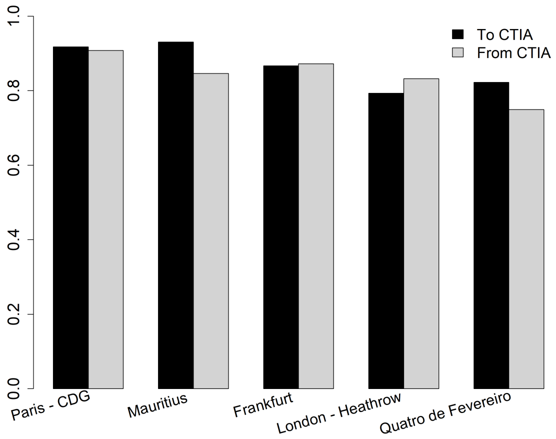
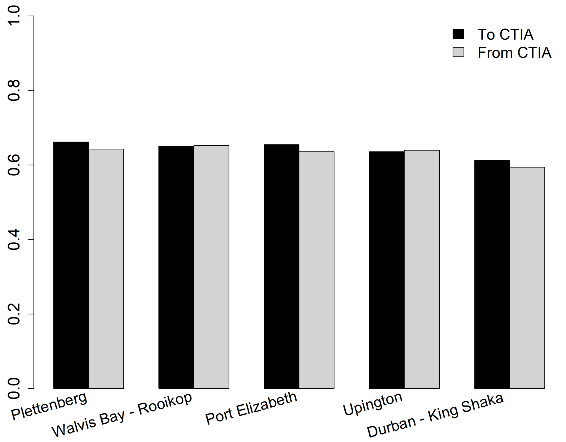


Figure 3.7: The number of passengers taking three different routes from JFK to CTIA from 2010-2016



(a) The top 5 most utilised routes direct to and from CTIA.



(b) The least utilised routes direct to and from CTIA

Figure 3.8: The top and bottom 5 ranked routes in terms of utilisation to and from CTIA. The black bars represent the routes to CTIA, and the gray bars the routes from CTIA. The y-axis is the utilisation of the route between 0 and 1, where utilisation = passenger numbers/seat capacity.

3.5 The link between data analysis and conceptual design

The next chapter on conceptual design will use this data analysis to formulate the network and build the model. In particular, Section 3.2 showed how important it is for CTIA to develop direct routes overseas. Section 3.3 added another dimension to development decisions based on how many business-class passengers can be expected on a route. Finally, the heuristics outlined in Section 1.3 will be applied to *RoutingData* to model the change in connectedness and accessibility of CTIA under different scenarios.

Chapter 4

Conceptual design

This chapter presents the mathematical formulation of the network of all the airports in the world within 2 links from CTIA (Section 4.1). This network is therefore limited to 825 airports, compared to the 4 648 airports in the WAN. However, this is sufficient for the model's purposes because the majority of CTIA's traffic comes from airports reachable within two flights. Section 4.2 then describes the metrics that will be used to analyse the network, and the reasons for choosing them. The two heuristics introduced in Chapter 1 are explained in more detail in Section 4.3. The model is verified in Section 4.4 to ensure its accuracy and appropriateness. Finally, the methodology behind the scenario testing is explained in 4.5.

4.1 Network formulation

In the formulation of the network, airports are represented by nodes. A link between two nodes is established when a flight exists between them. This is required to make the explicit connection between the real-world elements. Let the network be defined by the graph $G = (N, L)$, where N is the set of nodes and L is the set of links that connects the nodes. The node set is defined by,

$$N = \{N_1, N_2, N_3, \ldots, N_n\} \tag{4.1}$$

The links between nodes are defined by,

$$a_{ij} = \begin{cases} 1, & \text{if } N_i \text{ is connected to } N_j \\ 0, & \text{otherwise} \end{cases} \qquad \forall i, j \in \{1, 2, ..., n\} \text{ and } i \neq j \tag{4.2}$$

A weighted network is one in which a measure of importance is placed on links between nodes, depending on the relationship between the nodes. In this case, the weight is equal to the number of passengers p_{ij} travelling on flights from node i to node j in a semi-annual period, defined by,

$$w_{ij} = \begin{cases} p_{ij}, & \text{if } N_i \text{ is connected to } N_j \\ 0, & \text{otherwise} \end{cases} \qquad \forall i, j \in \{1, 2, ..., n\} \text{ and } i \neq j \tag{4.3}$$

Therefore, the non-zero list of links is defined by,

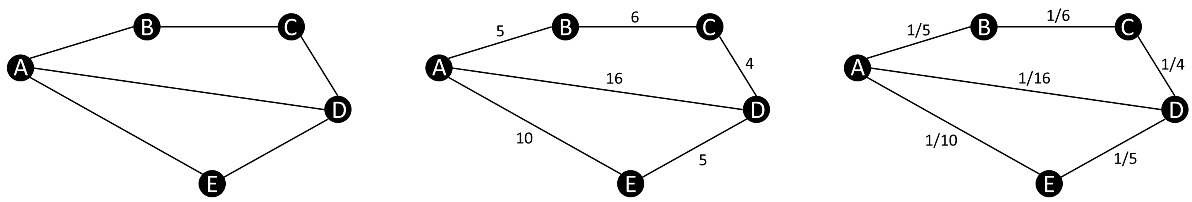
$$L = \{w_{ij}\} \qquad \forall i, j \in \{1, 2, ..., n\},\ i \neq j \text{ and } w_{ij} \neq 0 \tag{4.4}$$

The graph $G = (N, L)$ is now fully defined by the sets N and L. Four network metrics are presented which will be used to analyse the effect of route changes on the network.

4.2 Network metrics

The aim of this section is to introduce and explain the metrics to be used in the model. These metrics are crucial because they will measure the increase/decrease in CTIA's accessibility under different scenarios. Accessibility is a broad term to describe how well connected an airport is. The Western Cape Air Access Strategic Framework [9] expands this definition by adding considerations such as the "cost-efficiency, time-efficiency and flexibility of travel". These considerations are converted into quantifiable metrics by using the principles of CNT.

Average shortest-path length: The length of a path is defined as the number of links between a node-pair i and j. A *geodesic path* (or shortest path) is the shortest of the paths connecting i and j [7]. In Figure 4.1a, there are a number of paths to get from point A to point D. Line A→D is the geodesic path because it only takes one link to move between the two points, whereas the other routes require more than one. This network is described as 'unweighted' because there is no discrepancy between the links. To make the shortest path calculation more meaningful, Figure 4.1b has weights attached to the links. Let these weights define the distance between the points (not to scale). Now, the shortest path is no longer line A→D, but either A→B→C→D or A→E→D. Using the same values for link weights, let Figure 4.1c represent an airport network were the nodes are airports and the links represent seat capacity on the flights between the airports. A 'shortest path' is now one which has the greatest seat capacity, therefore the reciprocal of the weights is used and line A→D is the shortest path.



(a) Unweighted network.

(b) Weighted network with distances between nodes

(c) Weighted network with reciprocal of seat capacity

Figure 4.1: Three simple networks which illustrate the use of link weights in the calculation of the average shortest-path length

The average shortest-path for node i is defined by:

$$l_i = \frac{1}{n(n-1)} \sum_{j}^{n} d_{ij} \text{ , where } d_{ij} \text{ is the shortest path connecting } i \text{ and } j \tag{4.5}$$

The problem with this equation is that if there are unconnected nodes in the network, it diverges. To avoid this, only connected pairs are used. However, Costa et al. [7] states that doing this creates a distortion in a network with many unconnected pairs. It will show a low average shortest-path distance, which is only expected for highly connected networks. Therefore, Costa et al. [7] suggests using the *local efficiency* instead. Equation (4.5) then becomes:

$$e_i = \frac{1}{n(n-1)} \sum_{j}^{n} \frac{1}{d_{ij}} \tag{4.6}$$

Since the model uses seat capacity as opposed to distance for link weights, the geodesic path length (d_{ij}) is not used. Instead, the reciprocal of the seat capacity weighting ($\frac{1}{w_{ij}}$) is used so that the equation becomes:

$$e_i = \frac{1}{n(n-1)} \sum_{j}^{n} w_{ij} \tag{4.7}$$

This equation also causes a problem because if two airports are not directly connected, their link weight will be zero. However, they may still be indirectly connected, therefore a new variable s_{ij} defines the summative weight between two nodes. The equation is then defined by:

$$e_i = \frac{1}{n(n-1)} \sum_{j}^{n} s_{ij} \tag{4.8}$$

Using equation 4.8, the efficiency of CTIA in transporting passengers between itself and other airports can be measured.

Vulnerability: It is important to know which nodes are the most critical to the performance of a network. High degree nodes (hubs) are generally assumed to be the most critical, however this is not always the case. Costa et al. [7] use an example of a binary tree where all the nodes have the same degree. There are no hubs, however, the nodes closer to the roots are more important to the network's performance than the nodes closer to the leaves. In the case of Figure 4.2, node 1 has the highest vulnerability, as its disruption would stop the functioning of the entire network. The subsequent layers after node 1 are less and less critical to the performance of the network because they are lower in hierarchy. The vulnerability of a node is defined by [7],

$$V_i = \frac{E - E_i}{E} \tag{4.9}$$

where E is the global efficiency of the unchanged network, and E_i is the global efficiency after removing node i and its corresponding links. For the purpose of this study, the vulnerability of other airports with respect to CTIA is more important than their effect on the entire network. In other words, the aim is to determine the level of damage that removing a node will do to CTIA and its accessibility, rather than the entire network's accessibility. Therefore, equation 4.9 is modified as such:

$$V_i = \frac{e_C - e_i}{e_C} \tag{4.10}$$

where e_C is the local efficiency of CTIA in the unchanged network, and e_i is the local efficiency of CTIA when node i is removed from the network.

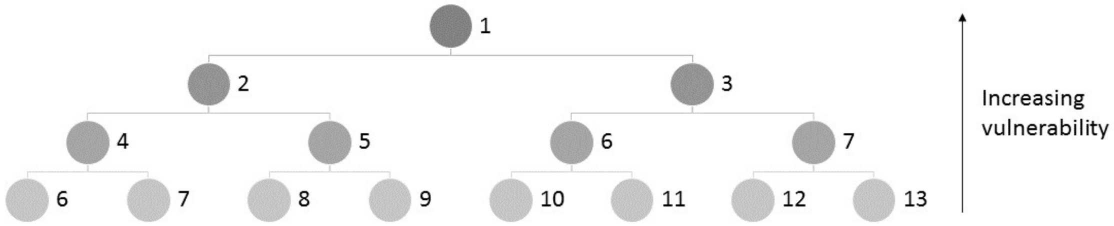
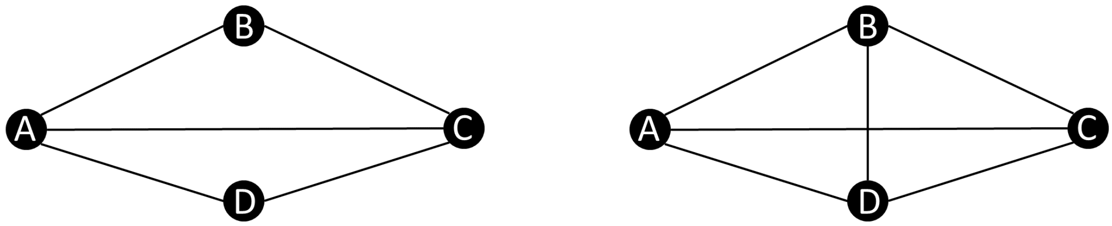


Figure 4.2: Binary tree showing the hierarchical nature of different node layers

Clustering coefficient: This measures how cohesive a node's immediate neighbourhood is by calculating the ratio of the number of *triangles* to the number of *triples*. A triangle is formed when two neighbours of a particular node are connected. Node A in Figure 4.3a has two triangles, ABC and ADC. A node is part of a triple when it has two neighbours, connected or not. Node A in 4.3a has three triples: The two triangles mentioned above, as well as the triple ABD. Essentially, the clustering coefficient calculates the number of triangles divided by the number of theoretically possible triangles.



(a) Node A has a clustering coefficient of $\frac{2}{3}$.

(b) Node A has a clustering coefficient of 1.

Figure 4.3: Two simple networks which illustrate the difference in cohesiveness of a node's immediate neighboured

The formula for the clustering coefficient of node i is stated as follows [7]:

$$C_i = \frac{N_\Delta(i)}{N_3(i)} \tag{4.11}$$

where,

$$N_\Delta(i) = \sum_{k>j}^{n} a_{ij}a_{ik}a_{jk} \tag{4.12}$$

$$N_3(i) = \sum_{k>j}^{n} a_{ij}a_{ik} = \frac{k_i(ki-1)}{2} \tag{4.13}$$

Equation (4.12) is the number of triangles that node i is involved in, and equation (4.13) is the number of occasions where node i is the central node of a connected triple. Rewriting equation (4.11) using equations (4.12) and (4.13) gives,

$$C_i = \frac{1}{k_i(k_i - 1)} \sum_{k>j}^{n} 2(a_{ij}a_{ik}a_{jk}) \tag{4.14}$$

Clustering coefficients can account for weighted links using the equation introduced by Barthélemy et al. [4], where the *weighted* clustering coefficient is given by,

$$C_i^w = \frac{1}{s_i(k_i - 1)} \sum_{k>j}^{n} \frac{w_{ij} + w_{ik}}{2} a_{ij}a_{ik}a_{jk} \tag{4.15}$$

This weighting adds more meaning to the metric because a neighbourhood with low seat capacity between its neighbours is not as cohesive as a neighbourhood with a higher seat capacity.

Betweenness centrality: One way of defining the importance of a node in a network is to measure how many shortest-paths it participates in [7]. Betweenness centrality measures "the extent to which a particular node lies between other nodes in a network" [16], and is defined by,

$$B_k = \sum_{ij} \frac{\sigma(i,k,j)}{\sigma(i,j)} \tag{4.16}$$

where $\sigma(i,k,j)$ is the number of shortest paths between a node-pair (i,j) that pass through a node k, and $\sigma(i,j)$ is the number of shortest paths between (i,j). The Western Cape is handicapped in this regard, because its geographic location on the Southern-most tip of Africa makes it unlikely to have a high 'betweenness' with respect to the WAN. However, adding direct flights from North America for example will make CTIA a gateway to the rest of Southern Africa, thus increasing its betweenness.

4.3 Implementation of heuristics

Proportionality heuristic: Figure 4.4 is a hypothetical case showing the proportion of passengers taking different connecting flights to reach CTIA. The circles represent connecting airports between JFK and CTIA, and the coloured lines represent flights. If the direct route is developed (shown by the dashed line), then capacity on other connecting flights will decrease for two reasons: Firstly, people will naturally prefer the more convenient direct route, and thus airlines will adjust their flight capacity to account for the lower demand on the connecting flights. Secondly, airlines have a finite number of aircraft, and thus in developing a new route, some of these aircraft will be redeployed, which will lower capacity on the connecting flights. What this image shows is that the stolen seat capacity on current connecting flights is a function of the number of seats on the proposed direct flight. For example, let the direct flight be developed to a seat capacity of $x = 10\,000$ seats per year. The stolen seats from the ORTIA route is described by the equation,

$$f(x) = \frac{45x}{100} \tag{4.17}$$

substituting in 10 000 gives,

$$f(10\,000) = \frac{45(10\,000)}{100} = 4\,500 \tag{4.18}$$

Thus, the route via ORTIA will have 4 500 fewer seats available as a result of the direct flight. This equation is applied to all the other routes so that Dubai and Heathrow lose 1 500 seats, Amsterdam loses 1 000 seats, Doha loses 800 seats and the rest lose 700 seats.

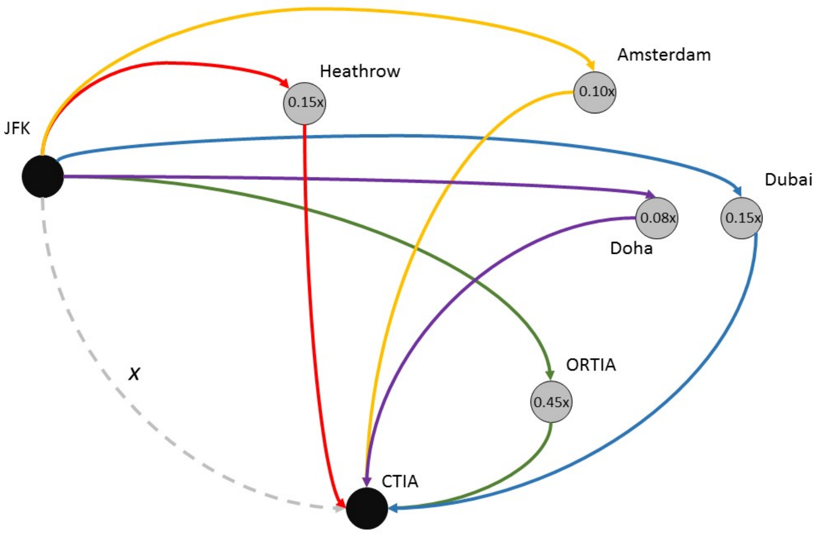
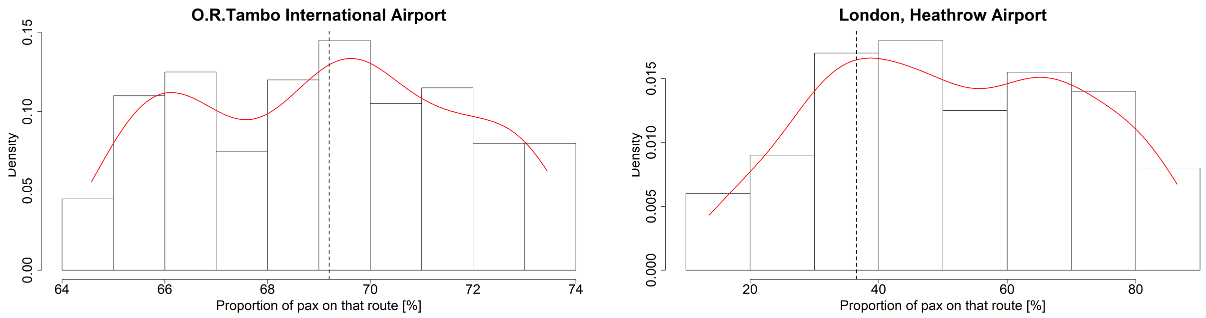


Figure 4.4: An example of how other routes are affected by adding a direct flight to CTIA from JFK. The coloured lines represent actual connecting flights. The dashed line represents the possible development of a direct route. The seat capacity of the proposed direct route is x, and the number of stolen seats is shown in each of the circles as a function of x

Since *BusinessData* contains passenger routing numbers from 2010–2016, the proportion of passengers on each route changes every year. Instead of simply using the average over the seven years, a uniform function will be used to simulate the passenger numbers on each route. Figure 4.5 shows the histograms and resulting uniform density function of the top two connecting flights from JFK to CTIA.

Switching behaviour heuristic: This is a similar concept to the proportionality heuristic, however, the proportion of passengers on each route is calculated differently. Firstly, it is assumed that all the business class passengers will transfer to the direct flight, because it is more convenient, and they can afford the increase in price (*RoutingData* provides the detail on the proportion of business-class passengers on each flight). Secondly, in the absence of flight costs and travel-time data between every airport in the world, only data for the most important connecting routes has been gathered from airline websites. This means that the connecting airports that make up about 99% of the traffic are included, and the other 1% of connecting airports are grouped together and kept constant. For example,



(a) (b)

Figure 4.5: The proportion of passengers travelling through the two most popular connecting airports from JFK to CTIA. The red line is an approximation of the uniform density function, and the vertical dotted line is the mean percentage.

in 2015, passengers used 89 different routes to get from Pearson International Airport in Canada to CTIA. However, as illustrated in Figure 4.6, only 4 connecting flights make up 99% of the traffic (Schipol, Heathrow, ORTIA – Abhu Dhabi and Frankfurt). Therefore, data for these four routes is used to rank them according to cost and travel-time and the number of stolen seats from these routes will be calculated based on these rankings. This ranking is done in the following way: Each flight's cost and time values are divided by the direct flight's estimated values. This gives a ratio of how much better/worse the connecting flight is in relation to the direct flight. The two ratios for each flight are then summed, and expressed as a proportion of all four flights' ratios. This proportion determines what percentage of passengers are stolen from a particular route. This process is detailed in Table 4.1, where the last column shows the percentage of the direct route's seats that will be stolen from each connecting route. For example, if the direct route is developed to 10 000 seats per year, then 18% of 10 000 (1 800 seats) will be stolen from the Amsterdam connecting route, 24% (2 400 seats) from London Heathrow, 34% (3 400 seats) from Abhu Dhabi and 24% (2 400 seats) from Frankfurt.

Table 4.1: The method used to calculate the percentage of stolen seats from each connecting route.

Connecting flight	**Cost ratio**	**Time ratio**	**Sum of ratios**	**Proportion of total ratio sum**
Amsterdam Schiphol	0.53	1.46	1.99	0.18
London Heathrow	0.57	2	2.57	0.24
Abhu Dabhi – ORTIA	0.4	3.27	3.67	0.34
Frankfurt	0.57	2	2.57	0.24
Total	**2.07**	**8.73**	**10.8**	**1**

4.4 Verification

Verification is the process of ensuring that the model accurately reflects the network which it aims to model and that the results can be trusted. Four methods are detailed below which make sure that the model has been built correctly.

Verification of input data: The input data was provided by CTAA who bought it from a company specialising in aviation data. It is therefore assumed that the data is valid

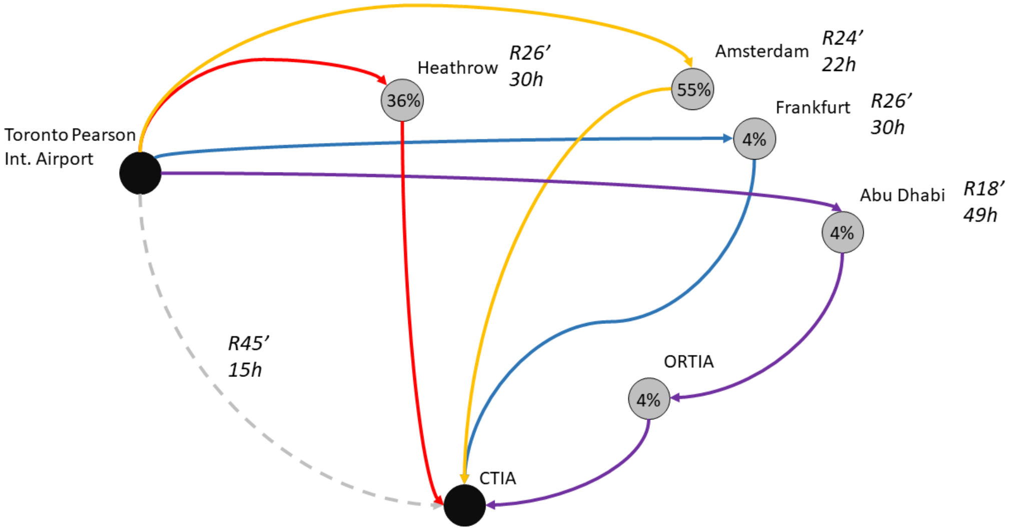


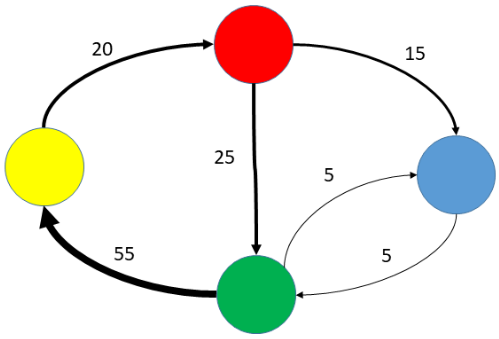
Figure 4.6: An example of how the switching behaviour heuristic uses the most popular routes' time and cost data to simulate the stolen seats on each route. The coloured lines represent actual connecting flights. The dashed line represents the development of a direct route from Toronto Pearson International Airport to CTIA. The numbers in the circles are the percentage of passengers who use that route.

because it comes from a reputable company who are also data providers for the likes of British Airways and RyanAir.

Sense-checking: This involves looking at the results of the data analysis at face value and deciding whether or not they make sense. For example, the incoming and outgoing degrees of Istanbul Atatürk Airport (206 and 201 respectively) are both at the top of the degree lists for this network. A quick internet search shows that this airport is consistently ranked in the top 5 busiest airports in Europe. Therefore, it makes sense that Istanbul Atatürk Airport is ranked so highly in this dataset. It must be noted, however, that this is only a gateway to further analysis. If an anomaly is spotted, (e.g. Dubai airport having a low degree), then this must be investigated further to reveal the cause.

Model definition: The model only includes airports reachable within two flights from CTIA. While this does not cover every possible passenger flying to the airport, it covers all the passengers required for the scenarios being tested. This is because all of the international airports in the five countries being studied can reach CTIA within two flights. The North-American airports generally use European countries for connecting flights, and the African airports generally use ORTIA.

Black-box analysis: Since it is impossible to manually calculate these metrics for a network as large as this one, a few smaller networks with four or five nodes were created (see Figure 4.7 as an example) so that each of the metrics described in Section 4.2 could be easily calculated. The results of these simple calculations were checked against the results produced by the `igraph` package in R to ensure that the program is functioning properly.



Incoming	Circle	Degree	Strength	Efficiency
	Yellow	1	55	0,085556
	Red	1	20	0,054949
	Blue	2	20	0,228788
	Green	2	30	0,073737
	Sum	6	125	0,44303
	Average	1,5	31,25	0,110758
Outgoing	**Circle**	**Degree**	**Strength**	**Efficiency**
	Yellow	1	20	0,098182
	Red	2	40	0,128788
	Blue	1	5	0,106061
	Green	2	60	0,11
	Sum	6	125	0,4430
	Average	1,5	31,25	0,110758

Figure 4.7: Simplified network used in Black-Box analysis. The metrics on the right were calculated using excel and compared to the output of the `igraph` package in R for the same network.

4.5 Simulation methodology

The model will be tested using three different scenarios as introduced in Section 1.4. Scenario 1 will increase the capacity on key flights, scenario 2 will decrease the capacity on key flights, and scenario 3 will introduce new direct flights to CTIA. The airports to be tested in each scenario were chosen using the following method:

Scenarios 1 & 2: The 'key flights' will be determined by finding the airports that are the most important to CTIA's connectivity. The vulnerability metric described in equation (4.10) was used and the results are found in Table 4.2. The vulnerability column shows the percentage increase or decrease of the efficiency of CTIA when that particular airport is removed from the network. The airports are ranked by their absolute value (or magnitude), where a negative value means CTIA's efficiency improves when the airport is removed, and a positive value means CTIA's efficiency declines when the airport is removed. It may seem counter-intuitive that removing hubs such as Dubai and Schiphol results in a negative vulnerability value. This implies that by removing these nodes, CTIA becomes more efficient at connecting passengers to other airports in its network. This implication is correct, however, CTIA's reach is greatly reduced by removing these hubs and therefore while its efficiency may increase, its connectivity decreases. Therefore, Table 4.2 ranks both positive and negative vulnerability values because these are the airports which have the biggest impact on CTIA when removed.

Scenario 3: Table A.1 ranks the busiest international airports in each of the 5 strategic destinations. For Canada and the USA, the top five busiest airports will be tested. For Angola, Quatro de Fevereiro Airport already has a direct flight, and is being tested in scenarios 1 & 2, therefore only Lubango Airport will be tested. For Ethiopia, Bole Airport in Addis Ababa also has a direct flight, therefore only Aba Tenna D Yilma Airport will

Table 4.2: The top 10 most vulnerable airports which have direct flights to CTIA and are not in South Africa.

Rank	Airport	Country	Vulnerability [%]
1	Changi Airport	Singapore	-6.618
2	Dubai Airport	UAE	-4.7198
3	Schiphol Airport	Netherlands	-4.617
4	Addis Ababa - Bole Airport	Ethiopia	4.588
	Hong Kong Int. Airport	Hong Kong	-4.563
5	Paris - Charles De Gaulle Airport	France	-3.882
	Johannesburg - O.R. Tambo Int. Airport	South Africa	-3.448
	Rome - Fiumicino Airport	Italy	-3.167
	Sao Paulo - Guarulhos Int. Airport	Brazil	-3.021
	Madrid - Barajas Airport	Spain	-2.824
6	Luanda - Quatro de Fevereiro Airport	Angola	2.616
	Lisbon Airport	Portugal	2.508
	Beijing - Capital Airport	China	2.339
7	Frankfurt Int. Airport	Germany	2.084
8	Doha Airport	Qatar	-1.599
	Moscow - Domodedovo Airport	Russia	1.547
	Delhi - Indira Gandhi Int. Airport	India	-1.442
9	Munich - Franz Josef Strauss Airport	Germany	1.409
...	...	...	...
10	Windhoek Hosea Kutako Int. Airport	Namibia	1.274

be tested. Finally, the first two Kenyan airports will be used (Jomo Kenyatta and Moi International), as not enough data is available for Wajir and Wilson airports. This comes to a total of 14 airports, all of which are listed in Table 4.3 along with their associated vulnerabilities.

Table 4.3: The 14 airports and their vulnerabilities to be used in the testing of scenario 3

Airport	Country	Vulnerability [%]
Lubango Airport	Angola	0.00619
Toronto - Pearson International Airport	Canada	0.00080
Vancouver International Airport	Canada	0.42768
Montreal - Pierre Elliott Trudeau International Airport	Canada	-0.08373
Calgary International Airport	Canada	0.00229
Edmonton International Airport	Canada	0.00719
Dire Dawa - Aba Tenna D Yilma Airport	Ethiopia	0.15945
Nairobi - Jomo Kenyatta International Airport	Kenya	-0.24139
Mombasa - Moi International Airport	Kenya	0.00373
New York - John F. Kennedy International Airport	United States	-0.75369
Miami International Airport	United States	0.59917
Los Angeles International Airport	United States	0.00048
Chicago - O'Hare International Airport	United States	0.00156
New York - Newark Liberty International Airport	United States	0.00382

Now that the airports have been identified, the specific details of each simulation need to be explained. There are four parameters which make up an individual simulation: The *scenario*, the *heuristic*, the *metric* and the *season*. There are 3 scenarios, 2 heuristics, 3 metrics and 2 seasons, therefore, there are $3 * 2 * 3 * 2 = 36$ possible simulations. Here is an example of one simulation:

Scenario — 3: The airports in Table 4.3 will have a direct flight from and to CTIA introduced into the current network one at a time. The number of direct flights per week will be increased from 0 flights, to 7 flights per week. The aircraft is assumed to be a Boeing 747 with a standard two-class layout. This results in a seat capacity of 524 passengers per flight.

Heuristic — Proportionality Heuristic: The proportionality heuristic 'steals' seat capacity from connecting flights with the same proportion as passengers currently using those connecting flights. The proportion of stolen seats from each connecting flight is calculated by randomly sampling values for each route until the cumulative sum is between 99 % and 101%. The samples are taken from the approximated uniform distributions (see Figure 4.5. As the number of flights per week increases from 0 to 7, more seat capacity will be 'stolen' from each route.

Metric — Clustering coefficient: The clustering coefficient will be measured for each airport at each additional flight added per week. This will result 8 measurements for each airport (from 0 to 7).

Season — Summer: The network is defined by the number of passengers flying between every airport on a six month period from October to March. The flights added will therefore only be defined for a six month period. For example, 2 flights per week are added between CTIA and JFK. This results in $524(\frac{\texttt{pax}}{\texttt{flight}}) * 2(\frac{\texttt{flights}}{\texttt{week}}) * 21(\frac{\texttt{week}}{\texttt{6 months}}) = 22\,008$ passengers between the two airports in that six month period.

Simulation result: The result of this simulation is a line graph with 0 to 7 flights added per week on the x-axis, the weighted clustering coefficient on the y-axis, and a line for each of the 14 airports describing how the clustering coefficient changed as a result of adding each flight (see Figure 4.8).

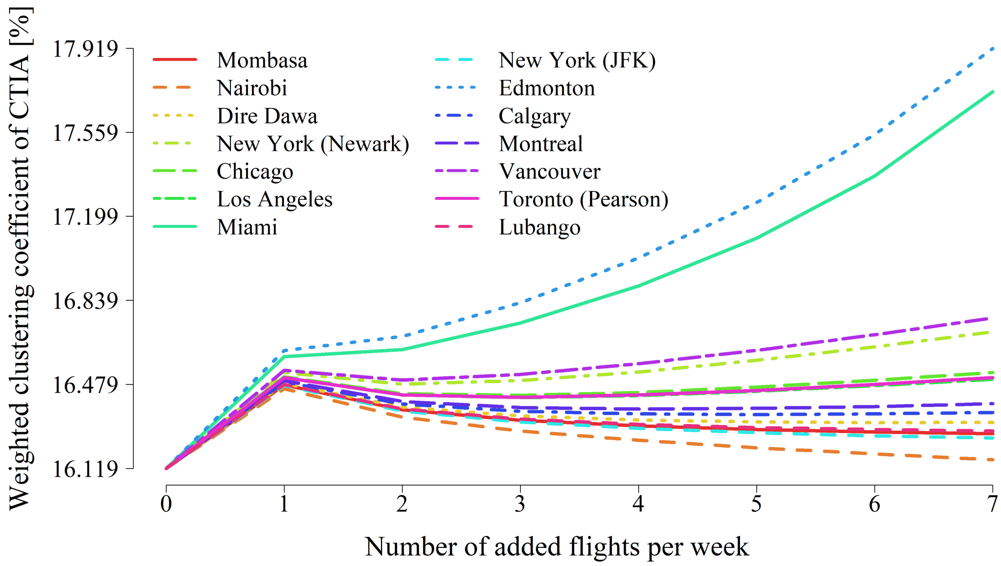


Figure 4.8: The weighted clustering coefficient of CTIA resulting from introducing new direct flights between CTIA and each of the above airports.

4.6 Summary of conceptual design

This chapter explained the formulation of a network model G, which allows the graph to be programmed in R. Four CNT metrics were then introduced and explained (only three are used in testing, because *vulnerability* was used to identify the important airports with regards to CTIA's connectedness). These metrics will measure the connectedness and accessibility of CTIA. The two heuristics to be used in accounting for 'stolen' seats were explained in detail. The model was then verified to ensure its accuracy and applicability to the real-world situation. The airports to be used in the scenario-testing were identified, and the conceptual design of the testing was outlined.

Chapter 5

Results and discussions

This chapter provides the results of the model's simulation runs and discusses the graphs. The purpose of these results is not only to rank the airports according to their efffect on CTIA's connectedness, but also to identify patterns in the graphs and the reasons for these patterns. Section 5.1 examines the effect of increasing the capacity on key direct flights to and from CTIA, while Section 5.2 examines the effect of decreasing the capacity on key flights. Section 5.3 examines the effect of introducing new direct flights to and from key airports in the 5 countries identified as 'strategic global destinations'. The results given here do not show the graphs of all 36 tests. Instead, only the results that are either new or interesting are shown because this will help to answer the question about whether CNT can assist the decision-makers at CTAA.

5.1 Increasing capacity on key flights

The key flights in this section were determined by calculating their effect on CTIA using the vulnerability metric (4.10). In decreasing order of vulnerability, these airports are: Changi Airport - Singapore, Dubai Airport - United Arab Emirates, Schiphol Airport - Amsterdam, Bole Airport - Addis Ababa (Ethiopia), Charles De Gaulle - Paris, Quatro de Fevereiro Airport - Luanda (Angola), Frankfurt Int. Airport - Germany, Doha Airport - Qatar, Franz Josef Strauss Airport - Munich and Hosea Kutako Int. Airport - Windhoek (Namibia). Figure 5.1 shows that with the exception of Doha, Addis Ababa and Windhoek, the efficiency of CTIA will increase as the capacity on key direct flights is increased. It may seem counter-intuitive that these destinations would decrease CTIA's efficiency as the direct flight capacity is increased. However, the connecting routes are swallowed-up by the direct route as more capacity is added, therefore destinations with smaller numbers of passengers bound for Cape Town will have one high capacity direct route and few (if any) connecting routes. The same logic applies for high capacity routes such as Amsterdam and Munich. Due to the large passenger volumes on routes to Cape Town from these cities, the connecting routes are not swallowed up as easily, therefore, while the direct route continues to increase, the overall efficiency of CTIA also increases because the connecting routes generally remain intact.

Figure 5.2 is the Winter equivalent of Figure 5.1. The most noticeable difference between these graphs is the drop in the Munich line. This is indicative of the mostly tourism-driven demand during the Summer months, and the lack of a direct flight during the Winter period. Also interesting is that Paris is now included in the group of destinations that cause a drop in the efficiency of CTIA. This result suggests that if expansion of the Paris to Cape Town route occurs, it must be seasonal to account for the drop in demand during Winter.

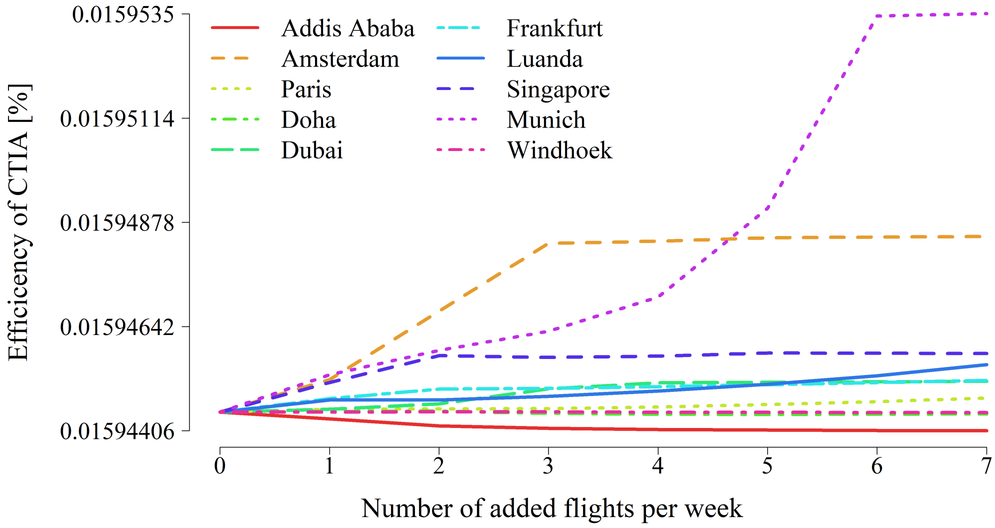


Figure 5.1: The *Summer* efficiency of CTIA resulting from increasing capacity on direct flights between CTIA and each of the above airports. The proportionality heuristic is used.

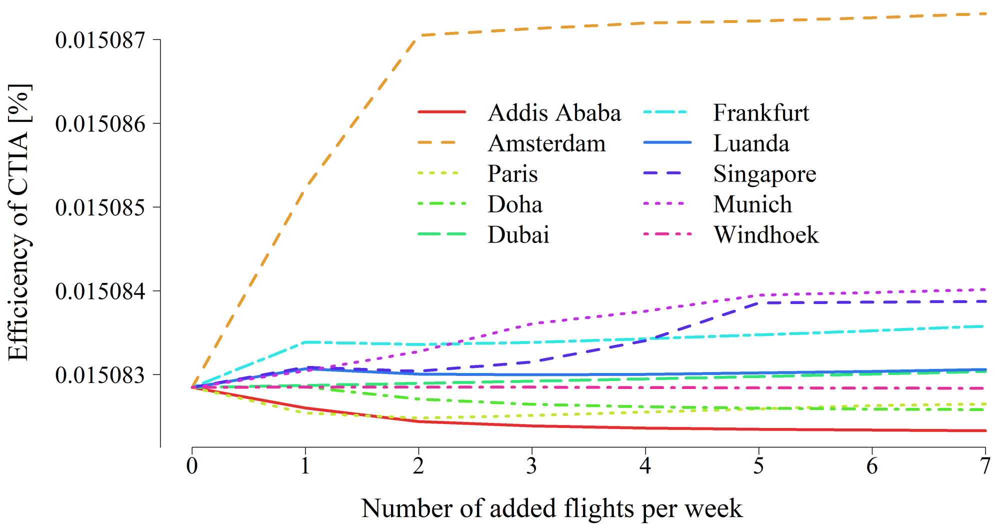


Figure 5.2: The *Winter* efficiency of CTIA resulting from increasing capacity on direct flights between CTIA and each of the above airports. The proportionality heuristic is used.

Figure 5.3 shows a sharp increase in the weighted clustering coefficient after adding one flight. This increase is not because CTIA is now involved in more triangles than before, because no new flights are being added, only existing flights are increasing. Thus, this increase is only due to the greater capacity on the direct flight and thus CTIA becomes more connected to its neighbours. Hubs such as Dubai, Doha and Paris are the most influential here because they are connected to many of CTIA's neighbours. They are more effective at increasing this metric than Singapore for example, an airport which despite being a hub, is not connected to CTIA's neighbours as much as airports in Europe and the Middle-East are. This figure is very similar to Figure 5.4, which uses the switching behaviour heuristic. The shapes of both figures are almost identical, which shows that there is little difference between the two heuristics. The switching behaviour heuristic only accounts for between 4 and 6 connecting airports per direct route, as these airports make up 99% of the traffic. Some destinations have over 50 different routes to reach CTIA, but the smaller routes making up the other 1% are largely insignificant because of how low their seat capacity is. Both Figure 5.3 and 5.4 confirm that the big European routes such as Paris, Munich and Amsterdam have significantly positive impacts on the accessibility of CTIA.

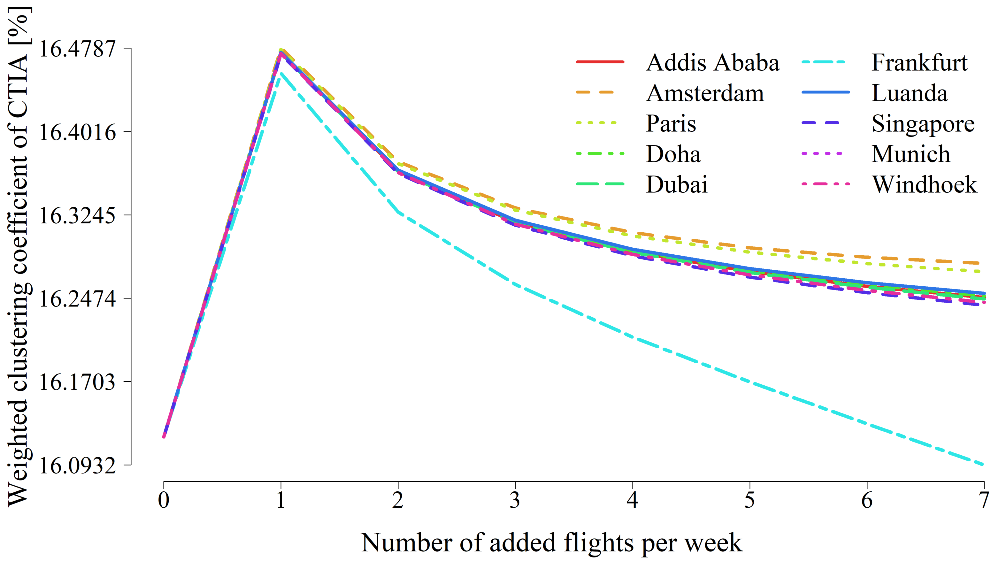


Figure 5.3: The *Summer* weighted clustering coefficient of CTIA resulting from increasing capacity on key direct flights between CTIA and each of the above airports. The proportionality heuristic is used.

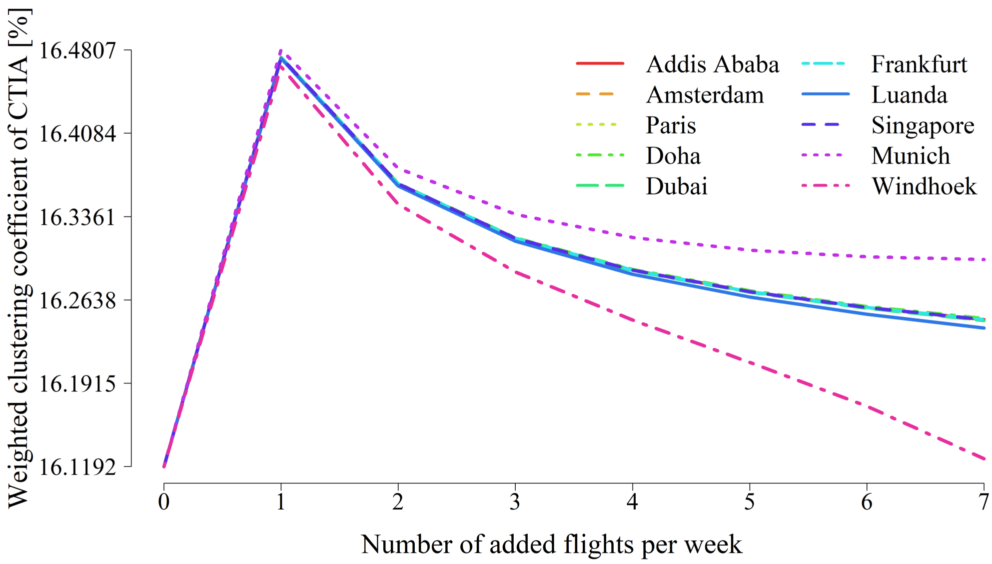


Figure 5.4: The *Summer* weighted clustering coefficient of CTIA resulting from increasing capacity on key direct flights between CTIA and each of the above airports. The switching behaviour heuristic is used.

Based on the efficiency and weighted clustering coefficient graphs for scenario 1, the recommendation would be to prioritise capacity development on the Amsterdam and Munich routes. This challenges the current thinking around route expansion at CTAA, which tends to focus on developing smaller routes. These results imply that there is more room for development on the already well established routes with relatively high passenger volumes. This development will increase the flexibility of travel for passengers, as well as improve CTIA's connectedness within its immediate neighbourhood.

5.2 Decreasing capacity on key flights

This section uses the same key airports as in scenario 1, where the 10 airports were chosen using the vulnerability metric (4.10). Since capacity is *decreased* on key flights as opposed to increased, the model *adds* seat capacity to connecting flights as opposed to stealing seat capacity. For example, removing a certain number of flights per week from Amsterdam – CTIA causes the connecting route from Amsterdam – ORTIA – CTIA to increase by a certain proportion. While adding capacity on connecting routes may not happen in reality, the model assumes that the same number of passengers intend on flying to CTIA as before the direct route capacity reduction. The removal of flights will stop when no capacity is left on a route, as shown in Figure 5.5. After removing 4 flights, 9 out of 10 destinations have no capacity left, and therefore the efficiency stabilises. The Dubai to CTIA route is the exception, because it has a significantly higher capacity, and therefore continues to deviate until 7 flights have been removed. Interestingly, this graph confirms that both Amsterdam and Munich are suitable choices for route expansion because decreasing their capacity results in lower efficiency for CTIA. The reverse is true for Doha, Dubai and Singapore, as they tend to increase the efficiency of CTIA in this scenario.

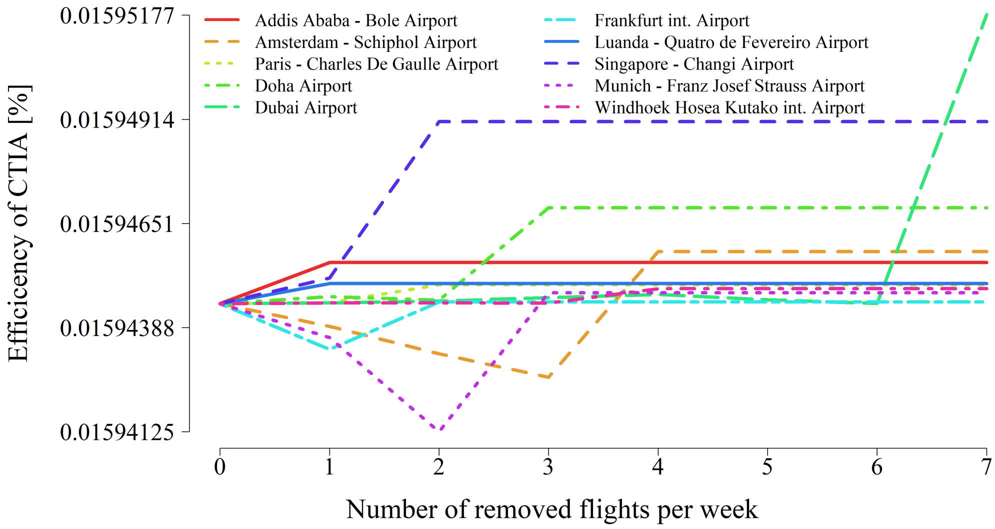


Figure 5.5: The *Summer* efficiency of CTIA resulting from decreasing capacity on key direct flights between CTIA and each of the above airports. The proportionality heuristic is used.

Figure 5.6 shows the betweenness centrality for this scenario. None of the destinations increase the betweenness centrality of CTIA, which is expected because these destinations are key to CTIA's connectedness. Lower capacity on key flights means that fewer passengers will use

CTIA as a connecting airport. A good example of this is Singapore: It is the most isolated and far-reaching destination connected to Cape Town. When in operation, this flight can be used as a link between South-East Asia and Southern Africa. However, removing this flight results in a significant drop in betweenness as CTIA can no longer perform this role of linking the two regions. The Dubai line only drops sharply after removing 6 flights per week, which is explained by its high traffic volumes to CTIA. This route is more robust than others because of its high capacity, but when totally removed, it has a devastating impact on CTIA's betweenness.

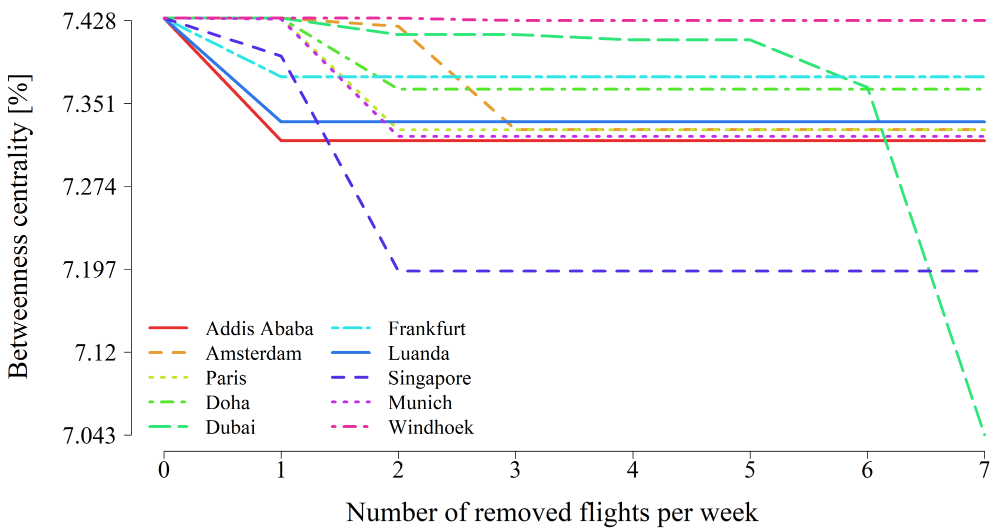


Figure 5.6: The *Summer* betweenness centrality of CTIA resulting from decreasing capacity on key direct flights between CTIA and each of the above airports. The proportionality heuristic is used.

Since the mandate of CTAA is to *increase* the accessibility of CTIA and not decrease it, this scenario is not necessarily as informative as the other two. It does, however, confirm that Amsterdam and Munich are suitable destinations for route expansion, as well as the importance of Singapore to CTIA's role as a link between South-East Asia and Southern Africa.

5.3 Adding flights to the network

The key flights in this section were chosen by using *CapacityData* to determine the 5 busiest international airports in each of the 5 strategic destinations as given by CTAA. These airports can be found in Table 4.3. Starting with efficiency, Figure 5.7 has two noteworthy airports. The small Ethiopian airport in Dire Dawa decreases the efficiency of CTIA significantly after adding one direct flight because it relies heavily on two main connecting routes. These routes (via Addis Ababa and ORTIA) are quickly swallowed up by the newly developed direct route. This is exacerbated by the fact that the airport is so small, as indicated in Table A.1. In contrast, Edmonton International increases CTIA's efficiency significantly in comparison to other destinations. This is interesting because Edmonton is only the 5th busiest Canadian Airport. Bigger airports in North America such as Pearson International in Toronto, Los Angeles International and JFK do not impact CTIA's efficiency as much as would be expected. The effect of Edmonton is partly due to its location in the North-West of Canada. This increases the reach of CTIA significantly due to the distance between the two airports, and thus the efficiency as more passengers from further away can fly direct to Cape Town. The Winter equivalent of this figure, Figure 5.8, confirms that Edmonton and Miami are the outstanding destinations, with Miami triumphing in this case. Mombasa joins Dire Dawa as an airport with a low impact on CTIA's efficiency, which suggests that demand from this region is mostly tourism-driven during the Summer months.

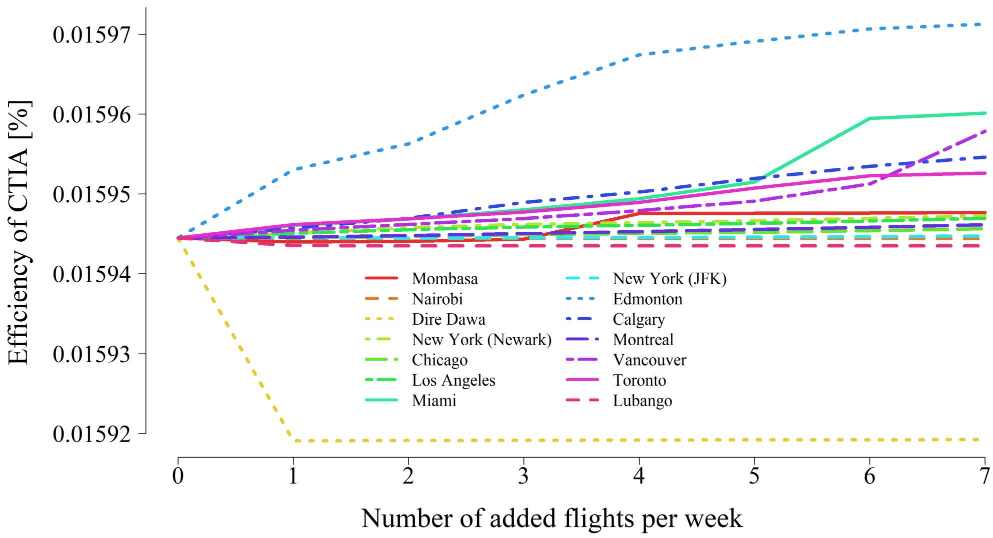


Figure 5.7: The *Summer* efficiency of CTIA resulting from introducing new direct flights between CTIA and each of the above airports. The proportionality heuristic is used.

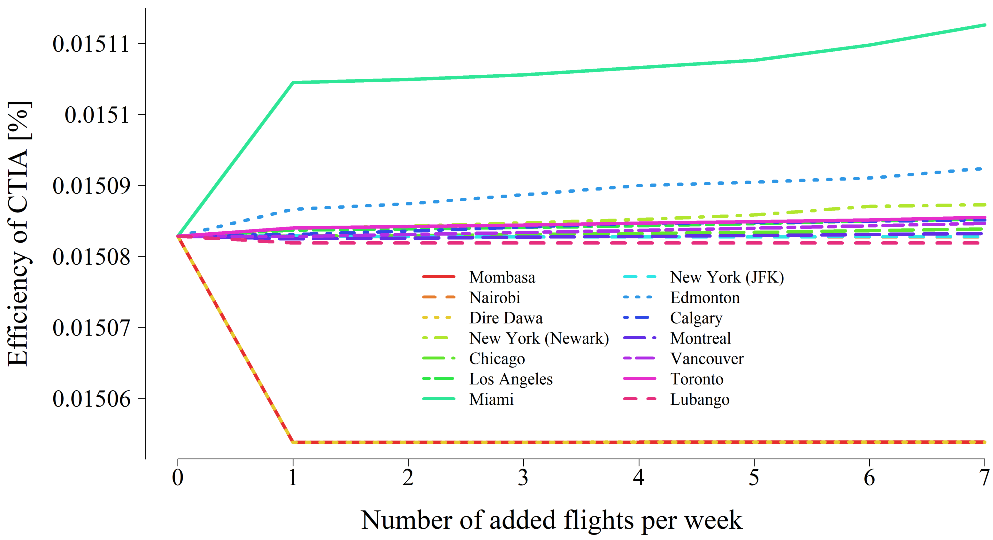


Figure 5.8: The *Winter* efficiency of CTIA resulting from introducing new direct flights between CTIA and each of the above airports. The proportionality heuristic is used.

The lines in Figure 5.9 show a sharp increase in the weighted clustering coefficient after adding one flight, and then the trajectories tend to even out. This sharp increase occurs for all the nodes because CTIA is now involved in more triangles, thus its neighbours are more connected. For example, CTIA is connected to hubs such as London Heathrow, Schipol and Dubai. JFK is also connected to these hubs, therefore a triangle is formed between these nodes

when a direct flight is added and therefore the clustering coefficient increases. The evening out of the line occurs as more flights are added, because smaller routes between CTIA and other airports are swallowed up by the direct route, which breaks up other triangles. This graph confirms the impact of both Edmonton and Miami on CTIA's connectedness and accessibility, as they are the only two lines to continue increasing after the initial spike.
Figures 5.10 and 5.11 compare how the betweenness centrality metric differs for each of the two heuristics. The top performers, JFK and Nairobi, are largely unchanged, with the only noteworthy difference being Newark Liberty International in New York. This difference is a result of the heuristic only stealing seat capacity from the top 5 connecting routes from Newark (London, Paris, Amsterdam, Dublin and Addis Ababa). These connecting routes are all similar in terms of location, thus the discrepancy between travel time and cost for each route is not very big. The model will steal a similar amount of seat capacity from each route (about 20% of the direct route's capacity), therefore, all the connecting routes should remain intact. With the major connecting routes remaining intact, the betweenness centrality of CTIA increases.

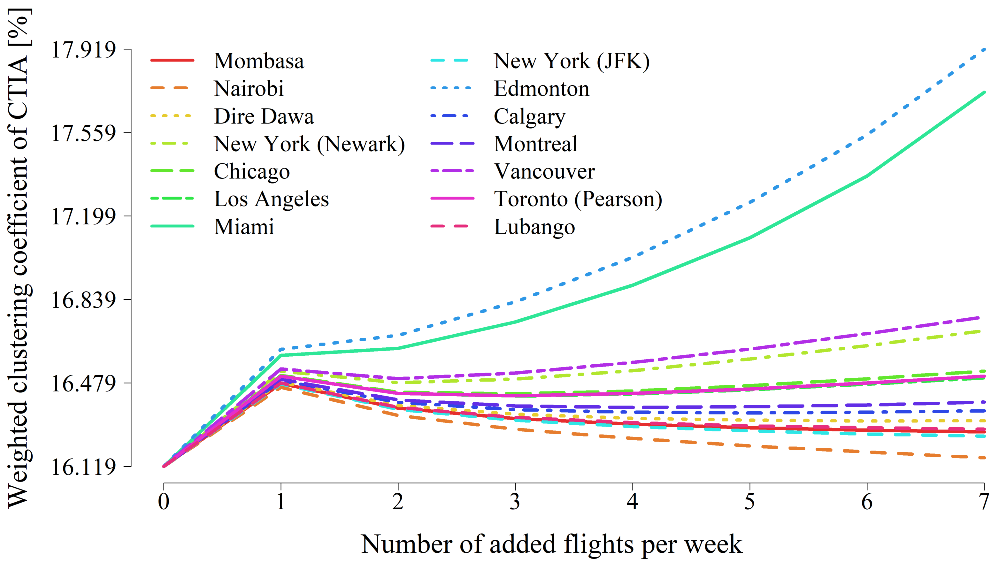


Figure 5.9: The *Summer* weighted clustering coefficient of CTIA resulting from introducing new direct flights between CTIA and each of the above airports. The proportionality heuristic is used.

Based on the graphs for scenario 3, the recommendation would be to develop new routes to JFK, Miami International and Edmonton International. The latter two impacted CTIA's efficiency and weighted clustering coefficient in both Summer and Winter using both heuristics. The JFK route stood out in the betweenness centrality graphs, which is a result of its high degree and popular choice as a connecting airport between North America and South Africa.

5.4 Summary of results and discussions

Scenario 5.1 suggested that the Munich and Amsterdam routes should be expanded. This challenged the current thinking around route expansion because the routes are already well-established. Surprisingly, expanding smaller routes does not significantly improve CTIA's accessibility and connectedness. Scenario 5.2 confirmed this as decreasing the capacity on the Munich and Amsterdam routes results in a decrease in CTIA's efficiency. Scenario 5.3 suggested 3 routes for development, namely JFK, Miami International and Edmonton International.

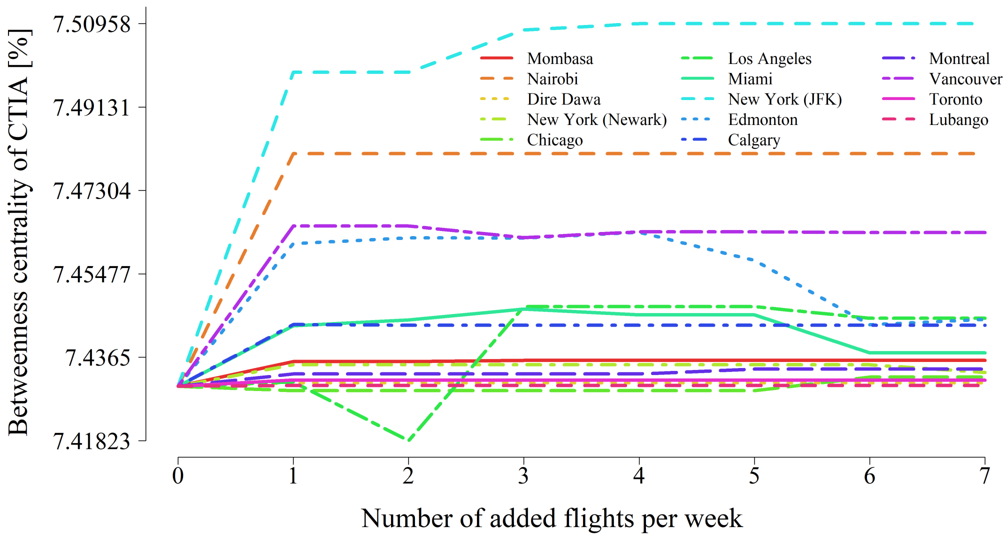


Figure 5.10: The *Summer* betweenness centrality of CTIA resulting from increasing capacity on key direct flights between CTIA and each of the above airports. The proportionality heuristic is used.

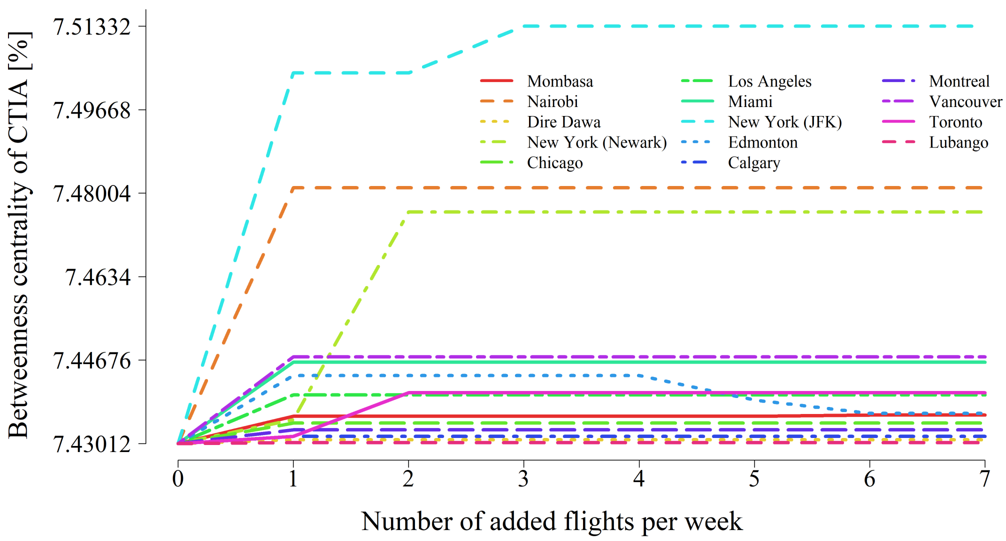


Figure 5.11: The *Summer* betweenness centrality of CTIA resulting from increasing capacity on key direct flights between CTIA and each of the above airports. The switching behaviour heuristic is used.

Chapter 6

Conclusion

This chapter concludes the report using three sections. Section 6.1 discusses the feedback from the company and the answer to the research question. Section 6.2 explains the value of this project to the client and its novelty in terms of CNT application. Section 6.3 is an honest assessment of assumptions made and data gaps which could be improved.

6.1 Company feedback

The research question for this project is: "Can a CNT approach assist decision makers in evaluating the impact of route changes on the Western Cape's air access?". After presenting the project and results to the project manager, David King, positive feedback was received. His answer to the question was: "This model can assist us in some of our decisions" and the analysis is "novel and worthwhile for what we are doing at CTAA". "The methodology and the way it is structured makes a lot of sense", however, he noted that in order to fully function as a DST, the results needed to be easier to understand and explain. In meetings with politicians or the CTAA steering committee, the concern is less to do with the technical aspects of the model and more to do with why a particular route is good for expansion. He suggested that the percentages on the y-axis should be changed to some form of rating which makes it easier for people to understand. Overall, he was happy with the project and noted that he would be interested to see the results using more up-to-date data.

6.2 Research contribution

As mentioned in the literature study in Chapter 2, this project is unique in that it used CNT to focus on only one node in a network. Traditionally, CNT is used to study the behaviour of entire networks or smaller communities within large networks. The concept of a vulnerability study was also used differently in that new links were added to the network in certain scenarios, as opposed to removing links. Ultimately, the value of using CNT in this way was to give the client a network view of the impact of route developments. This network view is important because route developments do not happen in isolation, as other airports feel the knock-on effect of route changes. In this way, the project gives a greater insight into the network than what is currently being offered. The model is also both descriptive *and* prescriptive, therefore future route development decisions can be assisted through using this model.

6.3 Limitations

Due to the complex, dynamic and (at times) unpredictable nature of the WAN, the model cannot predict with absolute certainty the effect of route developments on CTIA. However, the most important factors that affect flight patterns were incorporated into the model to attain a good estimate.

The two heuristics assumed that current flight capacity on routes, the cost of the flight and the length of the flight are the only three factors that influence passenger behaviour. Other factors such as airline alliances and government intervention were not considered. In terms of the input data, there are no doubts about its accuracy, however, it is out of date. The most recent data needs to be used in order to get the most out of the model.

Only three metrics were used to assess the connectedness and accessibility of CTIA. While all three are appropriately used, there are other metrics in the field of CNT that can be used to further develop this project.

Appendices

Appendix A

Table A.1: The top 10 airports in Angola, Canada, Ethiopia, Kenya and the United States by international seat capacity in 2015

Country	Rank	Airport code	Airport name	City	Incoming capacity	Outgoing capacity
Angola	1	LAD	Quatro de Fevereiro Airport	Luanda	1,352,914	1,365,567
	2	SDD	Lubango Airport	Lubango	17,965	17,965
Canada	1	YYZ	Pearson International Airport	Toronto	15,523,932	15,506,619
	2	YVR	Vancouver International Airport	Vancouver	6,405,693	6,410,507
	3	YUL	Pierre Elliott Trudeau International Airport	Montreal	5,681,585	5,673,804
	4	YYC	Calgary International Airport	Calgary	2,931,155	2,940,457
	5	YEG	Edmonton International Airport	Edmonton	1,186,312	1,201,137
	6	YOW	Ottawa International Airport	Ottawa	795,092	784,030
	7	YTZ	Toronto Island Airport	Toronto	677,322	677,322
	8	YHZ	Halifax Stanfield International Airport	Halifax	453,853	414,107
	9	YWG	Winnipeg Airport	Winnipeg	413,014	400,456
	10	YQB	Jean Lesage International Airport	Quebec	344,218	344,318
Ethiopia	1	ADD	Bole Airport	Addis Ababa	6,580,944	6,623,252
	2	DIR	Aba Tenna D Yilma Airport	Dire Dawa	49,207	40,752
Kenya	1	NBO	Jomo Kenyatta International Airport	Nairobi	4,881,456	4,898,976
	2	MBA	Moi International Airport	Mombasa	280,943	284,285
	3	WJR	Wajir Airport	Wajir	68,476	1,200
	4	WIL	Wilson Airport	Nairobi	30,364	30,364
United States	1	JFK	John F. Kennedy International Airport	New York	21,560,345	21,372,085
	2	MIA	Miami International Airport	Miami	15,158,315	15,204,287
	3	LAX	Los Angeles International Airport	Los Angeles	14,634,833	14,494,751
	4	ORD	O'Hare International Airport	Chicago	8,264,935	8,247,323
	5	EWR	Newark Liberty International Airport	New York	8,087,853	8,230,776
	6	SFO	San Francisco International Airport	San Francisco	7,871,481	7,903,092
	7	ATL	Hartsfield-Jackson Atlanta International Airport	Atlanta	7,145,787	7,187,944
	8	IAH	George Bush Intercontinental Airport	Houston	7,157,571	7,112,456
	9	DFW	Dallas/Ft. Worth International Airport	Dallas	4,347,574	4,476,729
	10	IAD	Dulles International Airport	Washington	5,234,054	5,121,830

Appendix B

Department of Industrial & Systems Engineering
Final Year Projects
Identification and Responsibility of Project Sponsors

All Final Year Projects are published by the University of Pretoria on *UPSpace* and thus freely available on the Internet. These publications portray the quality of education at the University and have the potential of exposing sensitive company information. It is important that both students and company representatives or sponsors are aware of such implications.

Key responsibilities of Project Sponsors:

A project sponsor is the key contact person within the company. This person should thus be able to provide the best guidance to the student on the project. The sponsor is also very likely to gain from the success of the project. The project sponsor has the following important responsibilities:

1. Confirm his/her role as project sponsor, duly authorised by the company. Multiple sponsors can be appointed, but this is not advised. The duly completed form will considered as acceptance of sponsor role.
2. Review and approve the Project Proposal, ensuring that it clearly defines the problem to be investigated by the student and that the project aim, scope, deliverables and approach is acceptable from the company's perspective.
3. Review the Final Project Report (delivered during the second semester), ensuring that information is accurate and that the solution addresses the problems and/or design requirements of the defined project.
4. Acknowledges the intended publication of the Project Report on UP Space.
5. Ensures that any sensitive, confidential information or intellectual property of the company is not disclosed in the Final Project Report.

Project Sponsor Details:

Company:	**Cape Town Air Access**
Project Description:	**Analysis of the connectivity of Cape Town International Airport and development of a decision making model to improve air access.**
Student Name:	**Kyle Robertson**
Student number:	**13103408**
Student Signature:	
Sponsor Name:	David King
Designation:	Project Manager
E-mail:	david@wesgro.co.za
Tel No:	021 487 8691
Cell No:	072 622 0510
Fax No:	--
Sponsor Signature:	

Appendix C

Reflection on learning

In the beginning of the year I was highly motivated to make a success of BPJ. The project began well when I was accepted by a company in January to investigate a warehouse location problem that seemed like an ideal Industrial Engineering project. This project fell through in mid-February when I learned of the amount of travelling around Gauteng that would be required. I then found another project, however, within two weeks it became clear that my scope overlapped with another student's project and thus it had to be abandoned. I felt uncomfortable about leaving after committing to it, however, I realised that it was a necessary decision to make given the circumstances. I then found and committed to the CTAA project with much lower hopes of success given the amount of wasted time and two setbacks. I also had no experience with CNT prior to this project, which lowered my hopes even further.

The project seemed overwhelming at first both in terms of scope and complexity. I was however encouraged by my supervisor regarding the vast amount of literature available. This led me to a paper which analysed the topology of the AAN, which was extremely useful in that it gave a good background on the application of CNT in ATNs, as well as describing 19 different network metrics. It also led me to 5 other papers which analysed the topology of various countries. I was however concerned during the first two weeks of reading literature that CNT could only provide a *descriptive* model of the network. The purpose of my project is to build a model which is both *descriptive* and *prescriptive*, which seemed very difficult given what I had read about the use of CNT. Eventually I started reading papers suggested by my supervisor and found the concept of vulnerability studies to be the missing link. I learned that instead of trying to do everything by myself, using the help provided can save a lot of time, especially when navigating the ocean that is academic literature.

From past experience with projects involving a lot of data, I understand that the data-gathering part of a project can be tedious and time-consuming. Fortunately, the data was already available from CTAA, all I had to do was email them and ask for it. This provided one of the biggest learning opportunities for me, because my email to the project manager describing the type of data I needed lacked a lot of detail. I learned that one has to be very specific when asking something from an organisation, because if the email is sent to someone else, they must be able to understand the context and purpose of the request from reading the email alone. I have found communication with the Cape Town-based company to be a challenge, because unlike past projects, I am not able to visit them and ask questions face-to-face. However, the email and phone call correspondence with the project manager has been fruitful and given me a clear idea of the company's requirements.

Once I had the data, the next phase was to analyse it and shape it into a useful format. The famous quote attributed to Einstein about how he would spend 55 minutes defining the problem and then 5 minutes solving it was the attitude I tried to adopt. I decided to enrol for the 'R Programming' Coursera course which improved my R skills significantly. This, along with 'stackoverflow.com' allowed me to quickly find solutions to what I found were common problems. I have come to appreciate LaTeX in recent weeks, despite the early frustration of taking hours to do simple tasks such as creating a table or positioning a figure correctly. It has resulted in a professional-looking document, and is a program that from now on will replace a lot of the work I do in Microsoft Word. Git gave me a large amount of anxiety early on because it would not clone with my repository for the first few weeks of the semester. I had to remain calm and sift through the numerous solutions found online before I found that a simple underscore character in one of the PDF filenames was the cause.

Despite the complexity of the problem and the vast amount of data, the biggest challenge in the next few months will not be entirely a technical one. The point of the project is to provide a useful tool to help people make decisions. I have to therefore be mindful of the end product and how it will be received by the company. Up to now, the company has been very helpful in terms of giving the information and data that I need, and so the hope is that the end product will successfully communicate the findings of the model and that management will have bought in to the concept. In saying this, I have kept in mind that CNT might not be very useful to the company, but there is no way of knowing this without creating an artefact that is well-packaged and user-friendly.

In terms of the broader impact of the project, the benefits of successfully developing more international seat capacity to CTIA are numerous and it is clear that Cape Town has huge potential for growth. I have been following updates of CTAA's new route developments and expansions, and it is clear that the company is making an impact. The new direct flight to Victoria falls in May and the expanded flight to Luanda planned for October are good examples of this, and the hope is that more international flights (particularly from overseas) will be landing at CTIA in the near future.

Bibliography

[1] (2006). World region map. *AnnualReport*.

[2] 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.

[3] Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.

[4] Barthélemy, M., Barrat, A., Pastor-Satorras, R., and Vespignani, A. (2005). Characterization and modeling of weighted networks. *Physica a: Statistical mechanics and its applications*, 346(1):34–43.

[5] Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., and Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature*, 464(7291):1025–1028.

[6] Cardillo, A., Zanin, M., Gómez-Gardenes, J., Romance, M., del Amo, A. J. G., and Boccaletti, S. (2012). Modeling the multi-layer nature of the European Air Transport Network: Resilience and passengers re-scheduling under random failures. *arXiv preprint arXiv:1211.6839*.

[7] Costa, L. d. F., Rodrigues, F. A., Travieso, G., and Villas Boas, P. R. (2007). Characterization of complex networks: A survey of measurements. *Advances in physics*, 56(1):167–242.

[8] Csardi, G. and Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, Complex Systems:1695.

[9] CTAA (2017). Western Cape Air Access Strategic Framework.

[10] Du, W.-B., Zhou, X.-L., Lordan, O., Wang, Z., Zhao, C., and Zhu, Y.-B. (2016). Analysis of the Chinese Airline Network as multi-layer networks. *Transportation Research Part E: Logistics and Transportation Review*, 89:108–116.

[11] Erdös, P. and Rényi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60.

[12] Glass, R. (2000). On design. In *The Journal of Systems and Software*, volume 52.

[13] Guida, M. and Maria, F. (2007). Topology of the Italian airport network: A scale-free small-world network with a fractal structure? *Chaos, Solitons & Fractals*, 31(3):527–536.

[14] Guimerá, R. and Amaral, L. A. N. (2004). Modeling the world-wide airport network. *The European Physical Journal B*, 38(2).

[15] hopper.com (2017). Low Oil Prices and Lower Demand Cause Mean BIG Airfare Savings. [Online; accessed 20-April-2017].

[16] Hossain, M. M. and Alam, S. (2017). A complex network approach towards modeling and analysis of the Australian Airport Network. *Journal of Air Transport Management*, 60:1–9.

[17] iata.org (2017). Air Passenger Monthly Analysis – February 2017. [Online; accessed 19-April-2017].

[18] Investopedia (2017). Purchasing Managers' Index (PMI). [Online; accessed 19-April-2017].

[19] Lordan, O., Sallan, J. M., Simo, P., and Gonzalez-Prieto, D. (2014). Robustness of the air transport network. *Transportation Research Part E: Logistics and Transportation Review*, 68:155–163.

[20] macrotrends.net (2017). Crude Oil Prices - 70 Year Historical Chart. [Online; accessed 20-April-2017].

[21] Manson, N. (2006). Is operations research really research? *Orion*, 22(2):155–180.

[22] mypivots.com (2017). Global Services PMI. [Online; accessed 19-April-2017].

[23] News24.com (2017). Becoming a global aviation leader: Is SA on track? [Online; accessed 21-April-2017].

[24] Travers, J. and Milgram, S. (1967). The small world problem. *Phychology Today*, 1:61–67.

[25] Viljoen, N. M. and Joubert, J. W. (2016). The vulnerability of the global container shipping network to targeted link disruption. *Physica A: Statistical Mechanics and its Applications*, 462:396–409.

[26] vox.com (2017). Heres an example of how Trumps travel ban could really hurt the US economy. [Online; accessed 20-April-2017].

[27] Wang, J., Mo, H., Wang, F., and 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.

[28] Wasserman, S. and Faust, K. (1994). *Social network analysis: Methods and applications*, volume 8. Cambridge University press.

[29] Weng, G., Bhalla, U. S., and Iyengar, R. (1999). Complexity in biological signaling systems. *Science*, 284(5411):92–96.

[30] Xu, Z. and Harriss, R. (2008). Exploring the structure of the US intercity passenger air transportation network: a weighted complex network approach. *GeoJournal*, 73(2):87.