

# Delay Propagation in Air Transportation

MSc. Thesis

J. van den Elshout



# Delay Propagation in Air Transportation

## MSc. Thesis

by

J. van den Elshout

in partial fulfillment of the requirements for the degree of

**Master of Science**  
in Aerospace Engineering

at the Delft University of Technology,  
to be defended publicly on Friday August 25th, 2017 at 14:30.

Student number: 4165063  
Date: August 8, 2017  
Version: Final Report  
Committee: Prof. dr. R. Curran Delft University of Technology  
Dr. ir. B.F. Santos, Delft University of Technology, supervisor  
Dr. A. Menicucci Delft University of Technology  
Dr. M. Menezes, Kedge Business School, supervisor

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.



# Preface

This research project was carried out to partially fulfill the requirements for the degree of Master of Science in Aerospace Engineering at the Delft University of Technology. After diving into the literature of the subject, a clear gap was identified in the approach to analyzing delay propagation within air transportation. The problem of delays in this industry causes both financial consequences and damage to the brand of stakeholders. With this in mind, a research framework was set-up to further expand the scientific knowledge in this topic. All of the research was completed in about 9 months of full-time work.

First of all, I would like to thank my supervisor dr. B.F. Santos for assisting me during my project. Although he has a fairly busy schedule, there was always room for me to ask any questions I might have had or just plain brainstorm with him. He created a pleasant environment for me to thrive in with a perfect balance between giving me freedom and providing guidance. Depending on the situation, dr. Santos provided a critical attitude towards my work while also supporting me.

Next I would also like to thank my other supervisor, dr. M. Menezes from the Kedge Business School in Bordeaux. He was able to join our meetings through Skype providing the much needed critical view from outside Aerospace Engineering. Often he was able to come up with creative questions which provided direction for the research. For the mid-term I was also able to travel to his school where he introduced me to his work and colleagues. I had the honor to present the work I had done up until that point to his department, which was a valuable experience for me.

Finally, I would like to thank my friends and family for constantly supporting me throughout this busy period. First of all, my girlfriend Marit was able to keep me motivated during the more difficult periods. During my research I have performed my work from room 3.15 at the Aerospace faculty and I would like to thank all of the students that accompanied me during that time span, it was a pleasure to find people with the same struggles. My roommates from Spoorsingel 13 were also very supportive when I was at home, providing pleasant company after a long day of work. Towards the end of the research, I took a much needed 3-week vacation with my friends from Dobbel. I would like to thank them for this welcome distraction in the final period. Finally I want to thank my parents for supporting me during this final year, but also throughout my entire study period.

*J. van den Elshout  
Delft, August 2017*



# Abstract

In 1967, Stanley Milgram described the small-world problem which examined the average path length between two individuals in the United States. The conclusion was that this network showed small-world characteristics, where any pair of individuals can be connected by six steps or fewer ('six degrees of separation'). This research is aimed at analyzing the air transportation system using this small-world approach, in particular how delay propagates throughout this network.

Almost all of the stakeholders in the air transportation industry are affected by delays. Delays will increase the operational costs for airlines, stress the limited capacity at airports and increase workload for air traffic control. Passengers are also affected by missing connections and losing time traveling. Overall this problem is estimated to cost billions of dollars annually [Ball et al., 2010].

The root cause of a delay can be a wide variety of issues, for example technical problems with the aircraft, bad weather or airspace congestion. This research focuses on the indirect effects already delayed aircraft have on other aircraft, so-called delay propagation. Aircraft interact with each other, comparable to individual people in a social network. The main research objective is as follows:

*To analyze the delay propagation in the top level of the US air transportation system by combining a complex network theory and an epidemic model approach.*

Three different ways delay is propagated are identified in this research, which are all modeled independent of each other. The first is delay propagation through the transfer of passengers or crew. When these are delayed on a flight, the assumption is made that the connecting flights might be affected as well. The second way is the use of the runway at airports, where the assumption is made that if aircraft are using the runway around the same period of time they will affect each other. Finally the third way is identified as using other airport resources such as fuel trucks, gates and baggage handling.

Compared to a social network, these networks of aircraft are dynamic and constantly evolving. This effect has been countered by defining a way to create a 'snapshot' of the situation. This snapshot aims at providing a static representation of aircraft which have interacted in the recent past. To create them, the status of an aircraft at a certain time is checked. Depending on whether the aircraft is in the air or on the ground, the last two or three movements (departure or arrival) are included in the network.

Once these networks are built, an analysis is done on the neighbors of each node. The nodes are split into two categories, either infected ( $\geq 15$  minutes delay) or susceptible ( $< 15$  minutes delay). For each node, the number of neighbors that are infected are identified. The two groups (infected and susceptible nodes) are compared using a Mann-Whitney U test to determine significant statistical difference between the two. Random sampling is used in order to get an independent subset of nodes and this test is performed 1000 times to achieve a stable result.

Each type of connection leads to a different type of network. The snapshots are proven to be relatively stable in terms of characteristics when snapshots are created in the afternoon and evening, while in the morning they show a growing behavior. The type 1 networks are the smallest in terms of number of nodes and edges, followed by type 2 and type 3 are the largest networks. In terms of small-world behavior, the degree distribution of the nodes are shown to follow a power-law distribution for type 1, while type 2 and 3 networks show a random distribution. The average shortest path length is around 5 for type 1 networks and around 3 for type 2 and 3 networks. Finally the clustering coefficient is around 0.35-0.40 for all types.

Delay propagation effects were finally proven to exist within each type of network. The results were dependent on the amount of delay on a certain day and the number of aircraft which were affected. The maximum degree of separation between which delay propagation effects were found was proven to be three for Type 1, two for Type 2 and one for Type 3. All in all this shows a lot of similarity with for example the spread of obesity in social networks [Christakis and Fowler, 2007], where three degrees was also the final effect.

The model was verified by changing input parameters and investigating the effect on the result. The model behaved as was expected, as the network changed in size according to the changes made. To validate the results, the model was run for completely different days and the results were compared. Most of the research focuses on August 2016, and during validation results from December 2016 and April 2017 are investigated. Similar results in terms of network representation and delay propagation were achieved in these months.

This research can be further improved by adding more data. The techniques of setting up the network are relatively abstract and can be further improved with more data. Also, the geographical scope is limited to domestic flights within the US. It would be interesting to add international flights or focus on other regions such as Europe and Asia to see if the same effects can be found.

# Contents

<b>Preface</b>	<b>i</b>
<b>Abstract</b>	<b>ii</b>
<b>Contents</b>	<b>vi</b>
<b>List of figures</b>	<b>vii</b>
<b>List of tables</b>	<b>ix</b>
<b>List of acronyms</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Air transportation system. . . . .	1
1.1.1 History of the air transport industry . . . . .	1
1.1.2 Business models . . . . .	2
1.1.3 Delays in air transport . . . . .	3
1.2 Report structure . . . . .	4
<b>2 Literature review</b>	<b>5</b>
2.1 Delay propagation . . . . .	5
2.2 Complex network theory . . . . .	7
2.2.1 Basic metrics . . . . .	8
2.2.2 Topological structures . . . . .	10
2.2.3 Centrality measures . . . . .	11
2.2.4 Analysis of complex networks . . . . .	13
2.3 ATS network analysis . . . . .	15
2.4 Delay propagation using complex networks. . . . .	19
2.5 State-of-the-art . . . . .	20
<b>3 Research framework</b>	<b>23</b>
3.1 Problem statement . . . . .	23
3.2 Research objective . . . . .	24
3.3 Model requirements . . . . .	24
3.4 Research scope . . . . .	25
3.5 Contribution. . . . .	25
<b>4 Methodology</b>	<b>27</b>
4.1 Data . . . . .	27
4.1.1 Input. . . . .	27
4.1.2 Processing . . . . .	30
4.2 Delay propagation . . . . .	31
4.2.1 Definitions. . . . .	31
4.2.2 Delay analysis . . . . .	32
4.3 Network representation . . . . .	32
4.3.1 Type of connections . . . . .	33
4.3.2 Snapshot . . . . .	35
4.3.3 Network measures . . . . .	36
4.4 Statistical analysis . . . . .	38
4.4.1 Sampling . . . . .	38
4.4.2 Mann-Whitney U test. . . . .	38
4.5 Summary . . . . .	40

<b>5</b>	<b>Results</b>	<b>43</b>
5.1	General delay analysis . . . . .	43
5.2	Network representation . . . . .	47
5.2.1	Type 1 network . . . . .	47
5.2.2	Type 2 network . . . . .	51
5.2.3	Type 3 network . . . . .	55
5.2.4	Comparison . . . . .	57
5.3	Delay propagation . . . . .	58
5.3.1	Neighbor analysis throughout an average day . . . . .	58
5.3.2	Comparison with other days . . . . .	70
<b>6</b>	<b>Verification and validation</b>	<b>73</b>
6.1	Verification and sensitivity analysis . . . . .	73
6.1.1	Type 1 time window . . . . .	73
6.1.2	Type 2 time window . . . . .	77
6.1.3	6 hours snapshot . . . . .	78
6.1.4	15 minutes delay . . . . .	80
6.1.5	Predecessors vs. successors . . . . .	81
6.2	Validation . . . . .	82
<b>7</b>	<b>Conclusion and recommendations</b>	<b>85</b>
7.1	Conclusion . . . . .	85
7.2	Limitations and recommendations . . . . .	87
7.2.1	Limitations . . . . .	87
7.2.2	Recommendations for future work . . . . .	87
<b>A</b>	<b>BTS data profile</b>	<b>89</b>
<b>B</b>	<b>Neighbor results</b>	<b>93</b>
B.1	Neighbor analysis 15-08-2016, Type 1, Degree 2-3 . . . . .	93
B.2	Probability of being infected 15-08-2016, Type 1, Degree 2-4 . . . . .	96
B.3	Neighbor analysis 15-08-2016, Type 2, Degree 2 . . . . .	98
B.4	Probability of being infected 15-08-2016, Type 2, Degree 2-3 . . . . .	99
B.5	Neighbor analysis 15-08-2016, Type 3, Degree 2 . . . . .	100
B.6	Probability of being infected 15-08-2016, Type 3, Degree 2-3 . . . . .	102
B.7	Multiple days, Type 1, Degree 1-4 . . . . .	103
B.8	Multiple days, Type 2, Degree 1-4 . . . . .	107
B.9	Multiple days, Type 3, Degree 1-4 . . . . .	111
<b>C</b>	<b>Validation results</b>	<b>115</b>
C.1	Snapshot time limit . . . . .	115
C.2	Delay definition . . . . .	117
C.3	Validation similar days . . . . .	120
	<b>Bibliography</b>	<b>121</b>

# List of figures

1.1	PP Network vs HS network [Cook and Goodwin, 2008]. . . . .	2
1.2	The causes of delay in US air transport for August 2016 [U.S. Department of Transportation, 2016]. . . . .	4
2.1	A delay propagation tree used by [AhmadBeygi et al., 2008]. . . . .	7
2.2	The seven bridges of Königsberg. . . . .	8
2.3	An example of a regular graph [Wilson, 1996]. . . . .	10
2.4	Random graphs for different $p$ [Perseguers et al., 2010]. . . . .	10
2.5	Small-world network [Watts and Strogatz, 1998]. . . . .	11
2.6	A complex network divided into $k$ -shells [Garas et al., 2012]. . . . .	12
2.7	Epidemic models (SIS vs SIR) with susceptible (S), infected (I) and recovered (R) states, along with infection rates $\beta$ and recovery rate $\mu$ [Pastor-Satorras and Vespignani, 2001].	14
2.8	The worldwide ATS, cities with highest degree vs. highest betweenness centrality [Guimerà et al., 2005]. . . . .	17
2.9	PP Network of Ryanair (left) vs MHS network of Lufthansa (right) [Zanin and Lillo, 2013].	18
3.1	Overview of the status quo and contribution of the proposed research, both for the scientific field as the industry. . . . .	26
4.1	US timezones . . . . .	30
4.2	Time-space network of type 1 connection from aircraft A to aircraft B. . . . .	34
4.3	Time-space network of type 2 connection from aircraft A to aircraft B. . . . .	34
4.4	Visualization of snapshot creation. . . . .	36
4.5	Rejection region of one-tailed tests and two-tailed tests. . . . .	39
4.6	Functional flow diagram of model. . . . .	41
5.1	The total arrival delay of each day in August 2016. . . . .	43
5.2	The number of aircraft performing flights and the number of infected aircraft. . . . .	44
5.3	The arrival delay versus the percentage of aircraft infected. . . . .	44
5.4	The development of arrival delays throughout individual days. . . . .	45
5.5	The magnitude of different types of delays throughout August 2016. . . . .	45
5.6	The magnitude of different types of delays on August 8th and 19th. . . . .	46
5.7	Delay development at the top 10 largest airports of the US on August 19th, 2016. . . . .	46
5.8	Development of network characteristics for type 1 snapshots on 15-08-2016. . . . .	48
5.9	Degree distribution for three snapshots of type 1 on 15-08-2016. . . . .	49
5.10	The degree distribution on logarithmic scale (solid line) with their best power-law fit (dotted line) for four snapshots of Type 1 on 15-08-2016. . . . .	50
5.11	20:00 type 1 snapshot on August 15h 2016. Green nodes are susceptible and red nodes are infected. . . . .	51
5.12	Development of network characteristics for snapshots on 15-08-2016. . . . .	53
5.13	Degree distribution for three type 2 snapshots on 15-08-2016. . . . .	54
5.14	Degree distribution for 20:00 snapshot on 15-08-2016 and its Poisson fit. . . . .	54
5.15	Development of network characteristics for type 3 snapshots on 15-08-2016. . . . .	55
5.16	Degree distribution for three snapshots of Type 3 on 15-08-2016. . . . .	56
5.17	Degree distribution for 20:00 snapshot on 15-08-2016 and its Poisson fit. . . . .	57
5.18	The distribution of the percentage of infected 1st degree neighbors of susceptible and infected nodes. . . . .	59
5.19	Boxplots of the z-score results from the Mann-Whitney U tests for 1st degree type 1 neighbors. . . . .	60

5.20	Bar chart summarizing median values for percentage of infected neighbors in type 1 networks. . . . .	62
5.21	Minimum number of infected first degree neighbors versus percentage of infected nodes for type 1 snapshots. . . . .	63
5.22	The distribution of the percentage of infected 1st degree neighbors of susceptible and infected nodes. . . . .	64
5.23	Boxplots of the z-score results from the Mann-Whitney U tests for 1st degree type 2 neighbors. . . . .	64
5.24	Bar chart summarizing median values for percentage of infected neighbors in type 2 networks. . . . .	65
5.25	Minimum number of infected first degree neighbors versus percentage of infected nodes for type 2 snapshots. . . . .	66
5.26	The distribution of the percentage of infected 1st degree neighbors of susceptible and infected nodes. . . . .	67
5.27	Boxplots of the z-score results from the Mann-Whitney U tests for 1st degree type 3 neighbors. . . . .	67
5.28	Bar chart summarizing median values for percentage of infected neighbors in type 3 networks. . . . .	69
5.29	Minimum number of infected first degree neighbors versus percentage of infected nodes for type 3 snapshots. . . . .	69
6.1	The development of arrival delays throughout similar days used in validation. . . . .	83
B.1	The distribution of the percentage of infected 2nd degree neighbors of susceptible and infected nodes. . . . .	94
B.2	Boxplots of the z-score results from the Mann-Whitney U tests for 2nd degree type 1 neighbors. . . . .	94
B.3	The distribution of the percentage of infected 3rd degree neighbors of susceptible and infected nodes. . . . .	95
B.4	Boxplots of the z-score results from the Mann-Whitney U tests for 3rd degree type 1 neighbors. . . . .	96
B.5	Minimum number of infected second degree neighbors versus percentage of infected nodes for type 1 snapshots. . . . .	96
B.6	Minimum number of infected third degree neighbors versus percentage of infected nodes for type 1 snapshots. . . . .	97
B.7	Minimum number of infected fourth degree neighbors versus percentage of infected nodes for type 1 snapshots. . . . .	97
B.8	The distribution of the percentage of infected 2nd degree neighbors of susceptible and infected nodes. . . . .	98
B.9	Boxplots of the z-score results from the Mann-Whitney U tests for 2nd degree type 2 neighbors. . . . .	99
B.10	Minimum number of infected second degree neighbors versus percentage of infected nodes for type 2 snapshots. . . . .	99
B.11	Minimum number of infected third degree neighbors versus percentage of infected nodes for type 2 snapshots. . . . .	100
B.12	The distribution of the percentage of infected 2nd degree neighbors of susceptible and infected nodes. . . . .	101
B.13	Boxplots of the z-score results from the Mann-Whitney U tests for 2nd degree type 3 neighbors. . . . .	101
B.14	Minimum number of infected second degree neighbors versus percentage of infected nodes for type 3 snapshots. . . . .	102
B.15	Minimum number of infected third degree neighbors versus percentage of infected nodes for type 3 snapshots. . . . .	102

# List of tables

2.1	Research modeling the ATS as a complex network . . . . .	16
4.1	List of airlines included in the dataset. . . . .	28
4.2	Data included in TransStats dataset. . . . .	29
4.3	Data included in <code>OpenFlights.org</code> dataset. . . . .	29
4.4	List of airports and airlines which can connect to flights of a certain airline. . . . .	33
5.1	Generating type 1 snapshot for several different times on 15-08-2016. . . . .	47
5.2	Best fit power-law coefficients for type 1 snapshots on 15-08-2016. . . . .	49
5.3	Generating type 1 snapshot for several different days in August 2016. . . . .	50
5.4	Generating type 2 snapshot for several different times on 15-08-2016. . . . .	52
5.5	Poisson distribution characteristics for type 2 snapshots on 15-08-2016. . . . .	52
5.6	Generating type 2 snapshot for several different days in August 2016. . . . .	54
5.7	Generating type 3 snapshot for several different times on 15-08-2016. . . . .	55
5.8	Poisson distribution characteristics for type 3 snapshots on 15-08-2016. . . . .	56
5.9	Generating type 3 snapshot for several different days in August 2016. . . . .	57
5.10	Mean value of network characteristics for different connection types. . . . .	58
5.11	Shape of degree distribution for the different connection types. . . . .	58
5.12	Number of 1st degree neighbors and percentage of neighbors infected for type 1 snapshots. . . . .	59
5.13	Z-scores calculated using the Mann-Whitney U test for first degree neighbors in type 1 snapshots. . . . .	61
5.14	Z-scores calculated using the Mann-Whitney U test for second degree neighbors in type 1 snapshots. . . . .	61
5.15	Z-scores calculated using the Mann-Whitney U test for third degree neighbors in type 1 snapshots. . . . .	61
5.16	Number of 1st degree neighbors and percentage of neighbors infected for type 2 snapshots. . . . .	63
5.17	Z-scores calculated using the Mann-Whitney U test for first degree neighbors in type 2 snapshots. . . . .	65
5.18	Z-scores calculated using the Mann-Whitney U test for second degree neighbors in type 2 snapshots. . . . .	65
5.19	Number of 1st degree neighbors and percentage of neighbors infected for type 3 snapshots. . . . .	66
5.20	Z-scores calculated using the Mann-Whitney U test for first degree neighbors in type 3 snapshots. . . . .	68
5.21	Z-scores calculated using the Mann-Whitney U test for second degree neighbors in type 3 snapshots. . . . .	68
5.22	Summarized results for different days of August. . . . .	70
6.1	Effect on network representation when changing type 1 scheduled time window. . . . .	74
6.2	Effect on first degree neighbors when changing type 1 scheduled time window. . . . .	75
6.3	Effect on Mann-Whitney U test when changing type 1 scheduled time window. . . . .	75
6.4	Effect on network representation when changing type 1 actual time window. . . . .	76
6.5	Effect on first degree neighbors when changing type 1 actual time window. . . . .	76
6.6	Effect on Mann-Whitney U test when changing type 1 scheduled time window. . . . .	76
6.7	Effect on network representation when changing type 2 time window. . . . .	77
6.8	Effect on first degree neighbors when changing type 2 time window. . . . .	78
6.9	Effect on Mann-Whitney U test when changing type 2 time window. . . . .	78

6.10	Effect on network representation when changing type 1 snapshot time limit. . . . .	79
6.11	Effect on first degree neighbors when changing type 1 snapshot time limit. . . . .	79
6.12	Effect on Mann-Whitney U test when changing type 1 snapshot time limit. . . . .	80
6.13	Effect on network representation when changing type 1 definition of infected neighbors. . . . .	80
6.14	Effect on first degree neighbors when changing type 1 definition of infected neighbors. . . . .	80
6.15	Effect on first degree neighbors when changing type 1 definition of infected aircraft and infected neighbors. . . . .	81
6.16	Effect on Mann-Whitney U test when changing type 1 delay definition. . . . .	81
6.17	Effect on network representation using successors instead of predecessors in type 1 networks. . . . .	81
6.18	Effect on first degree neighbors in type 1 networks when using successors instead of predecessors. . . . .	82
6.19	Effect on Mann-Whitney U test when using successors instead of predecessors. . . . .	82
6.20	20:00 snapshots for all connection types on 3 similar days. . . . .	82
6.21	Results for first degree neighbors on 3 similar days. . . . .	84
6.22	Mann-Whitney U test results for first degree neighbors of similar days. . . . .	84
A.1	Detailed definitions of BTS dataset. . . . .	89
B.1	Number of 2nd degree neighbors and percentage of neighbors infected for type 1 snapshots. . . . .	93
B.2	Number of 3rd degree neighbors and percentage of neighbors infected for type 1 snapshots. . . . .	95
B.3	Number of 2nd degree neighbors and percentage of neighbors infected for type 2 snapshots. . . . .	98
B.4	Number of 2nd degree neighbors and percentage of neighbors infected for type 3 snapshots. . . . .	100
B.5	Z-scores calculated using the Mann-Whitney U test for first degree neighbors in type 1 snapshots on multiple days. . . . .	103
B.6	Z-scores calculated using the Mann-Whitney U test for second degree neighbors in type 1 snapshots on multiple days. . . . .	104
B.7	Z-scores calculated using the Mann-Whitney U test for third degree neighbors in type 1 snapshots on multiple days. . . . .	105
B.8	Z-scores calculated using the Mann-Whitney U test for fourth degree neighbors in type 1 snapshots on multiple days. . . . .	106
B.9	Z-scores calculated using the Mann-Whitney U test for first degree contacts in type 2 snapshots on multiple days. . . . .	107
B.10	Z-scores calculated using the Mann-Whitney U test for second degree contacts in type 2 snapshots on multiple days. . . . .	108
B.11	Z-scores calculated using the Mann-Whitney U test for third degree contacts in type 2 snapshots on multiple days. . . . .	109
B.12	Z-scores calculated using the Mann-Whitney U test for fourth degree contacts in type 2 snapshots on multiple days. . . . .	110
B.13	Z-scores calculated using the Mann-Whitney U test for first degree contacts in type 3 snapshots on multiple days. . . . .	111
B.14	Z-scores calculated using the Mann-Whitney U test for second degree contacts in type 3 snapshots on multiple days. . . . .	112
B.15	Z-scores calculated using the Mann-Whitney U test for third degree contacts in type 3 snapshots on multiple days. . . . .	113
B.16	Z-scores calculated using the Mann-Whitney U test for fourth degree contacts in type 3 snapshots on multiple days. . . . .	114
C.1	Effect on network representation when changing type 2 snapshot time limit. . . . .	115
C.2	Effect on first degree neighbors when changing type 2 snapshot time limit. . . . .	116
C.3	Effect on Mann-Whitney U test when changing type 2 snapshot time limit. . . . .	116
C.4	Effect on network representation when changing type 3 snapshot time limit. . . . .	116
C.5	Effect on first degree neighbors when changing type 3 snapshot time limit. . . . .	117

---

C.6	Effect on Mann-Whitney U test when changing type 3 snapshot time limit. . . . .	117
C.7	Effect on network representation when changing type 2 delay definition. . . . .	117
C.8	Effect on first degree neighbors when changing type 2 definition of infected neighbors. . . . .	118
C.9	Effect on first degree neighbors when changing type 2 definition of infected aircraft and infected neighbors . . . . .	118
C.10	Effect on Mann-Whitney U test when changing type 2 delay definition. . . . .	118
C.11	Effect on network representation when changing type 3 delay definition. . . . .	118
C.12	Effect on first degree neighbors when changing type 3 definition of infected neighbors. . . . .	119
C.13	Effect on first degree neighbors when changing type 3 definition of infected aircraft and infected neighbors . . . . .	119
C.14	Effect on Mann-Whitney U test when changing type 3 delay definition. . . . .	119
C.15	Mann-Whitney U test results for second degree neighbors of similar days. . . . .	120
C.16	Mann-Whitney U test results for first degree neighbors of similar days. . . . .	120



# List of acronyms

Acronym	Description
<b>AA</b>	American Airlines
<b>AB</b>	Air Berlin
<b>AF</b>	Air France
<b>ATC</b>	Air traffic control
<b>ATM</b>	Air traffic management
<b>ATS</b>	Air transportation system
<b>BA</b>	British Airways
<b>BTS</b>	Bureau of Transportation Statistics
<b>CRS</b>	Computer reservation system
<b>CZ</b>	China Southern Airlines
<b>DL</b>	Delta Airlines
<b>DOT</b>	US Department of Transportation
<b>ER</b>	Erdős-Rényi
<b>FR</b>	Ryanair
<b>FSC</b>	Full-service carrier
<b>GMT</b>	Greenwich Mean Time
<b>HS</b>	Hub-and-spoke network
<b>IATA</b>	International Air Transport Association
<b>ICAO</b>	International civil aviation organization
<b>IQR</b>	Interquartile range
<b>KL</b>	KLM Royal Dutch Airlines
<b>LCC</b>	Low-cost carrier
<b>LH</b>	Lufthansa
<b>OS</b>	Austrian Airlines
<b>MHS</b>	Multi-hub-and-spoke network
<b>MU</b>	China Eastern Airlines
<b>NAS</b>	National airspace system (USA)
<b>PP</b>	Point-to-point network
<b>SA</b>	Star Alliance
<b>SF</b>	Scale-free
<b>SWN</b>	Small-world network
<b>U2</b>	EasyJet
<b>UA</b>	United Airlines
<b>US</b>	US Airways
<b>WN</b>	Southwest Airlines
<b>WS</b>	Watts-Strogatz



# Introduction

In research by Christakis and Fowler [2007], the spread of obesity throughout a group of US people is analyzed. The conclusion was that if the friends of your friends of your friends had obesity, it would increase the probability of you getting obesity. In the research described in this report, an analogy is made with the air transportation industry. How do aircraft interact and does this affect the spreading of their disease (delay) around the network?

This report is written to complete the Master of Science thesis project at the Faculty of Aerospace Engineering at the Delft University of Technology. The problem will be shortly introduced in Section 1.1 while the report structure is further explained in Section 1.2.

## 1.1. Air transportation system

The global air transportation system (ATS) plays an important role in our society and particularly in our economy. Along with systems such as power grids and communication networks, they allow society to stay connected on a global scale. In this chapter, the ATS will be reviewed by first presenting a short history in Section 1.1.1, then the different business models are discussed in Section 1.1.2 and furthermore the problem of delays is introduced in Section 1.1.3.

### 1.1.1. History of the air transport industry

The role of the air transport industry in creating a global economy was tremendous. Both its own operations and the effect on related industries such as manufacturing and tourism, have caused this industry to impact virtually every country in the world.

The beginning of the air transport industry is characterized as a highly regulated and subsidized market [Cook, 1996]. The Chicago convention in 1944 set the framework for the modern air transportation system [ICAO, 1944]. Currently, the freedoms of the air regulate international traffic and determine what airlines can fly to what countries. In the Chicago convention the first freedom of the air was introduced (under a different name), which allowed airlines to fly through foreign airspace without restrictions.

Major technological developments helped the industry grow during this regulated period. Examples were the introduction of the jet aircraft for commercial use in the 1950s, followed by the introduction of the wide-body 'jumbo-jets' in the 1970s [Belobaba et al., 2009].

In 1978, the Airline Deregulation Act was passed in the US which allowed the airlines to determine their own routes and fares [Cook, 1996]. This market freedom led to a period of rapid expansion of existing carriers and allowed new-entrants to proliferate. Only since this deregulation have cost efficiency, competitive behavior and profitability become important issues for airlines [Belobaba et al.,

2009]. Over the last 30 years, the growth of air traffic demand has increased by approximately 5% each year.

Since the events of 9/11 and the corresponding economic downturn, the air transport industry has been faced with problems concerning the profitability. The main challenges for the industry in the coming years have been defined by Belobaba et al. [2009] to be sustaining profitability, ensuring safety and security and developing sustainable air transportation infrastructure.

### 1.1.2. Business models

After the deregulation act in 1978, two different business models have developed in the scheduled passenger air transport industry. The first is the full-service carrier (FSC) and the other is the low-cost carrier (LCC) [Lordan, 2014b]. These two business models are also related to a network configuration, which determines the structure of the operating costs and the strategy of competitors [Aguirregabiria and Ho, 2010]. The network configurations that can be identified are the hub-and-spoke (HS) network, the multi-hub-and-spoke (MHS) network and the point-to-point (PP) network. The difference between the PP and HS network configurations is illustrated in Figure 1.1. The PP network offers direct flights between two airports while the HS network offers connecting flights through the hub airport.

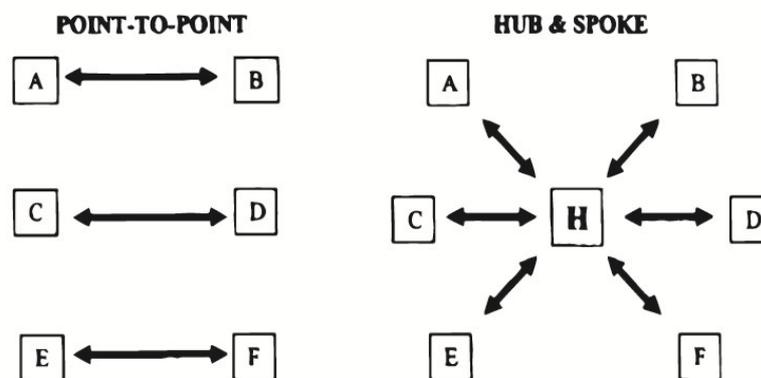


Figure 1.1: PP Network vs HS network [Cook and Goodwin, 2008].

Usually the national flag carriers which were founded during the heavily regulated period use the FSC model. This model focuses on serving the needs of the passengers by offering different classes and a wide variety of destinations through their hub(s), thus adopting a HS or MHS network configuration. These airlines usually form alliances in order to offer even more destinations using the networks of their partners. This does make the network extremely complex and increases the cost of scheduling and fleet allocation.

The main advantage of using a (M)HS network configuration is that it can efficiently create a large amount of origin-destination pairs, and thus create enough demand to fill the aircraft. Another advantage is the ability to centralize the main operations (maintenance, staff) at the hub airport. The major drawback of (M)HS networks are the traffic peaks which will occur at the hub airport in order to create the connections, leading to extra delays and increases in turnover time. Also, because all of the traffic is directed through the hub, the travel time is increased for passengers and the fuel/crew costs are increased for the airline.

Airlines who use the LCC model have started appearing since the deregulation of the air transport industry [Belobaba et al., 2009]. These airlines use a PP network, which leads to less destinations pairs when compared to the networks of FSC. The demand for these flights is increased by lowering the fares. A LCC reduces their costs by operating a limited amount of aircraft types in the fleet, thus lowering maintenance costs, and flying to secondary airports which have lower landing fees [Cook and Goodwin, 2008].

Pure PP networks almost never exist in the case of airlines, as LCCs plan the majority of their routes from a set of base airports [Belobaba et al., 2009]. However, as no connection services are offered in these airports the traffic peaks disappear and the operational efficiency is increased in terms of fuel and crew cost. The main drawback is of course that more routes need to be offered in order to link the same number of destinations. Usually LCCs are limited to regional routes that have a high enough direct demand to have a load factor that is high enough for profitability. Global routes are thus out of reach for LCCs with a PP network, as they have less direct demand than local routes.

### 1.1.3. Delays in air transport

Delays in air transport are a problem for numerous stakeholders, namely the airlines, passengers and society in general. For airlines the main issue is that delays drive up the operating cost. It is estimated that these increased costs for airlines annually add up to US\$ 8.3 billion (2007) in the USA [Ball et al., 2010]. As the air travel demand keeps growing, reducing these delays has been a long-term policy goal. It is also estimated that the cost from the loss of demand due to delays annually adds up to US\$ 3.9 billion [Ball et al., 2010].

Next to the increased cost for airlines, delays also cause direct problems for passengers. When using air transportation passengers go to catch the four o'clock flight to London, which is also the product they buy from the airline. It is up to the airlines to do everything in their power to keep that promise. The US Department of Transportation (DOT) handles complaints that commercial passengers have about their experiences using air transport. Each month the DOT presents these complaints in the Air Travel Consumer Report. The most recent issue available at this time is October 2016, which reports that the 40% of the total number of complaints is related to delays [U.S. Department of Transportation, 2016]. In the USA, the total cost which is incurred by the passengers adds up to US\$ 16.7 billion for 2007 alone [Ball et al., 2010]. Furthermore the whole society is affected as the negative impact on the GDP is estimated at US\$ 4 billion.

There are different definitions of delay in air transport. The most common definition and visible for passengers is the difference between scheduled and actual departure/arrival times. Throughout the rest of this literature review, this will be considered the definition of delay.

According to the monthly reports of the DOT, 77.6% of the flights operated in the US in August 2016 were on-time [U.S. Department of Transportation, 2016]. The definition for an on-time flight according to the DOT is that it arrives within 15 minutes after the scheduled arrival time shown in the Computer Reservations Systems (CRS) of the respective airline. The DOT categorizes the delay into five different categories depending on the cause of the delay:

1. Air carrier delay: the cause was within the control of the airline, for instance maintenance or crew problems. (6.04%)
2. Extreme weather delay: the meteorological conditions (forecast or actual) were significant enough to prevent or delay the flight, this is judged by the respective airlines. (0.67%)
3. National Aviation System (NAS) delay: these delays are caused by the NAS, such as airport operations, congested airspace or air traffic control (ATC) limitations. (6.35%)
4. Security delay: the cause of the delay lies in a security issue and it can be anything from long lines (> 29 minutes) at the passenger security check to the evacuation of the terminal. (0.04%)
5. Late arriving aircraft delay: if the aircraft arrived late due to a previously delayed flight. (7.65%)

The value within parenthesis behind each category shows what percentage of the total domestic flights in the US were affected by each respective category in August 2016 [U.S. Department of Transportation, 2016], showing that each category has a varying impact on the delays. Furthermore the DOT considers a flight to be canceled (1.38%) if it is not operated within seven days after the scheduled time in the CRS. Diverted flights (0.30%) are considered flights which were operated from the scheduled origin airport to a point different from the scheduled destination airport. Figure 1.2 summarizes these percentages in a pie chart.

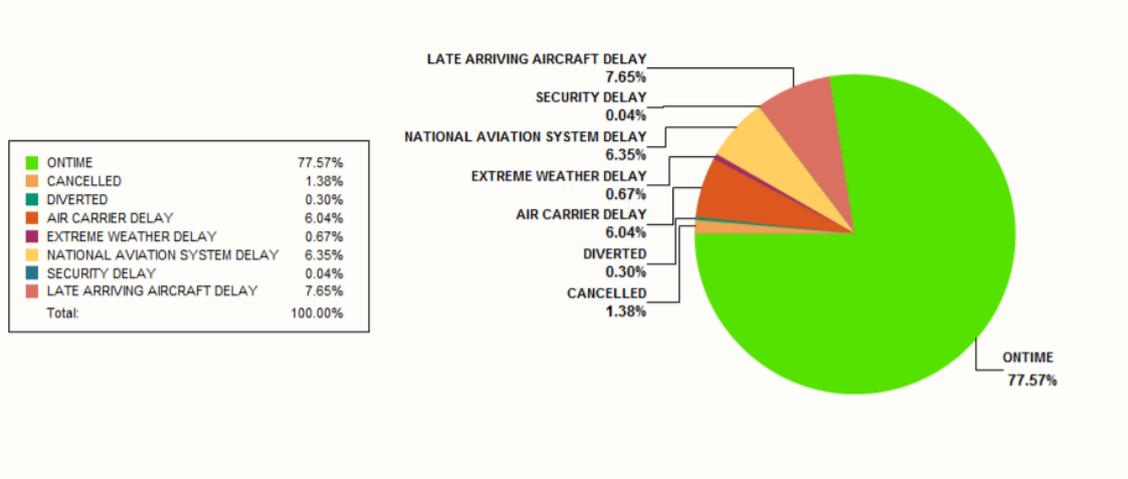


Figure 1.2: The causes of delay in US air transport for August 2016 [U.S. Department of Transportation, 2016].

## 1.2. Report structure

This report contains every step of the research performed. In Chapter 2 a literature review summarizes previous research and present the state-of-the-art in terms of scientific contributions. From this literature review, several research gaps have been identified. This research attempts to fill these gaps and the goals, objectives and requirements are summarized in Chapter 3.

The methodology is explained in Chapter 4. The model is explained from the data input to the output, carefully explaining each step. Also, certain assumptions were made in terms of the methodology and these are explained as well. The results are given in Chapter 5. Here a general delay analysis is presented, along with the resulting network and delay characteristics from the model. To prove the results have any meaning, the methodology is verified and the results validated in Chapter 6. Finally the main conclusions and recommendations are given in Chapter 7.

# 2

## Literature review

In this chapter a literature review is presented. The main goals are to summarize what research has been done and present the state-of-the-art in relevant scientific fields. For this research, the literature is split into 4 main sections: 1. General delay propagation research (Section 2.1), 2. Complex network theory (Section 2.2), 3. Analyzing the air transportation system as a complex network (Section 2.3) and 4. Analyzing delay propagation using complex network theory (Section 2.4). Finally, the state-of-the-art is presented in Section 2.5 along with the main conclusion.

### 2.1. Delay propagation

Primary delays in airline schedules are clear, an aircraft can have technical problems, airspace can be congested or weather conditions can limit the capacity of an airport. Initial delays usually lead to more delays, the so-called propagated delay, because of the tight scheduling most airlines use.

One of the first to analyze the effects of delay propagation on an airline's schedule were Boswell and Evans [1997]. This research introduced the concept of a delay multiplier; initial delay multiplied with this factor represents the total delay caused. The mean delay multiplier found was 1.8, so the mean propagated delay is 80% of the initial delay. Also a statistical model is presented which is able to quantify the likelihood of flight cancellation based on the initial arrival delay. Airlines will cancel flights in order to minimize the effect of delay propagation on their schedules.

To manage capacity in the US ATS, the Ground Delay Program (GDP) was introduced [Beatty et al., 1999]. In this program, the US ATC identifies airports that have a capacity problem in the near future. It then attempts to delay aircraft that are still on the ground but have an arrival time at those airports in the constrained period. This program attempts to prevent the unnecessary burning of fuel (i.e., preventing the need for flying holding patterns) but it also introduces delays for individual flights and puts great pressure on the tight schedules of airlines.

Beatty et al. [1999] attempted to quantify the effects of this GDP program on delay propagation, again using the concept of a delay multiplier. These delay multipliers can be used to calculate the total cost of an initial delay, which otherwise would be grossly underestimated. Beatty et al. [1999] only considers the delay which is caused by the late arrival of aircraft and crew. Having access to the database of American Airlines, it was easy to pinpoint the cause of the delay and thus quantify the delay caused by propagation. Delay multipliers were determined based on two variables: the amount of initial delay and the time of day it occurred. The results from this study were used in order to evaluate the changes in the procedures of the GDP program. At that time the savings of the program were clear (less fuel burn) but the total delay caused was unclear.

The first to actually compare the performance of airlines was Baden et al. [2003]. It compared the share

of initial delay to propagated delay for 20 airlines in the USA. The airline with the highest propagated delay was Southwest Airlines and Baden et al. [2003] hypothesize that this is caused by the LCC model, using short turnaround times to maximize the utilization of their aircraft.

Next to comparing airlines, Baden et al. [2003] also looked at the complete NAS to examine the effects of delay propagation. One of the conclusions was that bad weather days increased the share of propagated delay. It also assesses the performance of individual airports, how much arrival delay is propagated to departure delay. These results were linked back to the propagated delay for each airline, as the airports with the highest delay propagation were the hubs of the airlines which had the highest delay propagation.

The next step in analyzing delay propagation was the introduction of propagation trees [AhmadBeygi et al., 2008]. This research compared two airlines which use different business models (one HS and one PP) based on the delay propagation in their networks. In particular, it analyzes how delay propagates through an airline schedule in terms of crew and aircraft. The main conclusions are that combining crew and aircraft in scheduling reduces delay propagation, root delay in the beginning of the day leads to higher propagated delay and the different business models have no clear effect on delay propagation.

To minimize the effects of initial delays, airlines introduce slack times into their schedule. If a flight is delayed this slack time will be able to absorb the delay and prevent it from propagating. The disadvantage of using slack time is the decreased utilization of aircraft/crew in case flights operate on time. AhmadBeygi et al. [2008] proposes that a trade-off should be made between the placement of this slack time as the end of the day usually serves as a break in the schedule. Slack time in the beginning of the day provides more value in terms of delay propagation, but it might also lead to unused slack time and later delays will propagate fully. The optimal location is the middle of the day, making the expected delay minimal by trading off the length of propagation and the probability of the root delay.

Figure 2.1 shows an example of a propagation tree used in the analysis of AhmadBeygi et al. [2008]. Flight 1 has an initial delay of 180 minutes, which causes a total of 430 minutes of propagated delay with a depth of 3 (affecting 3 flights downstream). Clearly the splitting of aircraft and crew causes undesirable delay propagation by creating extra branches in the tree. The maximum depth found in the analysis is six for the HS network and ten for the PP network. The magnitude, propagated delay divided by initial delay, has a maximum value of 6.16 for the PP network and 5.78 for the HS network.

Daniel and Harback [2008] analyzed the congestion of airports using structural models of landing and takeoff queues. The 27 most important US airports were included in this analysis and separates delays into internal (from own airline) or external (from other airlines). Conclusions from this research indicate that any policy introduced to reduce airport congestion should be seen as a system-wide effort. It introduces congestion pricing as a delay mitigation technique and argues that it can even be advantageous for HS airlines if it is introduced at their major hub(s).

Research by Eurocontrol in 2009 has analyzed delay propagation for the specific case of the European network [Jetzki, 2009]. In this research, it was concluded that 50% of the delay in networks of LCC was reactionary, while it was only 40% for the networks of FSC. Another major conclusion was that FSC were able to absorb more delay in the turnaround phase, while the LCC were able to absorb more delay in the flight time. It also concluded that around 50% of the delays can be recovered within one leg after the initial delay.

To assess the individual performance of airports, Welman et al. [2010] calculated delay multipliers for 51 US airports. This research was also able to quantify the economic benefits of expanding airport capacity in order to reduce delay propagation using a cost-benefit analysis.

Kondo [2011] specifically compares a PP network with a HS network. The main hypothesis of his research was that delay propagation is larger for airlines using a PP network when compared to an airline using a HS network. The delay is considered propagated when the aircraft goes through three steps: arriving late at an airport, departing late from that same airport and arriving late at the next airport. Any reduction in delays, reduces the oldest propagated delay first using the first in, first out

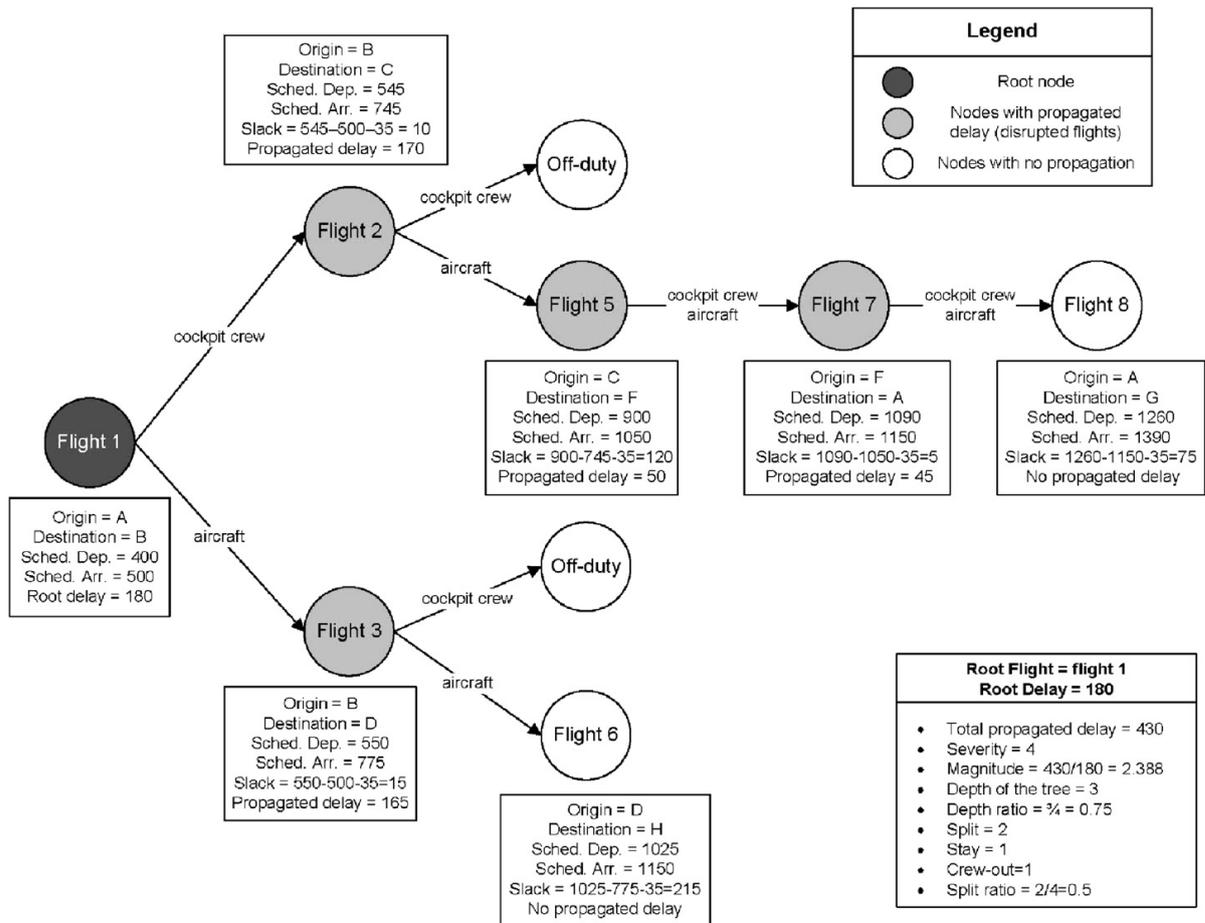


Figure 2.1: A delay propagation tree used by [AhmadBeygi et al., 2008].

logic. The conclusion of Kondo [2011] was that the LCC with a PP network showed more propagated delay than the FSC with HS network.

The next approach to analyzing delay propagation in the network of airports was to use a queuing engine and a delay propagation algorithm [Pyrgiotis et al., 2013]. The model operates by iterating between the two elements, the local delay is calculated by the queuing engine and the delay propagation algorithm uses this to update the flight schedules and demand at all the airports. The main application for this model is the ability to test new policies and strategies in a system of airports.

Until this point research into delay propagation has already presented interesting results. Delay multipliers seem good indicators to quantify the total delay and propagation trees provide a clear visualization of the delay spread. The most recent research into delay propagation is presented in Section 2.4 as it utilizes a complex network theory approach. Before this research is discussed, a general explanation about this methodology is deemed necessary. Section 2.2 will introduce this methodology and then discuss the most recent research on delay propagation.

## 2.2. Complex network theory

In the mathematical field, network theory is defined to be the study of graphs which represent relations between discrete subjects. The Seven Bridges of Königsberg problem (seen in Figure 2.2) was solved by Leonhard Euler and is assumed to be the foundation of modern network theory [Shields, 2012]. The city of Königsberg had seven bridges, and the question was whether one could take a stroll and cross all seven bridges exactly once.

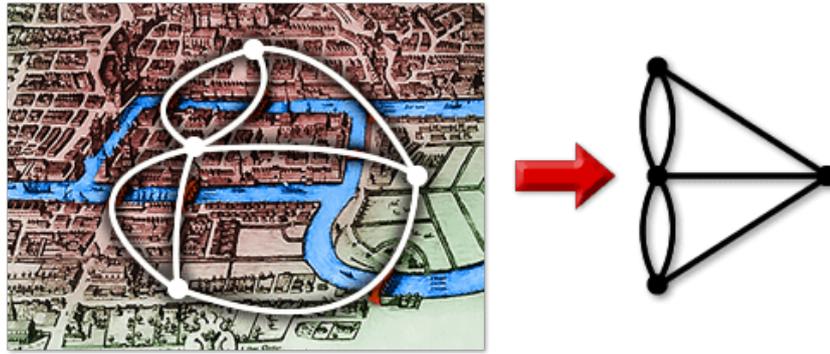


Figure 2.2: The seven bridges of Königsberg.

Euler realized that it did not matter what route you took inside each land mass. This allowed him to model the problem in an abstract way, retaining a list of land masses and the bridges between them. Each land mass is represented by a node and each bridge is a connection between the nodes. The resulting structure is called a graph, shown for the Königsberg problem in Figure 2.2. Euler concluded that this problem could not be solved, but offered some alternative solution methods by adding nodes and edges to the existing situation [Shields, 2012].

These graphs form the basis of complex networks and have been studied extensively since Euler's analysis. In this chapter the most important developments will be discussed. First, some basic metrics are given in Section 2.2.1 to provide a found basis for the use of complex networks. In Section 2.2.2 the basic topological structures which have been defined in recent years are reviewed. Next, the mathematical indicators of the importance of a node within the network are presented in Section 2.2.3 and the methods to assess the performance of complex networks are discussed in Section 2.2.4. The application of complex network theory to modeling the network of the ATS is given in Section 2.3 and to analyzing delay propagation in air transport in Section 2.4.

### 2.2.1. Basic metrics

As can be seen from Euler's representation of the Königsberg problem, a basic graph consists of a set of  $N$  nodes ( $n_1, n_2, \dots, n_N$ ) and a set of  $k$  edges ( $l_1, l_2, \dots, l_k$ ) connecting each node (i.e.  $l_{ij}$  is the edge between node  $i$  and  $j$ ). A graph is either directed or undirected, depending on whether the connection is mutual (i.e. in a directed graph  $l_{ij} \neq l_{ji}$ ). The graph can be modeled using an adjacency matrix  $A$ . This is a square  $N \times N$  matrix in which each element ( $a_{ij}$ ) is equal to 1 if node  $i$  and  $j$  are connected and 0 if they are not. For an undirected graph this matrix is symmetric and for a directed graph it is not. In this section some basic definitions will be defined which are used throughout complex network theory and in the rest of this review [Boccaletti et al., 2006].

The degree of a node ( $k_i$ ) is given in Equation (2.1), and it is equal to the number of edges each node has. In this case it is defined in terms of elements of the adjacency matrix  $A$ . If the graph is directed, an in-degree ( $k_i^{in}$ ) and out-degree ( $k_i^{out}$ ) is defined for each node  $i$ .

$$k_i = \sum_{j \in N} a_{ij} \quad (2.1)$$

Each node within the network has a certain degree and the degree distribution ( $P(k)$ ) describes the probability that a randomly chosen node has a degree  $k$ . Depending on the network structure, this distribution has a different shape. Again, if the graph is directed an ingoing and outgoing probability distribution can be identified.

When looking at the path connecting different nodes within the network, a characteristic path length  $L$  can be identified. Sometimes it can also be referred to as the average shortest path length and it is

shown in Equation (2.2).  $d_{ij}$  is defined as the geodesic distance (or the number of edges) between node  $i$  and  $j$ . The maximum value of  $d_{ij}$  is considered the diameter of the network and represents the maximum number of edges between any two nodes.

$$L = \frac{1}{N(N-1)} \sum_{i,j \in N, i \neq j} d_{ij} \quad (2.2)$$

Finally a clustering coefficient  $C$  can be identified, as shown in Equation (2.3). It consists of the average of the local clustering coefficient of each node, which is further elaborated in Equation (2.4). Clustering is used to measure what is the probability that two nodes  $j$  and  $m$ , which are both connected to node  $i$ , have a direct link.

In Equation (2.4), nodes  $j$  and  $m$  are all possible combinations of neighbors of node  $i$  (denoted as subset  $G$ ). The  $a$  coefficients are from the adjacency matrix and  $k_i$  is the total number of neighbors of node  $i$ .

$$C = \frac{1}{N} \sum_{i \in N} c_i \quad (2.3)$$

$$c_i = \frac{\sum_{j,m \in G} a_{ij} a_{jm} a_{mi}}{k_i(k_i - 1)} \quad (2.4)$$

The previous equations in this section have assumed that the network has homogeneous edges (either present or not). In real-world networks the edges are often characterized by a large variety of intensity/capacity. Barrat et al. [2004] was the first to use this feature in their analysis by introducing the weighted network, where each edge is assigned a proportional weight ( $w_{ij}$ ) depending on its capacity or intensity. Using these weighted edges, a completely different topological structure is uncovered. A weighted degree can be defined for each node, also called the strength  $s_i$  of the node (Equation (2.5)). Much like the degree distribution, a strength distribution ( $R(s)$ ) can also be defined for a network.

$$s_i = \sum_{j \in N} a_{ij} w_{ij} \quad (2.5)$$

In a weighted network, the characteristic path length can also be adapted ( $L^W$ ) to include the weight of the edges [Latora and Marchiori, 2001]. Equation (2.6) presents the equation that is proposed, replacing the geodesic distance with the inverse of the weights. In some literature this is also referred to as the efficiency of the network [Latora and Marchiori, 2003].

$$L^W = \frac{1}{N(N-1)} \sum_{i,j \in N, i \neq j} \frac{1}{w_{ij}} \quad (2.6)$$

Finally, a local weighted clustering coefficient ( $C^W$ ) can be identified, as shown in Equation (2.7). To calculate the total clustering coefficient for the network, the same approach is taken as in Equation (2.3) by taking the average local clustering coefficient. In this equation the weights  $w$  of the edges are included, along the strength  $s_i$ .

$$c_i^W = \frac{1}{s_i(k_i - 1)} \sum_{j,m} \frac{(w_{ij} + w_{im})}{2} a_{ij} a_{jm} a_{mi} \quad (2.7)$$

### 2.2.2. Topological structures

The graphs used in network theory all have a certain topological structure. Different approaches have been investigated to generate networks with the characteristics of real-world networks, the developments of which will be explained in this chapter. The structure of a network has an influence on the characteristics of each node, but also on the dynamics of the entire system.

Regular graphs are the most basic structures. In these graphs each node has the same number of edges [Wilson, 1996], thus the degree of each node is equal. Figure 2.3 shows an example of a regular graph, consisting of 8 nodes which are each connected to its 6 closest neighbors (i.e. degree = 6). These graphs describe real-world networks poorly as the degree of each node is usually very different. Consider a social network, the number of edges each person has will not be the same for each individual.

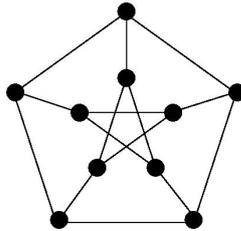


Figure 2.3: An example of a regular graph [Wilson, 1996].

A different network structure introduced in the 1950's was the random graph [Erdős and Rényi, 1959; Gilbert, 1959]. From a mathematical perspective, random graphs most often refer to the Erdős-Rényi (ER) model. In this model, a graph  $G$  is defined with a set of  $n$  nodes and each edge between every pair of nodes is included in the graph with a probability  $p$  (noted as  $G(n,p)$ ).

In Figure 2.4 it can be seen what the ER-model looks like for different choices of  $p$ . Intuitively, increasing  $p$  increases the overall number of edges and the connectivity of each node. Choosing  $p = 1$  would lead to a graph where every node is connected to every other node.

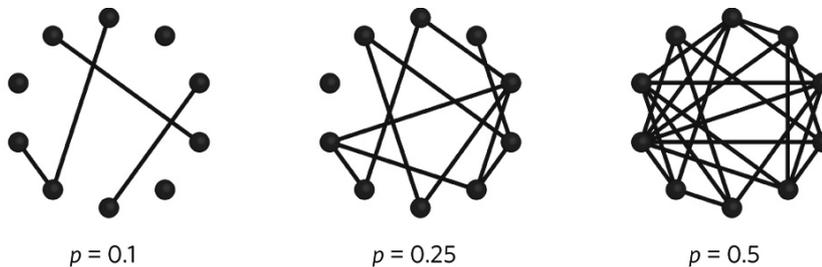


Figure 2.4: Random graphs for different  $p$  [Perseguers et al., 2010].

The main drawback of the ER-model is the inability to reconstruct certain characteristics of real-world networks (i.e. social networks, the Internet, biological networks). In particular, the lack of clustering in the network and an unrealistic degree distribution [Newman, 2002]. In the ER-model, the degree distribution is shaped by a Poisson distribution while many real-world networks show a power-law distribution.

One of the more recent models that approaches real-world networks more accurately are the small-world networks [Watts and Strogatz, 1998]. These networks combine the clustering features of regular graphs with the small characteristic path length of random networks. The small-world networks are named after the small-world phenomenon as described by Stanley Milgram [Travers and Milgram, 1969], where he states that the human society is a network of people with a short characteristic path length. Linking two randomly chosen, geographically separated people required only 6 edges (median number), also referred to as the 'six degrees of separation'.

Small-world networks (SWN) are usually constructed using the Watts-Strogatz (WS) model. It starts

with a regular graph consisting of  $n$  nodes and  $k$  edges connecting each node to its nearest neighbors. From this graph, each edge is rewired with a probability  $p$  to a random other node. Figure 2.5 shows the position of small-world networks, compared to regular graphs ( $p = 0$ ) and random graphs ( $p = 1$ ). Before this research, most of the work was done on networks which were either completely regular or completely random. This new approach allowed research to be done on networks that fall in between these two extremes, which is the case for many biological, technological and social networks [Watts and Strogatz, 1998].

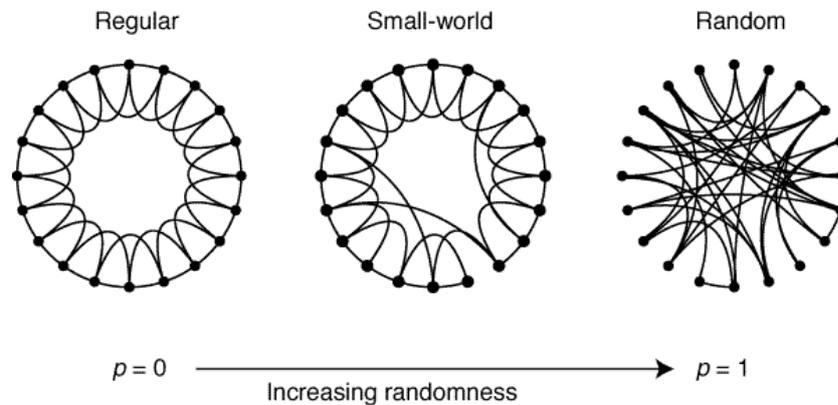


Figure 2.5: Small-world network [Watts and Strogatz, 1998].

Around the same time as the WS model, the scale-free (SF) network was introduced [Barabasi and Albert, 1999]. The main characteristic of the SF network is that the degree distribution of the nodes is shaped by a power-law distribution (i.e. majority of the nodes has a few edges and a few nodes are highly connected) instead of a Poisson distribution. This power-law distribution is usually characterized with the coefficient  $\gamma$  such that  $P(k) \approx k^{-\gamma}$ . When investigating real-world networks such as the Internet or the nervous system, the majority shows this behavior. Two explanations are given for this characteristic, namely that due to the addition of new nodes the networks continuously expand and secondly the new nodes have a preference to attach to nodes which are already highly connected.

The final type of network that will be discussed in this section are spatial networks. In these networks the nodes are located at a precise position in 2 or 3-dimensional space and the edges are real physical connections [Boccaletti et al., 2006]. Well-known examples include neural networks, where cells have a location and physical connections exist between them, or the power-grid network with physical power lines connecting power stations to individual houses. The degree of a node in these networks is limited, as the physical space to connect them is finite (i.e., only a certain number of streets can cross an intersection [Crucitti et al., 2006]).

After the development of these networks structures, most of the real-world networks have been shown to be a combination of small-world, scale-free and weighted networks. Examples are the Internet [Boccaletti et al., 2006], biological systems [Albert and Barabasi, 2002], transportation networks [Wu et al., 2006], the power grid [Kim et al., 2016] and social interaction systems [Barrat et al., 2004].

### 2.2.3. Centrality measures

Since the different structures have been defined, there has been much research into mathematical measures which can be used to assess the position of different nodes within the network. These are called centrality measures and are used to identify which nodes are most important within a graph.

The most simple centrality measure is the degree of each node [Sabidussi, 1966]. The higher the number of edges connected to a node, the more important that node is considered to be in the network. In a directed graph, a distinction can be made between in-degree and out-degree, ranking the nodes according to those respective characteristics.

Another centrality measure that has been extensively used is betweenness and the first use of it is

most often attributed to Linton Freeman [Freeman, 1977]. Equation (2.8) shows the methodology used to calculate the betweenness  $b_i$  of node  $i$ . In this equation  $n_{jk}$  is the number of shortest paths from node  $j$  to node  $k$ , and  $n_{jk}(i)$  is the number of shortest paths from node  $j$  to node  $k$  passing through node  $i$ . As each node has its own betweenness coefficient, a distribution can be determined for the entire network.

$$b_i = \sum_{j,k \in N, j \neq k} \frac{n_{jk}(i)}{n_{jk}} \quad (2.8)$$

The adjacency matrix  $A$  can also be used as a centrality measure [Bonacich, 1987]. In this method called eigenvector centrality, the idea is that relative scores are assigned to each node based on the importance of the nodes it is connected to (as shown in Equation (2.9)). That equation can be rewritten as  $A\mathbf{x} = \lambda\mathbf{x}$ , where there will be many values for  $\lambda$  for which a non-zero solution will exist. However, due to additional constraints only the greatest eigenvalue will yield the desired centrality measure. The  $i$ th component of the eigenvector  $\mathbf{x}$  will provide the relative centrality score of node  $i$ .

$$x_i = \frac{1}{\lambda} \sum_{j \in N} a_{ij} x_j \quad (2.9)$$

Closeness centrality was first proposed as a concept by Alex Bavelas [Bavelas, 1950]. The notion is that a node is central if it is close to all of the nodes in the network. Equation (2.10) shows the methodology to calculate the closeness of a node, using the inverse of the sum of the shortest path to each node in the network.

$$closeness_i = \frac{1}{\sum_{j \in N} d_{ij}} \quad (2.10)$$

In 1983, Stephan Seidman introduced the concept of identifying different layers in a network based on an integer index  $k_s$  [Seidman, 1983]. Each node is assigned a score based on its connectivity patterns, where a high value of  $k_s$  indicates the node is in the center of the network and a low value of  $k_s$  means the node is in the periphery. The network can then be divided into different layers based on the  $k_s$  value of each node, the so-called k-shells. An example of the resulting structure is shown in Figure 2.6.

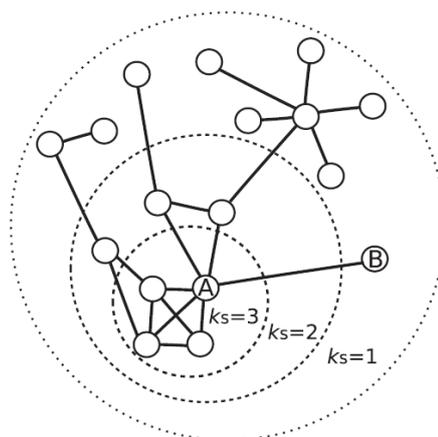


Figure 2.6: A complex network divided into k-shells [Garas et al., 2012].

The process to execute the k-shell decomposition method (as it is most often referred to) is as follows. It starts by removing all nodes from the network that have degree  $k_i = 1$  and these are placed in  $k_s =$

1. Keep iterating until there are only nodes left with degree  $k_i \geq 2$ . Repeat the process by deleting all nodes with degree  $k_i = 2$  and placing these in  $k_s = 2$ . These steps are done until all nodes in the network are assigned to a k-shell.

The main drawback of all these centrality measures is that they do not work well on weighted networks. This is why an alternative to the k-shell decomposition method was suggested, which includes the weights of the edges [Garas et al., 2012]. A weighted degree  $k'_i$  is calculated for each node, as defined in Equation (2.11). It is the product of the degree of the node and the sum of the weights of all its edges. The parameters  $\alpha$  and  $\beta$  can be tweaked to make either element more important, but it has only been investigated for cases where they are equally important (i.e.  $\alpha = \beta = 1$ ) [Garas et al., 2012].

Lately, research has also been done on decreasing the computational power which is needed to calculate the centrality measures. Considering a social network with millions of nodes, all the aforementioned measures are computationally very expensive. To counter this, local centrality measures were introduced which would prove to be an accurate representation of the entire network. Chen et al. [2012] introduced a semi-local centrality measure which is based on global measures such as betweenness and closeness. It also proves that this measure is an accurate identifier of influential nodes in a network. Another study uses the local clustering coefficient to identify what the most influential nodes are in the network [Chen et al., 2013]. This is compared to the most important global centrality measures to prove the performance in representing the global network.

$$k'_i = \left[ k_i^\alpha \left( \sum_j^{k_i} w_{ij} \right)^\beta \right]^{\frac{1}{\alpha+\beta}} \quad (2.11)$$

#### 2.2.4. Analysis of complex networks

Once complex networks have been established and their topology has been identified, they are usually analyzed using 2 important metrics. These 2 metrics are discussed in this section, namely the robustness and the spreading processes.

**Robustness** Complex networks are often used to assess the robustness of a network. In this case, it refers to the ability of a network to cope with disturbances and complete failures.

The static robustness of a network means that nodes will be deleted without redistributing the quantity being transported in the network. For example, in social networks relationships can be cut between individuals but that does not automatically lead to new relationships. Then there is also the dynamic robustness of a system, where the removed flows are diverted through other paths. An example of this can be the failure of a router, which will lead to the diversion of loads to other routers. The main difference is that static robustness can be measured analytically while the dynamic robustness requires numerical simulations [Boccaletti et al., 2006].

Percolation theory is usually the basis of static robustness analysis and it has been applied to many real-world networks such as the Internet [Cohen et al., 2000]. This example contains an analysis of random breakdowns (randomly deleting a fraction  $f$  of nodes) in the Internet and it tries to identify a critical point ( $f_c$ ) where the entire networks fails, also referred to as the percolation threshold of the network. However, this theory can not be used to analyze attacks on the networks, situations where nodes are not chosen at random but rather by their importance within the network. When looking at random networks (such as the ER-model), a threshold value can be identified for both the random removal of nodes and attacks.

Looking at the Internet example, different conclusions can be made about removing random nodes or attacking specific nodes based on their characteristics (i.e. their degree) [Albert et al., 2000]. The conclusion is that scale-free networks are highly robust against the random removal of nodes but highly

vulnerable when the most important nodes are removed. The percolation threshold value is absent for the case of random failures in scale-free networks [Cohen et al., 2000].

Of course, when removing nodes from a network this usually leads to dynamic effects. The main assumption is that the removal of those nodes leads to the overloading of others. This is most often analyzed using cascading failures in nodes or the congestion in communication systems [Boccaletti et al., 2006].

The example that is most used for describing dynamic robustness is the power grid. Sachtjen et al. [2000] analyzes the initial disturbance on August 14, 2003 which caused the largest blackout in the history of the US. In this research, random disturbances to the power grid network are analyzed. Each node has a maximum load it can handle and will fail when overloaded. An initial failure at a node causes neighboring nodes to overload and thus an avalanche of failures has been created.

**Spreading processes** Complex networks can also be used to analyze how something spreads across a network. The two main theories for this are epidemic spreading and rumor spreading. These two processes are fundamentally different, as in epidemic spreading the dynamics are fixed while in rumor spreading one can design the dynamics in such a way that the spreading is optimized [Boccaletti et al., 2006].

Modeling the spreading of disease in complex networks is usually done using an epidemic model. Based on the variant of the epidemic model used the nodes are in a discrete state, most often either susceptible (S), infected (I) or recovered (R). The two variants that are most often used are the SIS and SIR models, illustrated in Figure 2.7. When an individual comes into contact with an infected individual, it is infected with a certain rate  $\beta$ . Once infected, the individuals recover with a certain rate  $\delta$  and either return to the susceptible state (SIS) or are permanently recovered (SIR).

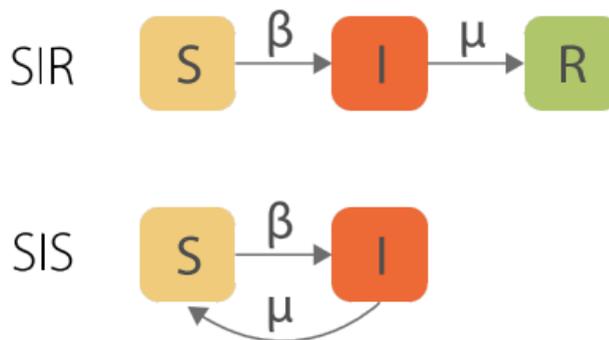


Figure 2.7: Epidemic models (SIS vs SIR) with susceptible (S), infected (I) and recovered (R) states, along with infection rates  $\beta$  and recovery rate  $\mu$  [Pastor-Satorras and Vespignani, 2001].

Epidemic models were first applied to complex networks (in particular scale-free networks) by Pastor-Satorras and Vespignani [2001]. In this research a dynamical model was introduced which defines the spreading of infections, using computer virus infections as an example. The average lifetime of the virus, the persistence of this virus, the absence of an epidemic threshold and the critical behavior are all analyzed. The conclusions from this research can be generalized to all scale-free networks with  $2 < \gamma \leq 3$ , with the degree distribution  $P(k) \approx k^{-\gamma}$ .

In 2008, the use of metapopulation models was connected with complex network theory, namely characterized by heterogeneous connectivity patterns [Colizza and Vespignani, 2008]. In a metapopulation model each node consists of a population of individuals. Combining this with an epidemic model where each individual can infect others leads to an interesting representation of cities (nodes) and their populations. These individuals can travel from one city to another, spreading the infection across a network. In this research, it was concluded that the coupling between populations has an effect on the ability of the infection to spread. It also defines a local and global epidemic threshold for an epidemic outbreak.

More recently, research has been done in identifying the most influential spreaders in an epidemic model [Kitsak et al., 2010]. In this research, the k-shell decomposition method was used in order to divide the network into layers (core and periphery). The conclusion was that the k-shells formed a good indicator in the spreading efficiency of a node. Identifying these nodes forms an important step in forming successful disease control strategies.

Two main reviews exist, which summarize the main findings thus far on the epidemic model in complex networks [Nowzari et al., 2016; Pastor-Satorras et al., 2015]. These papers present a comprehensive summary of the SIR/SIS model on complex networks, identification of influential nodes, strategies to maximize or prevent spreading and the use of metapopulation models.

Rumor spreading was first analyzed by Daley and Kendall [1965] and consists of a modified SIR model. The main difference with the epidemic model is that the rumor is propagated as long as it is new to the receiver. Once the receiver comes into contact with the rumor, the one spreading the rumor will lose interest in spreading it. Applying rumor spreading to complex topology leads to the conclusion that the heterogeneous characteristics block the dynamics of rumor spreading [Moreno et al., 2004].

## 2.3. ATS network analysis

To analyze the performance of the ATS, complex network theory has been extensively used. Recent developments by Watts and Strogatz [1998] and Barabasi and Albert [1999] have allowed a more accurate representation of real-world networks. This has led to the use of the ATS as an example of a complex network, from which conclusions have been drawn on the performance of this system. Table 2.1 presents an overview of the research performed on the ATS using complex networks, giving the year, authors, network scope, number of nodes, number of edges, the degree distribution power-law coefficient  $\gamma$  (more than one value indicates different power law approximations were used for different parts of the curve), the characteristic path length  $L$  and the clustering coefficient  $C$  found in each research. The scope can either be given geographically or it might only include certain airlines/alliances (refer to the list of acronyms).

In 2004, Barrat et al. [2004] was one of the first to use the worldwide ATS as an example of a complex network, at the same time introducing the concept of weighted complex networks. It models the airports as being nodes in the network and the edges are direct flights which connect the airports. The edges are weighted according to number of available seats in 2002. Barrat et al. [2004] was also the first to prove that the topology structure of the ATS shows scale-free and small-world properties. The characteristic path length  $L$  is shown to be 4.37 and the degree distribution  $P(k)$  has the form of a power-law distribution  $\approx k^{-\gamma}$ , with  $\gamma \approx 2.0$ .

Barrat et al. [2004] also introduces the concept of strength, as explained in Equation (2.5). Analyzing the strength of each airport led to the conclusion that the strength of a node grows faster than its degree. The weight of an edge connected to a node with a high degree tends to be higher than the scenario where weights are assigned randomly. This proves the correlation between the topological structure and the weight, airports which are larger also handle more traffic.

Further research used complex network specifically to analyze the ATS, instead of just using it as an example. Guimerà et al. [2005] analyzes the worldwide ATS, also proving the small-world and scale-free properties of the system. This research identifies communities within the network, which affects what airports are the most central in the network. It is shown that the most connected cities in the network are located in Europe and North America, while the most centrally located cities are located all around the world (illustrated in Figure 2.8). This leads to the conclusion that the nodes with the highest degree are not necessarily the nodes with the highest betweenness centrality. The opposite has been shown by Goh et al. [2003] to be the case when analyzing another SF complex network, the Internet.

Guimerà et al. [2005] also presents an explanation for this finding. Consider a region such as Alaska, with many smaller airports in the rural regions. All of these smaller airports are (directly or indirectly) connected to the largest airport of Alaska, Anchorage. This airport in turn, is one of the only airports in Alaska which is connected to the US mainland and international destinations. For this reason the

Table 2.1: Research modeling the ATS as a complex network

Year	Author	Scope	Nodes	Edges	$\gamma$	$L$	$C$
2003	Li and Cai	China	128	1,165	0.428, 4.161	2.067	0.733
2004	Barrat et al.	World	3,880	18,810	2.0	4.37	-
2005	Guimerà et al.	World	3,883	27,051	2.0	4.4	0.62
2007	Guida and Maria	Italy	42	310	0.2, 1.7	1.98	0.10
2008	Bagler	India	79	228	2.2	2.493	0.6574
2009	Reggiani et al.	LH Europe	111	522	2.01	-	-
		LH World	188	692	2.22	-	-
		SA Europe	111	3,230	2.49	-	-
		SA World	188	6,084	2.54	-	-
	Han et al.	OS	130	721	2.428	2.31	0.122
		BA	327	2152	2.894	2.66	0.116
		AF, KL	353	3835	1.856	2.20	0.376
		LH	430	4643	1.755	2.9	0.376
2010	Zhang et al.	China	150	-	0.49, 2.63	2.27	0.79
	Liu et al.	China	121	1,378	0.47, 2.07	2.263	0.748
2011	Wang et al.	China	144	1,018	-	2.23	0.69
2012	Cheung and Gunes	USA	850	6,478	1.05	3.24	0.62
2014	Jia et al. Lordan [2014a]	USA	$\approx 1,080$	$\approx 24,500$	-	3.02	0.52
		SA	1,150	4,240	-	3.24	0.77
		SkyTeam	896	3,226	-	3.13	0.74
		OneWorld	741	1,670	-	3.28	0.71
		LH	209	395	-	2.1	0.93
		UA	362	933	-	2.57	0.91
		US	203	408	-	2.26	0.96
		AB	119	361	-	2.31	0.51
		AA	272	523	-	2.3	0.94
		BA	186	223	-	2.87	0.15
		AF	178	258	-	2.42	0.46
		MU	182	571	-	2.5	0.55
		CZ	178	576	-	2.45	0.62
		DL	328	882	-	2.38	0.88
		FR	178	1,396	-	2.16	0.44
U2	131	601	-	2.19	0.39		
WN	86	507	-	1.97	0.72		

betweenness of Anchorage airport is relatively high compared to its absolute degree. Each community within the global network is shown to have such an airport and these are the ones with the highest betweenness centrality.

Next to the analysis of the worldwide ATS, research has been done on specific portions of the network. Taking a sample of the worldwide network has been done in two ways, by either taking a geographical

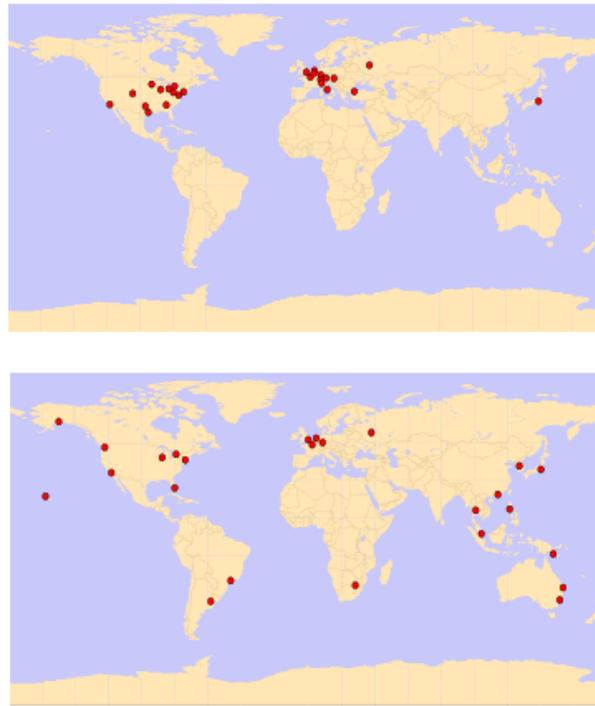


Figure 2.8: The worldwide ATS, cities with highest degree vs. highest betweenness centrality [Guimerà et al., 2005].

portion or only including specific airlines/alliances.

Looking at the geographical scope, Li and Cai [2003], Zhang et al. [2010], Liu et al. [2010] and Wang et al. [2011] have analyzed the Chinese airport network. The results of these studies, presented in Table 2.1, are not homogeneous. This might be due to a variety of factors related to the input data, for instance the temporal scope or the amount of airports included. The number of nodes and edges included in the network are similar, but not exactly the same (as shown in Table 2.1). Further regional studies have focused on Italy [Guida and Maria, 2007], India [Bagler, 2008] and the USA [Cheung and Gunes, 2012; Jia et al., 2014].

Other studies have focused on specific airlines and their alliances. Reggiani et al. [2009] has done a case study on the network of Lufthansa (LH), the biggest FSC of Germany, and Star Alliance (of which LH is a part of). It analyzes the network from four different aspects, the European network of Lufthansa, the worldwide network of Lufthansa, the European network of Star Alliance and the worldwide network of Star Alliance. It proves that all of these networks can be mapped on the SF model and it analyzes the role of the hubs in these networks (namely Munich, Frankfurt and Düsseldorf).

Han et al. [2009] has made a comparison between four different FSC in Europe. Concluding from this research, the networks of Austrian Airlines (OS) and British Airways (BA) are similar to each other and the networks of Air France - KLM (AF, KL) and LH are similar. The explanation for this is that the flight density for these airlines are roughly the same, but also that the flight distributions to smaller nodes are similar. The biggest difference between these two groups of airlines is that BA and OS operate a HS network, while AF-KL and LH operate a MHS network. This resulting difference in the network structure is shown by the clustering coefficient  $C$  and the power-law coefficient  $\gamma$  in Table 2.1.

To specifically compare FSCs with LCCs, research by Lordan [2014b] tried to specify a measure which determined the difference between these two networks. The conclusion was that the use of betweenness centrality was the most useful measure in determining the network configuration. A relatively high betweenness centrality was shown to correlate with a HS network. The power-law coefficient  $\gamma$  was shown to have little relation to the type of network configuration.

One main conclusion that can be made from Table 2.1 is that the ATS can be considered a SWN. The characteristic path length  $L$  can be seen to be relatively small to the amount of nodes included in the network. Another conclusion is that most networks can be considered SF as the power-law coefficient usually satisfies  $2 < \gamma \leq 3$ . When two power-law approximations are made it is harder to identify the overall degree distribution.

The clustering coefficient on the other hand can fluctuate depending on the scope of the network. Certain airlines such as BA who operate a HS network, are shown to have a clustering coefficient around 0.15. Airlines that operate a MHS network (this also includes the alliances) have varying clustering coefficients, but usually higher than 0.5. Finally LCC operating a PP network have varying clustering coefficients between 0.39 and 0.72. For the overall airport network (local and global) the clustering coefficient lies between 0.6 and 0.8.

To analyze the spreading of epidemic diseases around the world, Colizza et al. [2007] modeled the air transportation network in order to capture its role in this process. By comparing the spread of a disease to the structure of the ATS striking resemblances were discovered. In this research it is proven that the ATS is an important tool to spread diseases, but can thus also be used to effectively contain diseases.

A complete review of the static and dynamic analysis of the ATS as a complex network is given by Zanin and Lillo [2013]. It combines the results of studies on the topology, dynamics and resilience of the ATS. Figure 2.9 shows the networks set up for a LCC and a FSC, illustrating the difference between PP networks and HS networks.

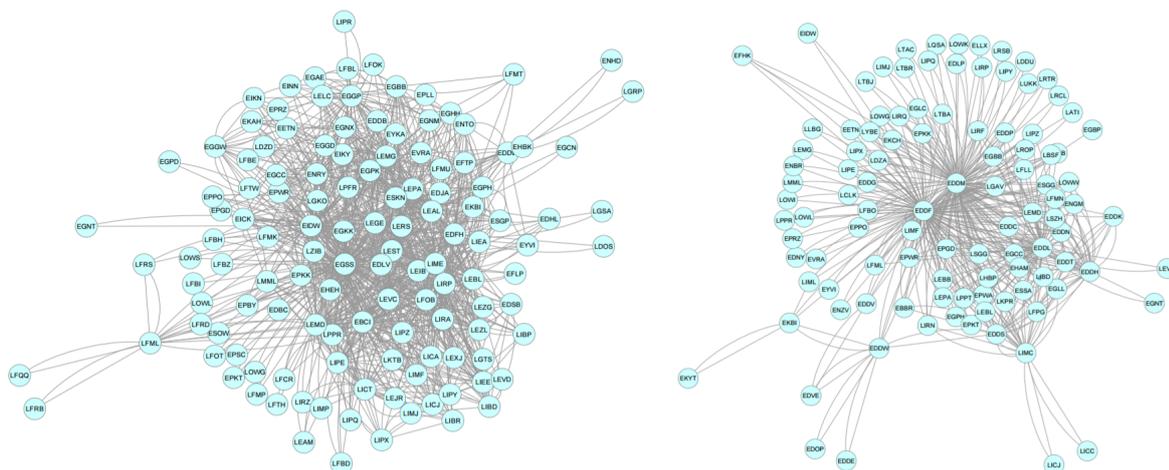


Figure 2.9: PP Network of Ryanair (left) vs MHS network of Lufthansa (right) [Zanin and Lillo, 2013].

More recently, studies have been performed into the layers within the ATS. Cardillo et al. [2013] analyzed the European ATS. Each airline is considered to be a layer in the total network and this study analyzes the effects of merging these individual layers. Next to analyzing individual airlines, the comparison between networks of FSC and LCC is made. From these results the hub-and-spoke structure of FSC can be identified due to the hubs having higher connectivity. Also the merging of FSC networks and LCC networks shows different behavior. Most notably, the clustering coefficient  $C$  saturates after merging 5 FSC layers while it keeps increasing when merging LCC layers. This can be explained as the merging of LCC networks leads to the continuous formation of new triangles, while that is not the case for FSC networks.

Verma et al. [2014] uses the k-core decomposition method to identify the layers within the worldwide ATS. It separates the airport network into a core layer, a bridge layer and a periphery layer. The most notable conclusion from this research is that the removal of airports from the core layer does not notably affect the connectivity of the network, while removal of highly connected airports from the periphery leads to a drop in resilience of the network. The same analysis was done on the Chinese ATS by Du et al. [2016]. In this research, removing the core layer led to the disintegration of the Chinese ATS and

the formation of many smaller clusters. This is attributed to the fact that the core of the Chinese ATS was relatively much larger compared to the core layer of the worldwide ATS.

Due to the heterogeneous results in Table 2.1, the way the network is sampled should be carefully examined. Stumpf et al. [2005] argued that by taking a subset of a SF network, it is very hard to reproduce the static and dynamic characteristics of the original network. This was studied by Belkoura et al. [2016] specifically for the case of the ATS, which considered eight air networks and several different datasets. Only one exception was found to show an acceptable error, which was the USA RITA database. This includes only 84 % of all airports in the USA, but because of the way this data is collected it shows an accurate representation of the complete US network. The best way of sampling the airport network is to take the subset of the most important airlines, ideally substituting some with smaller airlines.

## 2.4. Delay propagation using complex networks

Next to representing the ATS network with complex network theory, analyzing the delay propagation using this approach has recently gained attention. In Section 2.1 the first research into delay propagation was already discussed. In this section the most recent research using complex network theory is presented.

In 2013, the first research was done into using complex networks to analyze the delay propagation [Fleurquin et al., 2013]. This study analyzed the US airport network and identifies the cluster size of airports which are experiencing delays, as a function of the initial delay and fraction of initially delayed flights. It analyzes delay propagation due to plane rotation, passenger connectivity and airport congestion. The first is always included in the model and the latter two can be modulated at will to investigate the effects.

The model defined by Fleurquin et al. [2013] has two free parameters,  $\alpha$  controls the passenger connectivity at an airport and  $\beta$  accounts for the airport capacity. The passengers of an incoming flight have a certain probability of connecting to another flight, proportional with a factor  $\alpha$  to the connection levels provided by the Bureau of Transportation Statistics (BTS). The airport capacity is based on the actual scheduled arrival rate with the proportionality factor  $\beta$ . It assumes the initial delay is equal to the first delay each individual aircraft encounters. The conclusion is that the cluster size shows high variability from one day to the next, as the system is 'reset' at the end of every day.

To predict delays in the NAS, Rebollo and Balakrishnan [2014] used a temporal and spatial complex network approach. The main objective of this research was to predict a departure delay on a certain link or at a certain airport. It assesses two different prediction mechanisms, a binary output whether the delay is more or less than a threshold value and a continuous output of the estimated departure delay. Again the different clusters of congested airports are analyzed, showing the main delay centers to be Atlanta, Chicago and New York.

The first attempts to simulate the delay propagation in an airline network using complex networks was done by Ciruelos et al. [2015]. The European ATS is modeled as a complex network and an agent-based approach is utilized to simulate the process of delay propagation. Once an individual flight is delayed, a new pair of slots (one at the origin airport and one at the destination airport) is assigned to that flight given it is within a certain threshold. If this pair of slots cannot be found due to capacity constraints, then the flight is canceled. Results from this algorithm show promising similarities with the real patterns observed in the European ATS.

Studies have also compared the performance of the US network versus the European network [Campanelli et al., 2016]. This research uses two agent-based models to simulate the delay propagation and to analyze what the effects of disruptions are on the networks. As flights are delayed, they arrive at the airport when another flight was supposed to land. In this situation it is important to note that a difference exists between the US network compared to the European network. The US ATS uses a first come, first serve basis while in Europe the flight that arrives at its original slot has the priority. Campanelli et al. [2016] concludes that using the US method leads to larger delay propagation compared to the

European method.

The most recent research uses a complex network approach to calibrate an epidemic model which simulates the delay propagation in the European ATS [Baspinar and Koyuncu, 2016]. A metapopulation approach is used to represent the network where each node is an airport and consists of a population of individual flights. This allows the problem to be approached from 2 different perspectives, an airport-based approach and an individual flight approach. Baspinar and Koyuncu [2016] utilize a SIS epidemic model to simulate the delay propagation, as it argues that it is impossible for individual flights to permanently recover (i.e. once a flight absorbs the delay it is again susceptible to becoming delayed). It analyzes the recovery rates for each airport and the individual flights. Negative recovery rates for individual flights indicate congestion in en-route airspace or destination airport. Airports with a negative recovery rate are investigated and the conclusion is that they are utilized close to their maximum capacity.

Using complex network theory to analyze and simulate delay propagation is an active field of research. When looking at the networks created by this research, the common factor is that airports are represented by nodes and direct flights represent the edges. Only recently have separate flights been included by Baspinar and Koyuncu [2016], but only as individuals traveling between nodes. Opportunities for future research lie in including individual aircraft in the analysis.

## 2.5. State-of-the-art

This chapter has provided an in-depth analysis of complex network theory. It has also shown that this theory has been used to represent the network of the ATS and has been used to analyze delay propagation within the ATS. In this section the main conclusions from this research are summarized.

### Network analysis

There has been a lot of research into air transportation as it is an important industry with global effects. Since the deregulation the market has become enormously competitive, with two different business models (FSC and LCC) leading the industry.

Extensive analysis has been done into the performance of the ATS using complex network theory. First of all the topology is widely researched, with evidence showing that these networks have small-world and scale-free characteristics. Research has been done into different subsets of the network, geographically and for specific airlines/alliances. Recently it has been proven that using a subset of the network makes it very hard to capture the effects of the entire network. Also a lot of research has been done into the different layers of the network, created by airlines but also by the density of connections.

The ATS has been an active field of research for years, specifically using complex network theory for the last 10 years. All this research has solved many problems and forms a strong foundation for further research. The main gap that has been identified in this field is the fact that the network representation is always the same, the airports are modeled as nodes and direct flights are edges connecting them.

### Delay propagation

Delays are a problem for the air transportation industry. In the case of airlines it is both financially and for their brand reputation towards the passengers. The problem is even bigger than initial delays, as this delay is propagated throughout the tight schedule of an airline. A lot of research has been done into this topic, where the most recent papers use a complex network theory approach.

In 1999 the need arose to analyze the effects of the GDP, in particular to quantify the amount of propagated delay. The first research into delay propagation was aimed at the ability to translate initial delays into propagated delays. The main approaches were the use of delay multipliers and propagation trees, usually focused on individual airlines.

Since 2013 the use of complex network theory has been used in analyzing the delay propagation. A

variety of different approaches has been used, from analyzing the clustering coefficient to calibrating an epidemic model. Also some research focuses on simulating the delay propagation using various queuing algorithms and agent-based approaches.

The main gap that can be identified is the application of complex network theory to the delay propagation of individual flights. In all of the approaches the airports are modeled as nodes and individual members traveling between these nodes. Also the patterns are identified in the network, but the impact of delays on different days is not captured. Further research could potentially identify a threshold value for which the delay spreads across the network (like an epidemic) instead of being absorbed by the network.



# 3

## Research framework

To build a foundation for research, a framework has to be established within which the research is performed. The first thing is to identify the problem which will be solved, this is done in Section 3.1. The research objective, which is presented in Section 3.2, will present the expected achievement of the research. This is accompanied with a set of research questions and hypotheses. Next the model requirements are presented Section 3.3, which explains the expectations for the model and the software used to create it. Each research also needs a scope, which is described in Section 3.4. This scope defines the boundaries of the research, which are needed to keep the research goals realistic. Finally the contribution of this research is summarized in Section 3.5.

### 3.1. Problem statement

The air transportation industry has several distinct characteristics which make it unique in its operation. It is a global industry which experiences continuous growth and a cyclical nature due to the economic dependence and seasonal demand. The airlines operating in this industry also experience marginal profitability, high dependence on the fuel price and expensive capital. To finish it off the market is highly regulated to meet economic, environmental and safety standards.

These characteristics force airlines to increase the efficiency of their resources in order to sustain their profitability. One of the key problems with this increased efficiency is the vulnerability to disruptions. Delays will propagate throughout the network and impact the entire network. Three main stakeholders can be identified:

1. Airlines: delays lead to disruptions the flight schedules of airlines. As a result of these disruptions the operating time (and thus cost) is increased for key resources such as aircraft and crew. Also, passengers connections to other flights will be impacted, which again increases the cost for airlines.
2. Airports: as the capacity of airports is limited, delays will lead to extra congestion. This congestion will have an effect on the available resources at an airport such as runways, gates and ground handling. Also noise and emissions will be concentrated around congested periods, which affects the areas surrounding the airport.
3. ATC providers: the congestion at airports will also lead to congested airspace, increasing the workload for air traffic controllers. Future ATM modernization programs such as SESAR and NextGen will also be impacted by delays.

## 3.2. Research objective

The goal of this research is to fill the research gaps mentioned in Chapter 2. The proposed research objective is as follows:

**To analyze the delay propagation in the top level of the US air transportation system by combining a complex network theory and an epidemic model approach.**

By representing the aircraft as nodes in the system and defining certain contact points, a completely novel approach to model the network is presented. Applying an epidemic model to the delay in this network will be novel research as well. A few research questions are identified which will be answered throughout the research. Along with these questions, a hypothesis can be made for each about the expected outcome.

- Q1 General: Is the probability of an aircraft being delayed influenced when it comes into contact with other delayed aircraft? Yes this probability will increase.
- Q2 General: Can delay propagation be found every day in the ATS? The amount of delay on a specific day will have an effect on the magnitude of delay propagation. The hypothesis is that this probability will increase.
- Q3 General: Is the spreading of delays in the ATS comparable to the spreading of social behavior? The hypothesis is that these are comparable.

The first hypothesis H1 is made as it will provide important information at a very top-level. The general assumption is that delayed aircraft propagate their own delay throughout the network. This means that coming into contact with delayed aircraft will increase the probability of an aircraft becoming delayed. H2 is made because the assumption is that delay propagation effects will vary on different days. For example, a day with little delay will have less delay propagation than a day with a high amount of delay. The hypothesis could almost be that a threshold value exists for which certain effects can be found. The third hypothesis H3 is made as the behavior spreading through social networks can be compared to the delays in the ATS. Similarities between the spreading in these two systems are assumed to be found because the network topology has been proven to be similar (small-world, scale-free networks). The first step will be to verify that the networks actually show these characteristics and then compare the spreading behavior.

## 3.3. Model requirements

To achieve the research objective, a model will be built which will help achieve the research goals. In this section a few model requirements are explained and the software used will be discussed. The model requirements needed are as follows:

- The input for the model will be provided in Excel.
- The network of aircraft needs to be set up according to the methodology discussed in the next chapter.
- Delays should be analyzed for each individual aircraft.
- All flights included in the model should be handled within reasonable computational time.
- To be able to substantiate the results of the analysis, statistical support should be provided.
- The output of the data should be concisely presented in graphs and/or tables. Raw output data is needed to validate results.

These requirements will be translated to a software model, which will be built in Python. The biggest advantage of Python is that it is an open-source program which is free to use (including all add-ons).

It is heavily inspired by MATLAB, a software package requiring expensive licenses. The version used for this research is Python 2.7, downloaded through one of the most common distributions Python(x,y). As this is a relatively large project, Spyder will be used as the Interactive Development Environment which offers many extra features compared to the standard environment.

Improving the computational time of this model will be outside the scope of this research. The computer used to build and test the model is a HP Z-book Studio G3 with an Intel Core i7-6700 HQ processor and 8GB of RAM.

### **3.4. Research scope**

The scope of the research will provide delimitations within which the research is performed. This will keep the research goals realistic in terms of time and resources.

Geographically this research will be limited to domestic, passenger flights within the United States. No international flights (including Canada en Mexico) are included in this research. Overseas US territories such as American Samoa and Guam are included.

The temporal scope for the research will span over a set of individual days, as delays are usually absorbed at the end of each day. Data from other days can be used to validate conclusions made from one day of flights. Data is used from the month of August 2016, which was the last available at the start of this research. To validate the results, data from other months is used in order to prove the conclusions made in this report can be made for a general situation. These extra months are December 2016 and April 2017.

This research sticks to analyzing the delay propagation. Optimizing the schedules in terms of avoiding delay propagation is considered to be outside the scope of this research. Also predicting or simulating delay patterns is deemed outside the scope of this research.

### **3.5. Contribution**

To find a gap in the scientific body of knowledge, a literature review was performed in Chapter 2. Here the state-of-the-art research was presented and certain gaps were identified. Figure 3.1 presents an overview of the state-of-the-art and potential contribution of the proposed research from both a scientific and industry perspective.

	Scientific	Industry
Status quo	<ul style="list-style-type: none"> <li>• Network representation using airports as nodes and direct flights as edges.</li> <li>• Delay multipliers (Boswell and Evans, 1997)</li> <li>• Propagation trees (AhmadBeygi et. Al, 2008)</li> </ul>	<ul style="list-style-type: none"> <li>• Basic understanding of behavior of delay propagation at complete network level.</li> <li>• Initial delay and delay multiplier define the value of a delay.</li> </ul>
Contribution	<ul style="list-style-type: none"> <li>• Represent network using individual aircraft as nodes.</li> <li>• Defining degree of separation between maximum delay propagation.</li> <li>• Make connection with spreading of behavior in social networks.</li> </ul>	<ul style="list-style-type: none"> <li>• Understand effects of delay propagation at top level.</li> <li>• Defining the value of a delay in terms of propagation to other aircraft.</li> </ul>

Figure 3.1: Overview of the status quo and contribution of the proposed research, both for the scientific field as the industry.

# 4

## Methodology

In this chapter the methodology which was applied in the research is discussed. In Section 4.1 the data used is presented, focusing on the origin of the data and the processing needed to use it in the model. Next, the model is split into two parts: network representation and delay propagation. The network representation is discussed in Section 4.3, where a variety of different network types are discussed. The method used to analyze the delay propagation is presented in Section 4.2. Finally the statistics used to validate the significance of the results is described in Section 4.4.

### 4.1. Data

For the research, the main input is quantitative from external sources. In this section the different input data is discussed in Section 4.1.1 and the processing of this data is presented in Section 4.1.2.

#### 4.1.1. Input

The input data can be divided into the main dataset and three smaller datasets. The main dataset contains all of the data needed on individual flights, while the smaller ones are additions needed in order to process the data and set-up the network.

The United States Department of Transportation (DOT) has an independent statistical agency called the Bureau of Transportation Statistics (BTS). They supply information which shapes research, investments and policy of transportation across the world. The main data which is used in the research is taken from TransStats, in particular the database 'Airline On-Time Performance Data'.<sup>1</sup>

The data contains non-stop domestic (US) flights by major carriers, the list of which ones are included is given in Table 4.1. These airlines all account for at least one percent of the domestic scheduled passenger revenues in the US and are obliged to report on-time data for most of their flights (they have uniformly elected to voluntarily report their total domestic operations). In this database, the output can be shaped to the preference of the user. Checking the boxes of the data needed will lead to a concise dataset perfectly fit to the needs of the user. The data can be downloaded for each month separately and is available around two months after the operations (i.e., data for January is available in March).

In the columns of the dataset are the categories of data which were specified by the user and each row represents an individual flight. The categories which are included for this model are given in Table 4.2. For a more detailed explanation of the data available from this database, please refer to Table A.1 in Appendix A.

---

<sup>1</sup>[https://www.transtats.bts.gov/DL\\_SelectFields.asp?Table\\_ID=236&DB\\_Short\\_Name=On-Time](https://www.transtats.bts.gov/DL_SelectFields.asp?Table_ID=236&DB_Short_Name=On-Time): accessed on March 30th, 2017.

Table 4.1: List of airlines included in the dataset.

IATA airline designator	Carrier
AA	American Airlines
AS	Alaska Airlines
B6	JetBlue Airways
DL	Delta Air Lines
EV	ExpressJet Airlines
F9	Frontier Airlines
HA	Hawaiian Airlines
NK	Spirit Airlines
OO	SkyWest Airlines
UA	United Airlines
VX	Virgin America
WN	Southwest Airlines

The second dataset contains information on each individual airport and is provided by an online international flight database <sup>2</sup>. In this dataset information can be found on each individual airport regarding the location and timezone, further specified in Table 4.3.

The final input data relates to passenger connections. The two different airline business models discussed in Section 1.1 can be associated to a certain network. For FSC operating HS networks, the hub acts as a central airport and this is also the place where passengers connect to different flights. They also tend to have codeshare agreements with other airlines, which allows them to transfer passengers on flights of those other airlines. So in this dataset the airports and airlines which allow passengers connections are defined. This overview was built for the purpose of this research, the sources used are the code-share report published by the US DOT <sup>3</sup> (used to determine codeshare partners) and the websites of individual airlines (used to determine transfer airports). This means that for each FSC, the hub airports in both their network and their partners' network are included.

When looking at LCC operating a PP network, they usually do not transfer passengers. In the case of NK and WN (both LCC), they do transfer passengers but as they do not have hubs it is difficult to define where that happens. So this dataset will define all of the airports in the network of NK and WN to be considered for transferring passengers.

<sup>2</sup><http://openflights.org/data.html>: accessed on April 20th, 2017.

<sup>3</sup><https://www.transportation.gov/office-policy/aviation-policy/code-share-report>: accessed on June 24th, 2017.

Table 4.2: Data included in TransStats dataset.

Term	Definition
FlightDate	Flight date (DD-MM-YYYY)
Carrier	IATA airline designator
TailNum	Tail number
FlightNum	Flight number
Origin	IATA code origin airport
Dest	IATA code destination airport
CRSDepTime	Scheduled gate departure time in the Computer Reservation System (CRS) (local time: HHMM)
DepTime	Actual gate departure time (local time: HHMM)
DepDelay	Departure delay, difference between DepTime and CRSDepTime. Negative delay (i.e., early departure) is set to 0.
TaxiOut	Taxi out time, in Minutes
TaxiIn	Taxi in time, in Minutes
CRSArrTime	Scheduled gate arrival time in the Computer Reservation System (CRS) (local time: HHMM)
ArrTime	Actual gate arrival time (local time: HHMM)
ArrDelay	Arrival delay, difference between ArrTime and CRSArrTime. Negative delay (i.e., early arrival) is set to 0.
Cancelled	Cancelled flight indicator (1=Yes)
Diverted	Diverted flight indicator (1=Yes)
CarrierDelay	Carrier delay, in minutes
WeatherDelay	Weather delay, in minutes
NASDelay	National Air System delay, in minutes
SecurityDelay	Security delay, in minutes
LateAircraftDelay	Late aircraft delay, in minutes

Table 4.3: Data included in OpenFlights.org dataset.

Term	Definition
Name 1	Airport name
Name 2	City name
Country	Country
IATA	IATA airport code
ICAO	ICAO airport code
Latitude	Geographic coordinate specifying north-south position in degrees.
Longitude	Geographic coordinate specifying east-west position in degrees.
Altitude	Altitude of airport in feet.
Hours offset	Hours offset from GMT.
DST	Daylight savings time region.
TZ	Name of timezone in Olson format.

### 4.1.2. Processing

After the input of the raw data, some processing is done in order to be able to use the data. First of all, the canceled and diverted flights are removed from the data. For this analysis, these flights are not relevant as they do not provide information on actual delays. In the main input data from TransStats, two columns (one for cancellations and one for diversions) exist which mark a 1 if it occurs and 0 otherwise. This allows for easy filtering of the data.

As all of the departure and arrival times are given in local time, these needed to be adapted in order to compare flight times between two flights. For this the timezone information of the second input dataset is used. For each Origin and Destination airport, the corresponding time offset from Greenwich Mean Time (GMT) is found. Eastern standard time (EST, GMT -5) is used as the baseline and all flight times are translated to this timezone. Figure 4.1 shows the major timezones in the US. Without daylight savings time (so in the winter months), EST is used in the green area. With daylight savings time (the summer months), EST is used in the yellow area. Another point of notice, is that in a large part of Arizona there is no daylight savings time. This means that during the winter months they belong to the pink area and during the summer months to the purple area.



Figure 4.1: US timezones

As flights can also depart on one day and arrive on the next, the flight date is added to the scheduled departure time to make sure the correct combination is assigned to a flight. The flight date is given (which is the scheduled local departure date) and the scheduled departure time is added to this date. This departure time was translated from local time to EST as described. For example, if the local scheduled departure time was 22:00 on August 1st in San Francisco, this would translate to 01:00 EST on August 2nd. As three hours are added to correct this time, 25 hours ( $22 + 3 = 25$ ) are added to the departure date and this automatically changes the date to the next day in EST. After this, several conditions are included which adapt these flight dates further to correct for certain cases. To determine the actual departure date:

1. If the scheduled departure time is greater than the actual departure time AND the departure delay is  $> 0$ , then the actual departure date is the scheduled departure date + 1 day. This corrects cases where the actual departure is on the next day. For example, if the scheduled departure time is at 23:45 and the actual departure time is 0:15 (which is officially the next day). The departure delay will be 30 minutes in this case, which will satisfy both requirements and add 1 day to the actual departure date. This condition is needed for around 90 flights a day.
2. If the scheduled departure time is less than the actual departure time AND the departure delay is

equal to 0, then the actual departure date is the scheduled departure date - 1 day. This deals with cases where the actual departure is on the previous day. For example, if the scheduled departure time is at 00:05 and the actual departure time is 23:55. In this case the actual departure date is the day before, so 1 day will be subtracted from the actual departure date. This condition is used for only a handful of flights (17 in total in August 2016).

3. Otherwise the actual departure date is equal to the scheduled departure date.

Next, the scheduled and actual arrival dates are determined:

1. If the scheduled departure time is greater than the scheduled arrival time, the scheduled arrival date is the scheduled departure date + 1 day. As only the scheduled departure date is known, it might happen that the flight arrives on the next day. This happens if the departure time is greater, for example if the flight is scheduled to depart at 22:00 and scheduled to arrive at 03:00. This rule is needed for around 600 flights each day.
2. If the scheduled departure time is greater than the actual arrival time, the actual arrival date is the scheduled departure date + 1 day. This covers the same cases as described previously for the scheduled arrival date, but now uses the actual arrival date. For example it might happen that has a scheduled departure time of 22:00 and arrival time of 23:30. However due to some delay the actual arrival time is 00:15, so one day should be added. This rule corrects around 800 flights a day.
3. If these conditions do not hold for the arrival times then the scheduled and/or actual arrival date is equal to the departure date.

For the most part, the flights contained in this dataset do not cross the international dateline. One exception are flights between Honolulu (HNL) and Guam (GUM), a United States territory in the Pacific Ocean. Flights crossing this line from Honolulu towards Guam, do not need special attention as the hours offset to the GMT correct the flight dates (i.e. HNL is -10, GUM is +10). Flights crossing the line from Guam to Honolulu need a slight adaptation. For every flight where Guam is the origin, the scheduled and actual arrival dates are the flight date - 1 day.

The final check is to calculate the total flight time of each flight. If the flight exceeds 24 hours, one day is subtracted from the actual arrival date. This occurs when flights are scheduled to arrive on one day, but actually arrive a day earlier (i.e., scheduled arrival time is 00:05 and actual arrival time is 23:55). After this process all of the flight times are correctly assigned the right day and time in EST.

Finally, the wheels off and wheels on time are calculated. These times are determined using the TaxiOut and TaxiIn times given in the data in Minutes. They can be added to the actual departure time and actual arrival time, respectively, to calculate the wheels off and wheels on time. This will provide accurate timing of the use of the runway.

## 4.2. Delay propagation

In this section the method applied to analyzing the delay propagation is explained. First, a few definitions are given in Section 4.2.1 which will be used throughout the report. The actual method of analyzing delays is presented in Section 4.2.2.

### 4.2.1. Definitions

First a few definitions are defined which will be used throughout the report. These definitions are related to the delay of the aircraft. According to the Federal Aviation Authority in the US, a delay in air travel is considered to be at least 15 minutes [U.S. Department of Transportation, 2016]. The delay is calculated as the difference between the actual departure/arrival time and the scheduled departure/arrival time in the CRS.

To compare the spreading of delays in air transportation an analogy can be made with a disease spreading through a social network, where the most common two models are the SIS and SIR model. These names are derived from the possible states of individuals, namely Susceptible (S), Infected (I) and Recovered (R). If the disease is considered to be delay and the individuals are the aircraft these models can be applied to model the delay in air transportation. As aircraft will never permanently recover from a delay, a SIS model is more applicable in this case. The following two definitions are used to define a state of the aircraft:

- **Susceptible:** An on-time aircraft is an aircraft which has a delay smaller than 15 minutes.
- **Infected:** A delayed aircraft is an aircraft which has a delay greater than or equal to 15 minutes.

As there is no 'live' delay information throughout the flight, an aircraft is measured at the time of its last movement. When performing a movement, every aircraft will have a certain delay. Based on that delay it is either considered to be susceptible or infected. When comparing that delay to the delay at its previous movement, it is possible to identify whether that aircraft has changed states. It is thus possible for an aircraft to get infected with a delay in the morning and perform several flights with that delay. If the delay does not change throughout the day, it will become infected once and carry the infection around for the rest of the day.

#### 4.2.2. Delay analysis

If considering a set of individuals who have interacted with each other in some way and spread the disease, an analogy can be made to a network of aircraft who have connected with each other. Some of these aircraft are carrying a disease around called delay. What happens if the neighbors of all of the individuals are observed? Can something be said about the number of neighbors that are infected?

The first step will be to define the first degree neighbors of each individual node. The expected result is that there will be more infected first degree neighbors when comparing the infected nodes to the susceptible nodes. This process is repeated for the second, third and fourth degree neighbors of each node. After a certain degree the results of both the infected and susceptible nodes are expected to be comparable. The goal will be to define the maximum degree where a difference between the two groups can be found.

### 4.3. Network representation

In this section, the way the network is represented will be explained. When thinking of the air transportation network, the most obvious representation is that of airports (nodes) connected by direct flights (edges). Delays can spread from one airport to another by these direct flights and thus delay propagation can be observed.

Instead of using this standard approach, each individual aircraft will be considered a node and the edges represent points of contact between aircraft. By doing this, the analogy with individuals carrying a disease around can be made. Individual aircraft can be identified that have interacted in some way, which created an opportunity for the disease to spread. Three methods of these interactions are defined and each will generate edges in their own way leading to three different types of networks, as presented in Section 4.3.1.

In order to analyze the delay propagation, a snapshot of the network is taken to make an analysis of which aircraft have interacted with each other in the near past. The goal of creating this snapshot is to be able to identify which aircraft have interacted with each other, much like individuals traveling to different areas. The process which is used to accomplish this explained in Section 4.3.2. Finally, to analyze the network characteristics, a few basic measures and definitions are given in Section 4.3.3.

### 4.3.1. Type of connections

To build a complex network, a set of nodes and a set of edges are required. The set of nodes is easily defined as the set of aircraft which are performing flights during a given day. The set of edges are defined according to three different mechanisms through which delay can spread. In this subsection these three mechanisms are explained and visualized. These three mechanisms will create three different networks and are labeled Type 1, 2 and 3.

#### Type 1

A type 1 edge is defined to be an edge made representing crew and passenger transfers. In the US it is fairly common to transfer crew from one aircraft to another, just like passengers on connecting flights. However, both of these are only done at hubs of airlines so it is important to limit the generation of the connections to these airports. Also, airlines transfer passengers to other airlines on so-called codeshare flights. Table 4.4 shows for each airline at which airports these connections can be made, as well as the partner airlines. Flights from an airline can also make connections at the hubs of partner airlines.

Table 4.4: List of airports and airlines which can connect to flights of a certain airline.

Airline	Hub airports	Connecting airlines
AA	CLT, DCA, DFW, JFK, LAX, LGA, MIA, ORD, PHI, PHX	AA, AS, HA, EV (only at DFW), OO (only at LAX, ORD, PHX)
AS	ANC, LAX, PDX, SAN, SEA, SJC	AA, AS, DL, OO (only at PDX, SEA)
B6	BOS, FLL, JFK, LGB, MCO, SJU	B6, HA
DL	ATL, BOS, CVG, DTW, JFK, LAX, LGA, MSP, RDU, SLC, SEA	AS, DL, EV (only at ATL, DTW), OO (ATL, CVG, DTW, LAX, MSP, SLC, SEA)
EV	ATL, DFW, DTW, EWR, IAD, ORD	AA (only at DFW), DL (only at ATL, DTW), EV, UA (only at EWR, IAH, ORD)
F9	DEN	-
HA	HNL, OGG	AA, B6, DL, HA, UA, VX
NK	All	-
OO	ATL, CVG, DEN, DTW, IAH, LAX, ORD, MSP, PDX, PHX, SEA, SFO	AA (only at LAX, ORD, PHX), AS (only at PDX, SEA), DL (only at ATL, CVG, DTW, LAX, MSP), OO, UA (only at DEN, IAH, LAX, ORD, SFO)
UA	DEN, EWR, IAD, IAH, LAX, SFO	EV (only at EWR, IAH, ORD), HA, OO (only at DEN, IAH, LAX, ORD, SFO), UA
VX	SFO, LAX	HA, VX
WN	All	-

In Table 4.4 each airline is listed, alongside their hub airports and the possible connecting airlines. The regional airlines EV and OO operate as a feeder airline for different airlines at different airports. This is why exceptions are included to only connect to certain airlines at specific airport. For example, EV operates as American Eagle (linked to AA) at DFW and Delta Connection (linked to DL) at ATL. The source of this data is given in Section 4.1.1.

Next a time window needs to be set, within which these connections can influence each other in terms of delay. The hypothesis is that any incoming aircraft that has a scheduled arrival time between 30-120 minutes before an outgoing aircraft, will offer a possibility of connecting passengers and crew. Another hypothesis is that at some point, the possible connection will be deleted by the airlines if the incoming flight is too severely delayed. Since this decision point is a bit ambiguous (i.e. when does an airline decide this), it is decided to include an edge if the incoming flight actually arrives before the outgoing flight. The visualization of this type of edge is presented in a time-space network in Figure 4.2.

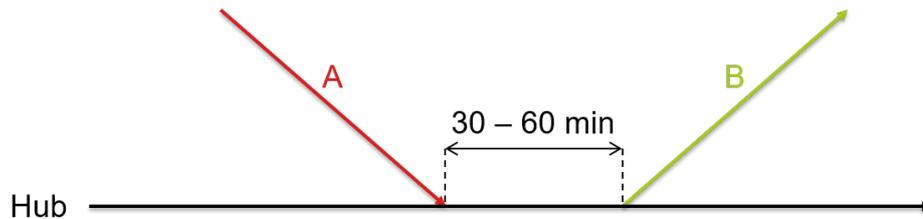


Figure 4.2: Time-space network of type 1 connection from aircraft A to aircraft B.

The formal definition for a type 1 connection is as follows:

**Type 1 - A connection from one inbound aircraft A towards one outbound aircraft B, with the following requirements:**

1. The scheduled (gate) arrival time of aircraft A is between 30-120 minutes before the scheduled (gate) departure time of aircraft B.
2. The actual (gate) arrival time of aircraft A is before the actual (gate) departure time of aircraft B.
3. The first and second requirement happens at the same airport.
4. Aircraft A and aircraft B are operated by the same airline or partner airline (refer to Table 4.4).
5. Aircraft A and aircraft B are located at the hub of (one of) their airline(s) (refer to Table 4.4).

#### Type 2

At type 2 edge is defined to be an edge made representing the use of the runways at an airport. In Europe, a slot system is in place which regulates how many aircraft arrive and depart at a specific airport. Airlines need to acquire timeslots at airports within which they can plan a flight. In the US this system is the other way around. The airlines make their flight schedules which are sent to the airports. This means that airports do not have control over how many aircraft are planned to arrive at a certain time.

The assumption in this type of connection is that when many delayed aircraft arrive or depart at a certain airport, they will have an effect on the aircraft also arriving or departing at that airport. The time window for these connections is defined to be plus or minus 15 minutes around the time of a movement. Due to separation restrictions between aircraft using the same runway, up to 7 or 8 aircraft can use the runway within 15 minutes. This means that an aircraft can influence up to this number of aircraft before and after its movement per active runway at an airport.

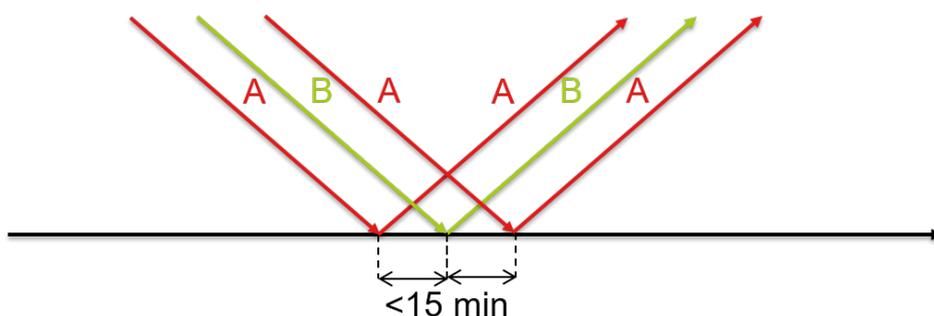


Figure 4.3: Time-space network of type 2 connection from aircraft A to aircraft B.

Figure 4.3 visualizes this type 2 connection in a time-space network. Any aircraft A that arrives or departs from the same airport within plus or minus 15 minutes of a movement of aircraft B. The formal definition is as follows:

**Type 2 - A connection from one aircraft A towards one aircraft B, with the following requirements:**

1. The actual wheels off/on time of aircraft A is within plus or minus 15 minutes of the actual wheels off/on time of aircraft B.
2. The first requirement happens at the same airport.

**Type 3**

A type 3 edges is defined to be a connection made representing the use of airport resources. These resources (such as gates, baggage handling, towing vehicles and fuel trucks) are in limited supply. The hypothesis is that if flights are delayed they are using up resources which are supposed to be used for other flights.

A type 3 connection is thus made if two aircraft are present at the same airport at the same time. The time used to determine if this happens, is the actual gate arrival/departure time of each aircraft. The formal definition is as follows:

**Type 3 - A connection from one aircraft A towards one aircraft B, with the following requirement:**

1. Aircraft A is located at the same airport as Aircraft B at a point in time.

**4.3.2. Snapshot**

When analyzing social networks, the network is a static representation of the situation. In this case, aircraft are constantly moving around and creating new edges. For this a reason, a process is defined which will create a snapshot of the situation and this is explained in this section.

The first step is to define the time at which the snapshot needs to be made. After this time has been defined, the flight status of each individual aircraft is checked. There are two possible scenarios:

1. The aircraft is in-flight at the defined time. If this is the case, the edges which were made at the previous airport are included in the network. In this case that would included edges made by the two movements (arrival and departure) at this airport. As the aircraft is in-flight at the defined time, the delay is measured at the departure from the last airport. It is assumed that any edges made at this airport might have had an effect on this delay.
2. The aircraft is on-ground at the defined time. If this is the case, the edges which were made at the previous and current airport are included in the network (i.e., this would include the arrival at the current airport and the arrival and departure at the previous airport). The delay of the aircraft is measured at the arrival at the current airport. The assumption is that this delay was affected by the edges made at this arrival, but also by edges made at the previous airport.

If any of the movements are more than 6 hours before the defined time, these connections are not included in the network. It is assumed that connections that were made before this time no longer have an effect on the current network and should not be used in the analysis of the snapshot. For example, an aircraft which arrived at the current airport 7 hours ago, will have generated edges from other aircraft at that point. The last delay for that aircraft will be measured at that point 7 hours ago, but the other aircraft might have already performed several more flights in the meantime. The longer this time limit, the more of this noise will be included in the network. A sensitivity analysis is performed in Section 6.1.3 which will show the effect of this assumption.

The last step is to exclude all the isolate nodes in the network. They have no effect on the further analysis and removing them simplifies the network characteristics. An aircraft can be isolated due to two reasons:

1. If it has not performed a movement within the last 6 hours.
2. If it is in-flight at the reference time and its previous airport was so small that no connections were made.

The whole process is summarized and visualized in Figure 4.4, including an overview of what edges are made where. The flat lines represent time spent at an airport while the dotted arcs represent the aircraft being in-flight. The arrows with number represent the different types of edges. Type 1 edges are generated at the departure of an aircraft, type 2 at both a departure and an arrival and type 3 throughout the stay of an aircraft at an airport as well as at the last arrival (if on-ground). The red-dotted line represents the time of the last movement of an aircraft. After this no more information is available for a certain aircraft so no more incoming connections are recorded. However, outgoing connections of type 2 can be made for in-flight aircraft and outgoing connections of all types can be made for on-ground aircraft. So an aircraft can no longer be infected by other aircraft after the red line in Figure 4.4, but they can still spread the infection to other aircraft.

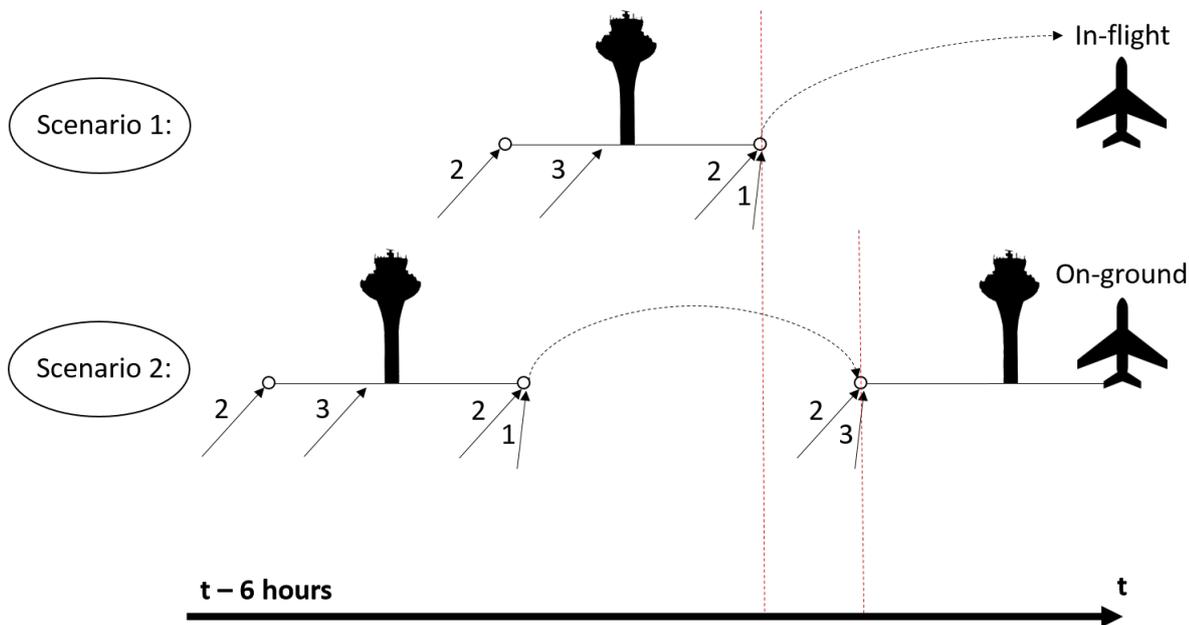


Figure 4.4: Visualization of snapshot creation. Two scenarios are presented, the aircraft is either in-flight or on-ground. Based on this status, either connections made at the last two or three movements are included. The arrows define the points where connections can be made for the different types of edges.

### 4.3.3. Network measures

Several measures can be identified which can be used to analyze the networks. Some have been mentioned in the literature review in Section 2.2.1 but they will be repeated in this subsection to provide a complete overview of the measures that will be used. A basic graph consists of a set of  $N$  nodes ( $n_1, n_2, \dots, n_N$ ) and a set of  $k$  edges ( $l_1, l_2, \dots, l_k$ ). In this case a directed graph is used, which means that the edges are further defined by the direction (i.e.,  $l_{ij}$  is not necessarily equal to  $l_{ji}$ ).

The most basic metric is the degree of a node. In the graph, every node will have connections to neighboring nodes. The degree of a node is the number of edges that are connected to that node. In a directed graph, which it will be for Type 1, an in-degree and out-degree can be identified. Respectively, these are defined the number of incoming edges and outgoing edges. Another characteristic is be the average degree of a node, how well is a node connected on average.

By combining the degree of all individual nodes, a degree distribution can be made. This distribution defines the probability of picking a random node from the network with a certain degree. The most important conclusion from these distributions is the shape, where the two most common shapes are

the power-law and the Poisson distribution. To make an accurate power-law fit, a Maximum Likelihood Estimation (MLE) approach is used. For scale-free networks, the higher values of degree should follow the power-law distribution Barabasi and Albert [1999]. The best fit for the power-law distribution can be calculated using MLE starting at each unique degree. Then the Kolmogorov-Smirnov distance is calculated between each fit and the data, the minimum value represents the most accurate starting point for the power-law distribution. The Poisson distribution is represented by a mean and standard deviation.

In a graph the number of weakly and strongly connected components can be identified. A strongly connected component of a graph is defined to be a subgraph within the network where there is a path from any node A to any node B and vice versa. A weakly connected component of a graph is a component of the network where if the directed edges are replaced with undirected edges a strongly connected component is created. In a directed graph there tend to be more strongly connected components than weakly connected components, while in an undirected graph these values are the same.

Another measurement which is important in defining the network characteristics is the average shortest path length. Equation (4.1) shows the mathematical formulation used to calculate this measurement. In this equation  $d_{ij}$  is the shortest path length from a node  $i$  to a node  $j$  (i.e., for a direct path  $d_{ij} = 1$ , if passing through 1 node  $d_{ij} = 2$ , etc.).  $N$  in this equation is equal to the total number of nodes in the network. In a network, the maximum value for the shortest path length is called the diameter.

$$L = \frac{1}{N(N-1)} \sum_{i,j \in N, i \neq j} d_{ij} \quad (4.1)$$

In a directed graph or a graph with unconnected subnetworks, it might not always be possible to determine a path from one node to every other node. In this case, the number of paths that are possible are included and the average length of these paths is determined [Mao and Zhang, 2013]. If this requires too much computational power, the largest strongly connected component can be taken. In this component a path exists between every pair nodes (both directions) thus the average shortest path length can be determined, however this was not a problem during this research.

When a certain degree neighbor of a node is considered, it refers to the distance between the two nodes. A first degree neighbor of a node has direct path, to get to a second degree neighbor one other node needs to be traversed, to get to a third degree neighbor two other nodes need to be traversed, etc.

Furthermore, in a directed graph it is important to make the distinction between predecessors and successors. As the edges in a directed graph point in one direction, the neighbors of a node can be split into these two categories. Predecessors are nodes that have an edges connecting them towards a specific node, while successors are nodes that are connected by an edge pointed away from that specific node.

The clustering coefficient of a directed graph can be calculated using Equation (4.2) [Watts and Strogatz, 1998].  $N_i$  is considered the neighborhood of a node  $n_i$  and  $k_i$  is the number of nodes in  $N_i$ . A node  $n_j$  is in that neighborhood if a direct edges exists between the  $n_i$  and  $n_j$  (denoted either  $l_{ij}$  or  $l_{ji}$  depending on direction). The clustering coefficient  $c_i$  of a node is calculated by summing up the number of edges that exist between 2 neighboring nodes of  $n_i$  and dividing by  $k_i(k_i - 1)$ , which is the maximum number of possible edges within the neighborhood.

$$c_i = \frac{|\{l_{jk} : n_j, n_k \in N_i, l_{jk} \in E\}|}{k_i(k_i - 1)} \quad (4.2)$$

## 4.4. Statistical analysis

After creating the network and analyzing the neighbors of individual aircraft, these values will need to undergo statistical tests. The goal of these tests will be to prove the significance of the results. A statistical test sometimes requires samples to be independent of each other, which can be achieved by different sampling techniques. The sampling method is discussed in Section 4.4.1, which will take a random subset of the entire dataset in order to prove the hypotheses. Next a statistical test is discussed (Section 4.4.2) which will be used throughout the report to substantiate the results, the Mann-Whitney U test.

### 4.4.1. Sampling

Sampling is used to take a subset of the entire population, in this case the population is considered to be the set of individual aircraft. These aircraft can split into an infected population and a susceptible population, according to their last known delay.

Several different sampling techniques were considered in this research. After testing several of these techniques, random sampling was chosen over geographic sampling. Using geographic sampling would create a bias towards the smaller areas (i.e., airports or US states). For instance, if considering sampling one aircraft per airport. A smaller airport might have only one aircraft so that one would always be sampled, while a larger airport has a larger group of aircraft to take a sample from. Also, the last known location for in-flight aircraft is the last airport, so several aircraft can be sampled which are all en-route to the same airport.

The simplest sampling method is random sampling, which is accomplished by randomly choosing a certain number of aircraft from the population. If there are two groups (i.e., an infected population vs. a susceptible population) that need to be compared, there is also the possibility of choosing a certain number from each population. If there are two groups with different sizes, relative random sampling might be used. Here a certain number of aircraft are randomly drawn from each group, depending on the group size. For example, if there is one group with 50 aircraft and one group with 100 aircraft, random samples will be drawn with the ratio 1:2. This last method will be used throughout the rest of the report, as it accurately represents a random (assumed independent) subset of the total population.

### 4.4.2. Mann-Whitney U test

In 1947 a statistical test was designed which determined whether one random variable was statistically larger than another [Mann and Whitney, 1947]. The test was named after the authors, the Mann-Whitney U test. The fact that this test is non-parametric offers one main advantage over a parametric test (i.e., the Students t-distribution), which is that the distribution of the variable is not fixed (i.e., a Students t-distribution would require a normally distributed variable). In a non-parametric test the distribution is not determined before the test, but rather follows from the test.

The Mann-Whitney U test attempts to compare two groups of data from different populations. Assume the groups of data are composed of individual measurements  $\{a_1, a_2, \dots, a_i\}$  and  $\{b_1, b_2, \dots, b_j\}$ . The test is based on comparing each individual member of population A ( $a_i$ ) with each individual member of population B ( $b_j$ ), the total number of possible pairs thus being  $i \times j$ .

The null hypothesis of the test assumes that the median of each of the two groups is equal, as presented in Equation (4.3). This means that when choosing one observation from each group and comparing them, the probability of  $a_i$  being larger than  $b_j$  is the same as the probability of  $b_j$  being larger than  $a_i$  (i.e.,  $P(a_i > b_j) = P(b_j > a_i)$ ).

$$H_0: P(a_n > b_m) = \frac{1}{2} \quad (4.3)$$

The alternative hypothesis depends on whether a two-tailed or one-tailed test is needed. In a two-tailed test the alternative hypothesis is that the probabilities are not equal, as shown in Equation (4.4). The

alternative hypothesis in a one-tailed test is that one of the probabilities is larger than the other, which is presented mathematically in Equation (4.5). In the two-tailed test the null hypothesis is rejected if the test statistic falls into either tail of the distribution, while in the one-tailed test the null hypothesis is only rejected if it falls into one specified tail of the distribution. Figure 4.5 illustrates this in a figure, where the rejection region is marked in green.

$$H1: P(a_n > b_m) \neq P(b_m > a_n) \quad (4.4)$$

$$H1: P(a_n > b_m) > P(b_m > a_n) \quad (4.5)$$

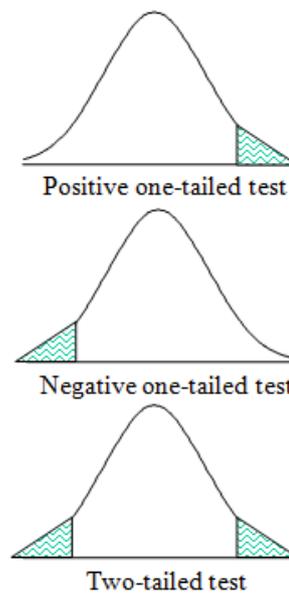


Figure 4.5: Rejection region of one-tailed tests and two-tailed tests.

To perform the Mann-Whitney U test, a statistic is calculated which is usually denoted with  $U$ . For large sample sizes this statistic can be assumed to have a normal distribution. The first step is to order all the individual observations according to their value (mixing the two populations). The smallest value is given the rank 1 and the highest value is given the rank  $n$ , the number of total observations. If two values are equal they are assigned half points (i.e., if they are number 10 and 11, they are both assigned 10.5).

Equation (4.6) and Equation (4.7) show how the test statistic for each group can be calculated.  $n$  is the number of observations that are used of each group, while  $R$  is the sum of all the ranks of each group. The maximum value of the test statistic of one group is the product of the two group sizes,  $n_a \times n_b$ . This would mean that all of the observations in one group are larger than all of the observations in the other. The sum of both statistics is also equal to the product of the two group sizes. Under the null hypothesis, the values of both test statistics are approximately equal.

$$U_a = (n_a n_b) + \frac{n_a(n_a + 1)}{2} - R_a \quad (4.6)$$

$$U_b = (n_a n_b) + \frac{n_b(n_b + 1)}{2} - R_b \quad (4.7)$$

As mentioned before, for large sample sizes ( $> 20$ ) the test statistic can be approximated by a normal distribution. Using this fact, a Z-score can be calculated based on the average value of U (Equation (4.8)) and the standard deviation of U (Equation (4.9)). In this case,  $N$  is equal to the sum of both group sizes ( $n_a + n_b$ ) The equation to calculate the Z-score is given in Equation (4.10).

$$\mu_U = \frac{n_a + n_b}{2} \quad (4.8)$$

$$\sigma_U = \sqrt{\frac{n_a n_b (N + 1)}{12}} \quad (4.9)$$

$$Z = \frac{U_{\max} - \mu_U}{\sigma_U} \quad (4.10)$$

Once this Z-score is calculated, a significance level needs to be determined. This significance level can be used to find a tabulated Z-score. The null hypothesis is rejected (and thus the alternative hypothesis accepted) if the calculated Z-score is larger than the tabulated z-score, as presented in Equation (4.11).

$$\text{Reject } H_0 \text{ if: } \textit{calculated } Z \geq \textit{tabulated } Z \quad (4.11)$$

## 4.5. Summary

In this section, the methodology is summarized and the final model is visualized in a functional flow diagram in Figure 4.6.

Three datasets are used as input in the model: 1. The RITA dataset provides delay data for individual flights, 2. OpenFlights dataset provides the timezones for each airport and 3. passenger connections are defined according to codeshare agreements from the DOT. The first two datasets are combined to process the schedules. The local times are transformed into one standard timezone (EST) as described in Section 4.1.2.

After selecting a certain day, the possible edges from other aircraft are defined for each movement. The different types of edges are presented in Section 4.3.1. Then, a time has to be defined at which the network is to be analyzed. The network is then built according to the snapshot process described in Section 4.3.2. This network is then used to analyze the delay in the network as presented in Section 4.2.2. This process of snapshot creation and delay analysis can be repeated for different times of day.

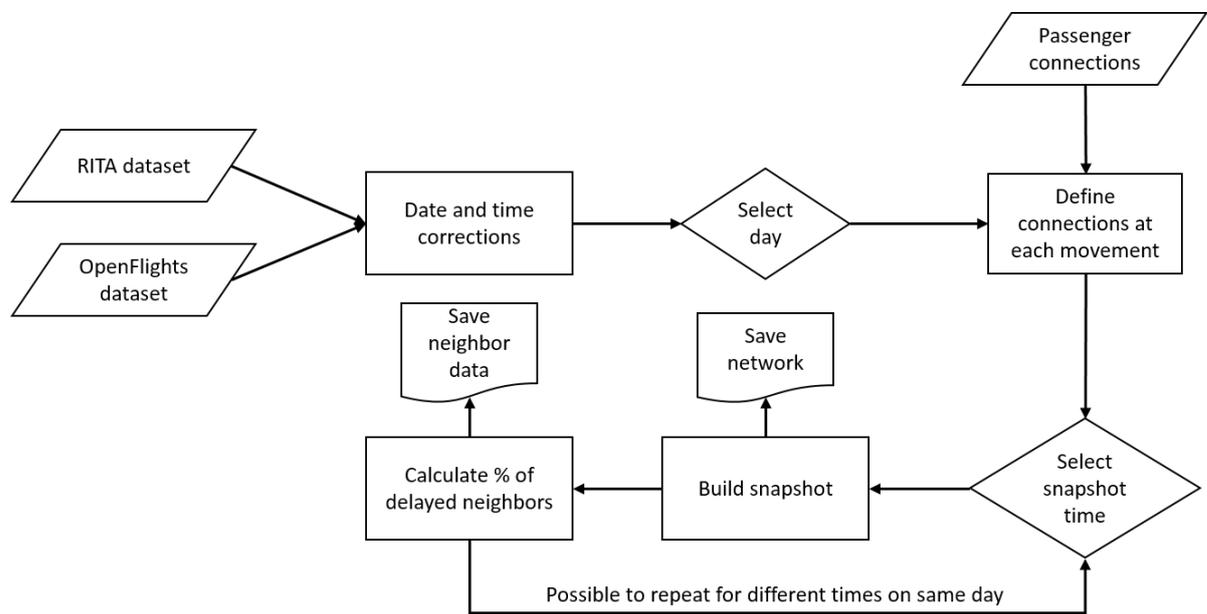


Figure 4.6: Functional flow diagram of model.



# 5

## Results

In this chapter the results of the implementation of the methodology discussed in the previous chapter is presented. First of all, a general analysis is done on delay patterns in air transportation and will be presented in Section 5.1. Next the resulting network structures will be discussed in Section 5.2. Finally the delay propagation is analyzed in Section 5.3.

### 5.1. General delay analysis

Delays in air transportation are caused many different factors. These factors will play different roles on different days (i.e., bad weather, unexpected closures of airspace, security problems, etc.). This nature makes it hard to predict when delays occur and with what magnitude.

Figure 5.1 shows the total arrival delay of all domestic US flights on every day of August 2016. The delay varies from 100,000 minutes up to more than 600,000 minutes. No pattern can be found in the arrival delay on a certain day. For instance, weekdays do not have higher delays than weekends and Fridays are not worse than a Sunday.

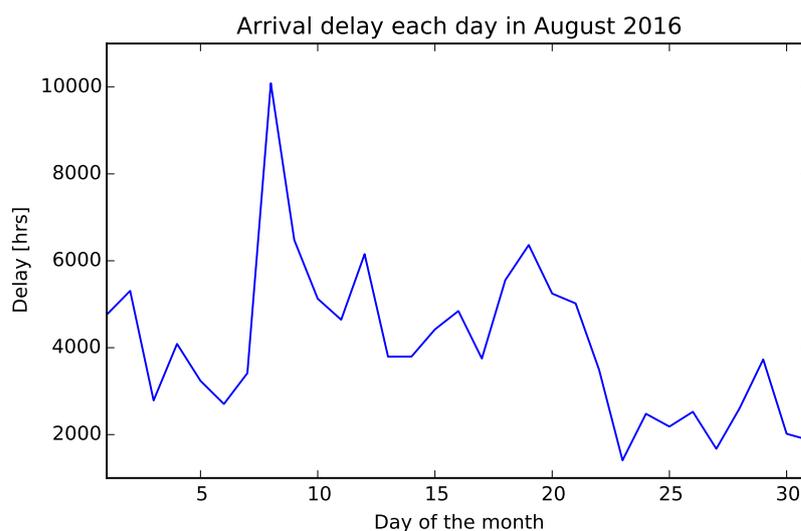


Figure 5.1: The total arrival delay of each day in August 2016.

The number of unique aircraft that are performing flights each day of August 2016 is plotted in Figure 5.2.

This can be seen to be a relatively constant number, with slight dips on Saturdays. The number of aircraft infected is random, much like the arrival delay pattern.

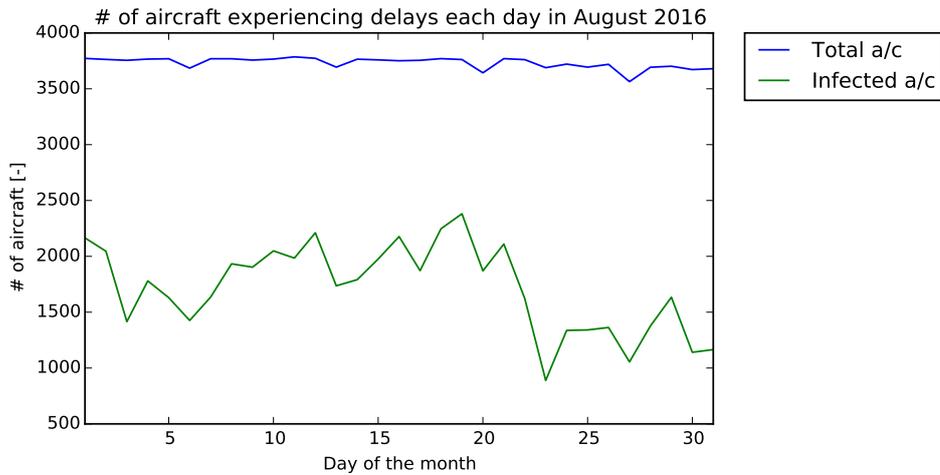


Figure 5.2: The number of aircraft performing flights and the number of infected aircraft.

Dividing the values of the two graphs from Figure 5.2 leads to the percentage of unique aircraft that are infected on a specific day. This percentage is plotted against the delay pattern in Figure 5.3. For the most part, it can be observed that these two graphs follow the same pattern. If the total delay increases, so does the percentage of aircraft infected and vice versa.

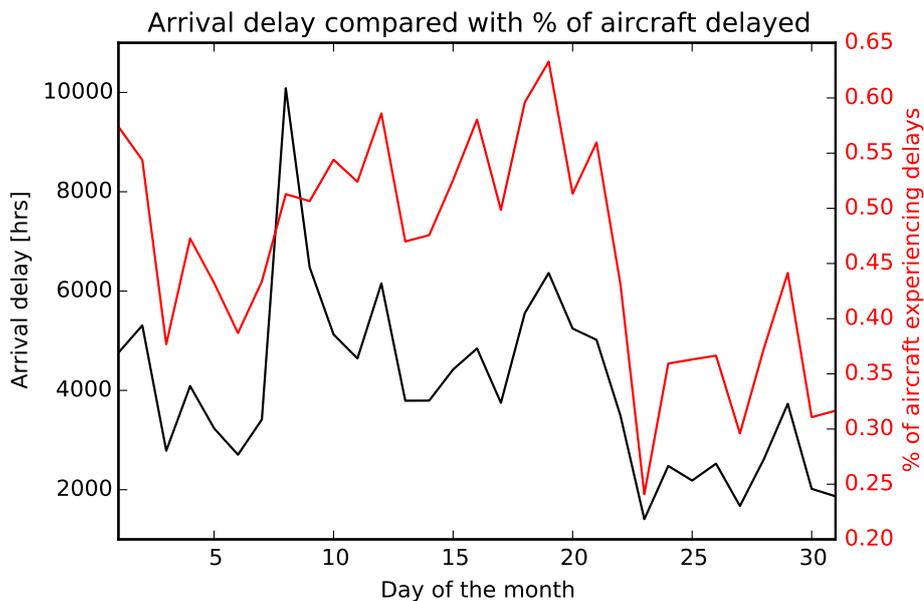


Figure 5.3: The arrival delay versus the percentage of aircraft infected.

One exception is the high peak of arrival delay on August 8th, 2016. With this massive amount of delay (>600,000 minutes), one would expect the number of aircraft infected to also peak. However, this is not the case as percentage of infected aircraft does increase but not as much as expected.

Research leads to the conclusion that on August 8th, 2016, Delta airlines experienced a 6-hour global computer system outage <sup>1</sup>. The massive amount of delay is thus caused by a failure of one airline,

<sup>1</sup><http://money.cnn.com/2016/08/08/news/companies/delta-system-outage-flights/>: accessed on June 1st, 2017.

which explains why the number of aircraft infected does not increase as expected. This extreme case will be avoided for the basic network analysis, but the delays on this day will be analyzed individually.

Throughout the day, a pattern can be observed in the amount of delay. This is visualized in Figure 5.4. The development of delay throughout the day follows a gradual increase throughout the day and a hard decline after around 10 PM. In Figure 5.4 the arrival time is plotted versus the arrival delay. The arrival times are split into 19 bins, 1 for all flights up until 06:00 and 1 for each other hour of the day.

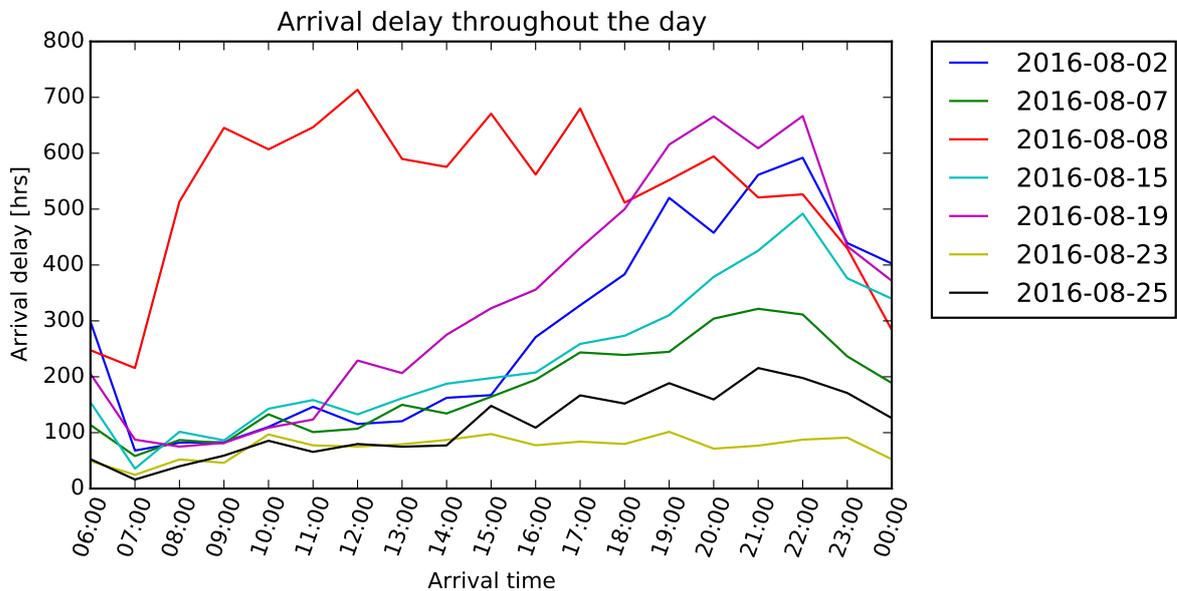


Figure 5.4: The development of arrival delays throughout individual days.

Again, August 8th forms an exception in this pattern. On this day the peak of the arrival day was reached in the morning, unlike the other days. This is the result of the computer system failure mentioned before, which happened during the early morning.

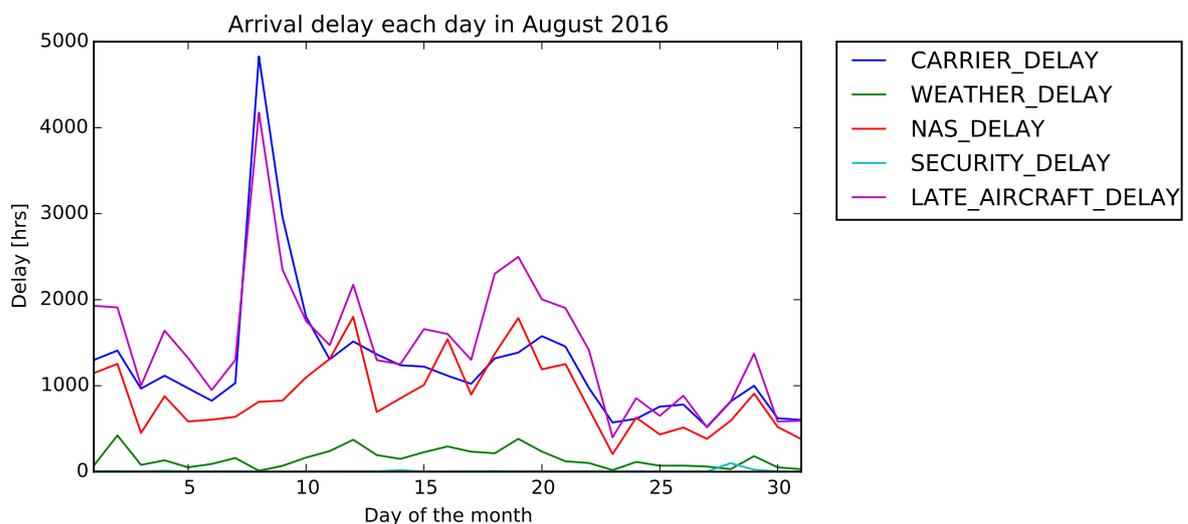


Figure 5.5: The magnitude of different types of delays throughout August 2016.

It is important to note that the shape of the delay development shown in Figure 5.4 might be similar for each day, the peak of the shape varies. This means that whatever conclusions are made for the behavior of delays on one day, must be validated with the behavior on other days.

The different types of delay (mentioned in Chapter 1) identified by the FAA are plotted in Figure 5.5. Clearly, the delay on August 8th was caused by an abnormal spike in carrier delay (caused by internal problems at airlines) and late aircraft delay (caused by late arrival of the previous flight). Again this was the cause of the computer failures at Delta Airlines.

Generally, the three major causes of delay are the carrier delay, late aircraft delay and airspace system delay (caused by congestion in airspace/airports). Weather delay is a fluctuating factor in causing delay, while security delay is a very minor factor.

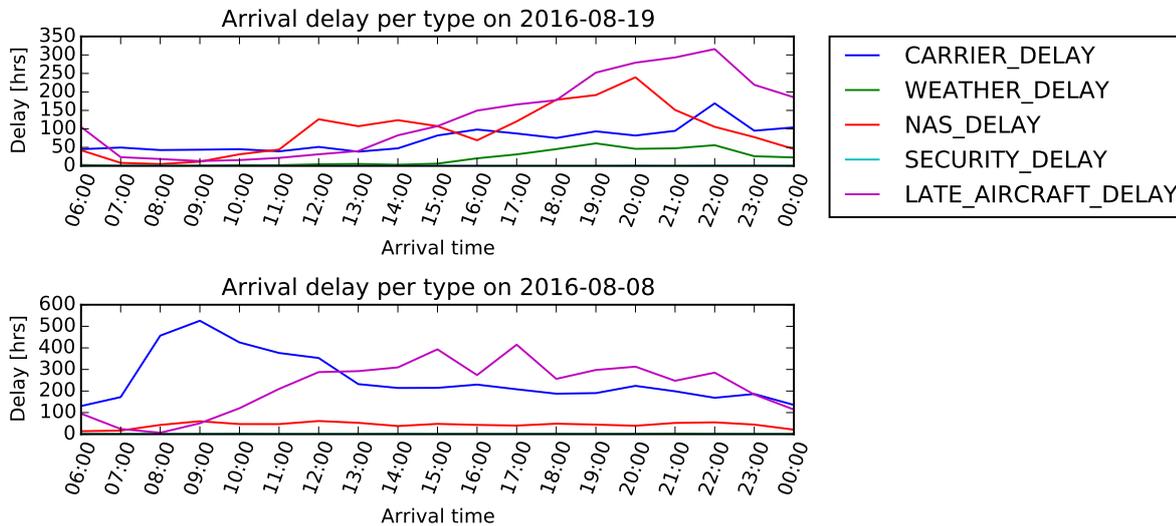


Figure 5.6: The magnitude of different types of delays on August 8th and 19th.

Preceding a peak in late aircraft delay will always be a peak in another type of delay. Figure 5.6 shows the development of the different causes of delay on August 8th and 19th, 2016. The morning peak of carrier delay on the 8th is a direct result of the computer failure at Delta Airlines. The rest of the day shows a high peak of late aircraft delay, which is the result of earlier flights being affected by the carrier delay. On the 19th, there is a peak of airspace delay towards the end of the afternoon. Again this is followed by a peak of late aircraft delay, which is caused by the initial airspace delay.

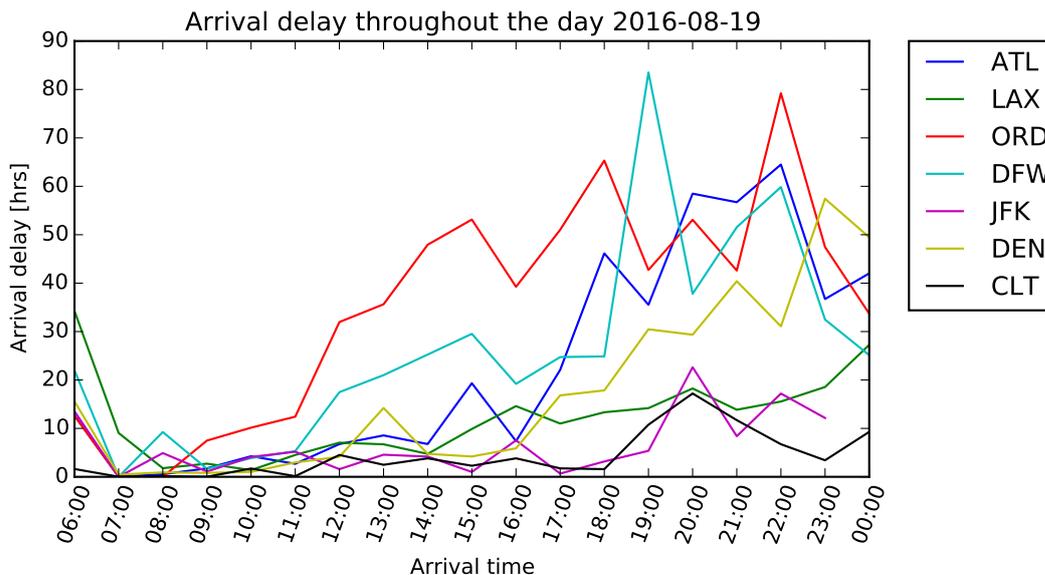


Figure 5.7: Delay development at the top 10 largest airports of the US on August 19th, 2016.

If individual airports are considered, delays show a similar pattern of development. This can be seen in Figure 5.7, where the arrival time blocks are plotted versus the arrival delay for the top 10 airports in the US in terms of movements. The delays develop in a similar matter as the delay in the entire network, as seen in Figure 5.4.

## 5.2. Network representation

Creating the network as described in Section 4.3.2, a snapshot of aircraft interacting with each other is created. The results of the three different types of contact discussed in Section 4.3.1 will be presented separately. For each type, the resulting network characteristics are discussed in Section 5.2.1, Section 5.2.2 and Section 5.2.3 respectively. After discussing the types individually, in Section 5.2.4 a comparison is made between the different types of edges and comparable real-world networks are identified.

The first effect that is discussed is the variation in setting the snapshot time throughout the day. What is the difference between a network at 7 A.M. and a network at 10 P.M. The other major effect is the difference between the days of the month. Does it matter if the network is built on Saturday or Monday? These questions will be handled per connection type. The individual day that is discussed will be August 15th, which experienced a little above average delay. The multiple days that are further compared are the ones given in Figure 5.4, with the exception of August 8th. This will include several days with a variety of total delay, as to get an average overview.

### 5.2.1. Type 1 network

To investigate the effects of building a network at a certain time of day, snapshots are built for several different times on August 15th. The resulting network characteristics are presented numerically in Table 5.1 and visually in Figure 5.8.

Table 5.1: Generating type 1 snapshot for several different times on 15-08-2016.

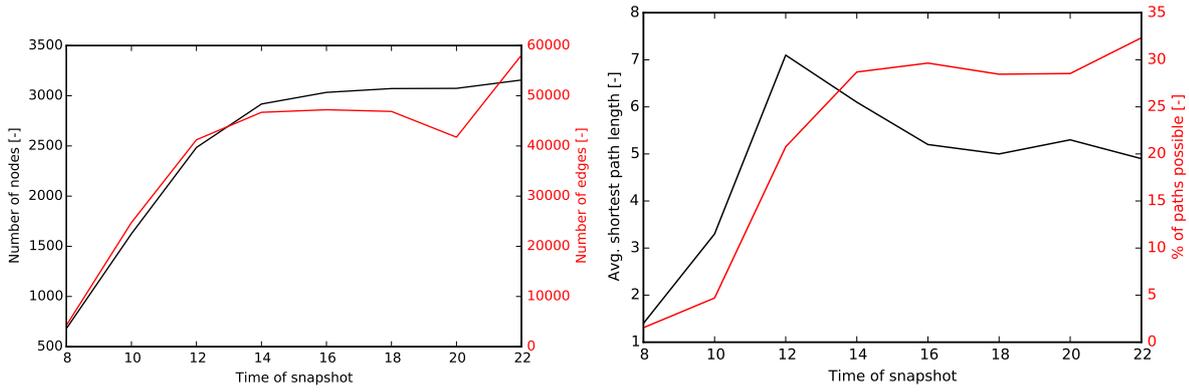
Time snapshot	Nodes	Edges	Avg. degree	Strong	Weak	% paths possible	Avg. shortest path	$C$
<b>08:00</b>	681	4263	6.3	600	29	1.7%	1.4	0.26
<b>10:00</b>	1625	24734	15.2	1074	44	4.8%	3.3	0.43
<b>12:00</b>	2485	41233	16.6	1292	39	20.8%	7.1	0.43
<b>14:00</b>	2918	46704	16	1325	27	28.7%	6.1	0.40
<b>16:00</b>	3034	47219	15.6	1340	18	29.7%	5.2	0.39
<b>18:00</b>	3072	46883	15.3	1375	18	28.5%	5	0.37
<b>20:00</b>	3074	41752	13.6	1404	11	28.6%	5.3	0.36
<b>22:00</b>	3157	58114	18.4	1269	24	32.4%	4.9	0.38

Figure 5.8a shows the development of the number of nodes and edges throughout this day. Eight snapshots were computed in order to make this graph (8:00 - 22:00, every two hours). The network at 8 and 10 AM are clear outliers, as the number of nodes is significantly lower (-60%) and the number of edges is close to 0.

Creating the snapshot at 10 A.M. will lead to an increase of the number of nodes and edges. This effect is caused by the fact that the day of operation is still starting up. 7 A.M. EST means 4 A.M. on the west coast of the US, so there are still many aircraft sitting on the ground at this point. Moving a bit further in the day leads to flights being operated all over the country and thus an increase in movements.

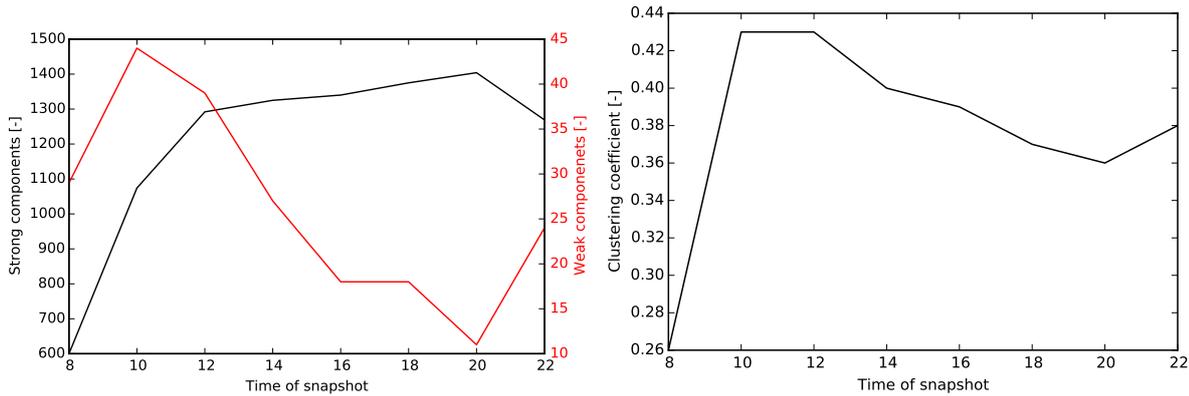
Looking at the number of nodes and edges for the rest of the day, the values seem to be relatively stable. The number of nodes definitely stabilizes from 12:00 onwards and the number of edges still varies (20% difference between 12:00 and 22:00).

Next, the degree distribution is computed for the different snapshots. Figure 5.9 is the degree distribution for three snapshots throughout the day. These distributions seem to follow a similar distribution,



(a) Number of nodes and edges.

(b) Average shortest path length and % of paths possible.



(c) Weak and strong connected components

(d) Clustering coefficient

Figure 5.8: Development of network characteristics for type 1 snapshots on 15-08-2016.

where for high degrees a power-law distribution is identified which can be approximated by  $P(k) \approx k^{-\gamma}$ . Using a maximum likelihood fit, these coefficients can be determined and they are presented in Table 5.2.

The minimum value of degree for which the best fit for the power-law distribution was found is also given in Table 5.2. To visualize what this power-law fit looks like, the degree distribution is plotted on a logarithmic scale in Figure 5.10. For the last snapshot at 22:00, the best power-law fit starts at 167 and the  $\gamma$  is extremely high at 28.25. The best power-law fit is thus found for only a very small part of the degree distribution (the maximum degree is 207). The rest of the snapshots show a power-law distribution for a larger part of the degree distribution, with  $\gamma$  varying between 2.5 and 3.4.

The node with the highest degree for the 20:00 snapshot on August 15th is N967AT, with a total degree of 165. This aircraft is operated by Delta Air Lines and has just departed from Atlanta at 19:54. As this is a hub airport for Delta 79 edges are created by this departure. Furthermore it arrived at Atlanta 18:07 on a flight from Birmingham, where it created outgoing edges (86) to aircraft departing from Atlanta.

In Figure 5.11 the total 20:00 snapshot is plotted in a graph. The nodes are colored by their last known status, the red nodes are infected and the green nodes are susceptible. The clustering is based on the edges between the nodes, triangles of edges between 3 nodes attract towards each other. Because of the nature of this connection type, clusters are made for airports. For example, the large cluster at the bottom is Atlanta (hub for Delta) and the top right is Dallas (hub for American).

Figure 5.8b shows the average shortest path length as a function of the snapshot time. Again the same characteristics can be found as in the number of nodes and edges. The value of the average shortest path length is short in the morning while the afternoon values are relatively stable. Also the number of

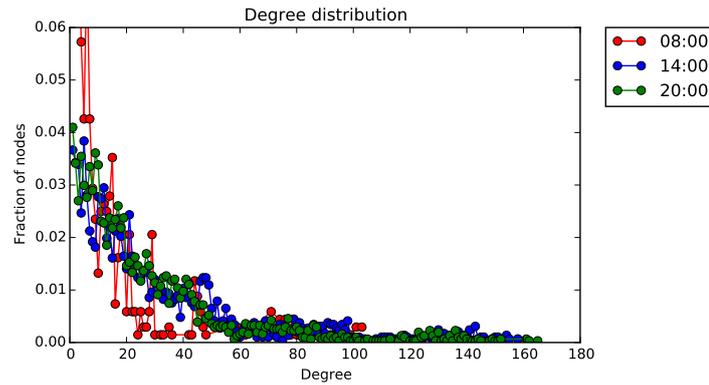


Figure 5.9: Degree distribution for three snapshots of type 1 on 15-08-2016.

Table 5.2: Best fit power-law coefficients for type 1 snapshots on 15-08-2016.

Time snapshot	$\gamma$	Minimum degree
08:00	2.50	12
10:00	2.62	37
12:00	2.70	34
14:00	2.80	37
16:00	3.42	58
18:00	3.06	45
20:00	2.84	31
22:00	28.25	167

paths which are possible is plotted in Figure 5.8b. This percentage of possible paths grows with the number of edges created. The more edges, the more possible paths between nodes.

The diameter of the 20:00 snapshot is equal to 18 and this occurs twice in the entire network. An example where this occurs is from N788AA to N863AA. These aircraft are both operated by American Airlines and have just departed from New York JFK and Kahului (Hawaii) respectively.

Figure 5.8c shows the number of these components versus the time of the snapshot. The number of strongly connected components increases throughout the morning and then reaches a relatively stable level, much like the number of nodes. In the end the number of strongly connected components ends up at around 1700. This means that the average size of a strongly connected component is around 2. The cause of this is that only one departure is included for each aircraft. This limits the connection possibilities to one airport only.

The number of weakly connected components varies a bit throughout the day and no clear pattern can be established. As this connection is limited to certain airports and airlines, the graph will never be fully connected (i.e., it is impossible for this number to be 1).

This has all been for just one single day, what does the network look like if other days are considered? To capture this, the network has been calculated for six different days in August at 4 different times (16:00, 18:00, 20:00 and 22:00). These times were chosen as the network of August 15th seemed to stabilize in terms of characteristics in the afternoon. The mean network characteristics are given in Table 5.3.

The coefficient of variation is the standard deviation divided by the mean. This leads to a (unitless) measure of spread relative to the mean, which is more valuable than the standard deviation itself. The number of nodes is found to be very stable, with 1.9% it has a very low coefficient of variation. The

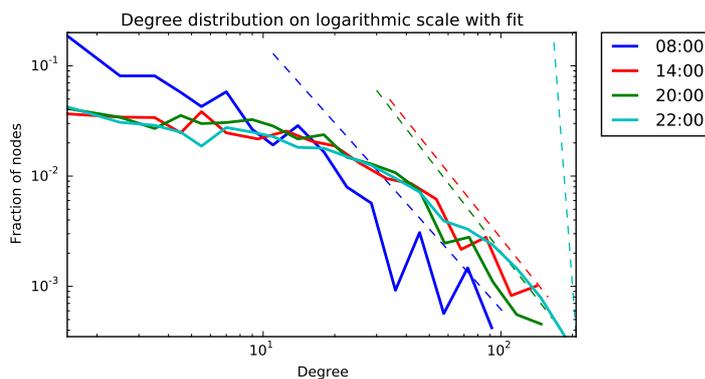


Figure 5.10: The degree distribution on logarithmic scale (solid line) with their best power-law fit (dotted line) for four snapshots of Type 1 on 15-08-2016.

Table 5.3: Generating type 1 snapshot for several different days in August 2016.

Characteristic	Mean	Standard deviation	Coefficient of variation	Minimum	Maximum
<b>Nodes</b>	3,050	58	1.9%	2,953	3,175
<b>Edges</b>	48,860	6,635	13.6%	37,999	64,300
<b>Avg. degree</b>	16	2	12.2%	13	21
<b>Strong</b>	1,292	73	5.7%	1,122	1,404
<b>Weak</b>	20	4	20.4%	11	29
<b>% paths possible</b>	30.4%	1.9%	6.2%	27.6%	34.7%
<b>Avg. shortest path</b>	5.3	0.2	4.3%	4.8	5.7
<b>C</b>	0.39	0.02	5.2%	0.36	0.43

largest coefficient of variation is found in the weakly connected components, with 17.9%. Even this coefficient is not considered high, as less than 1 is considered low-variance. From this table, it can be concluded that the type 1 network is relatively stable in terms of shape from 16:00 to 22:00.

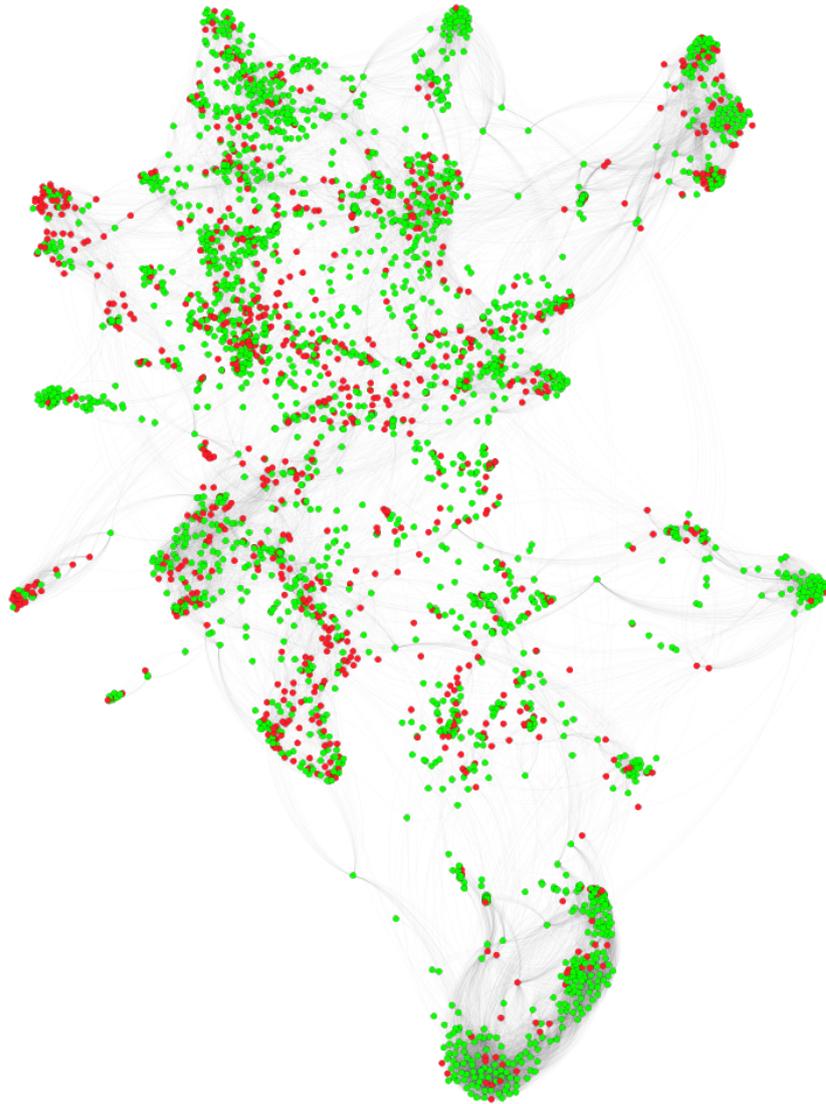


Figure 5.11: 20:00 type 1 snapshot on August 15h 2016. Green nodes are susceptible and red nodes are infected.

### 5.2.2. Type 2 network

Type 2 connections are made when two aircraft use the same runway at an airport. The same analysis as with Type 1 is done for this type of connection, again August 15th, 2016, is taken and snapshots are made every two hours from 08:00 to 22:00. Table 5.4 presents the network characteristics for the eight type 2 snapshots made on August 15th. The values show a significant change in the morning, and for the most part stabilize towards the afternoon and evening as can be seen in Figure 5.12.

Figure 5.12a shows the development of the number of nodes and edges throughout the day. The observation can be made that in the morning snapshots the number of nodes and edges is much lower than in the afternoon/evening snapshots.

Three degree distributions are presented Figure 5.13, for snapshots made at 08:00, 14:00 and 20:00. In this case, the distribution at 08:00 clearly has different shape when compared to the other two distributions. The 08:00 degree distribution can be considered a power-law distribution, but the other two follow a Poisson shape. This means that a type 2 network is a random graph in which small-world characteristics are not found. Table 5.5 shows the mean and standard deviation of the degree which can be

Table 5.4: Generating type 2 snapshot for several different times on 15-08-2016.

Time snapshot	Nodes	Edges	Avg. degree	Strong	Weak	% paths possible	Avg. shortest path	C
<b>08:00</b>	2,029	29,889	14.7	324	30	71.5%	4.8	0.53
<b>10:00</b>	2,942	96,689	32.9	208	6	89.0%	4	0.52
<b>12:00</b>	3,308	135,491	41	121	3	94.9%	3.3	0.41
<b>14:00</b>	3,416	158,711	46.5	113	1	95.6%	3.1	0.36
<b>16:00</b>	3,406	157,680	46.3	115	2	96.0%	3.1	0.35
<b>18:00</b>	3,433	154,719	45.1	109	1	96.6%	3.1	0.34
<b>20:00</b>	3,481	150,169	43.1	126	1	95.6%	3.2	0.34
<b>22:00</b>	3,549	173,023	48.8	75	1	97.6%	3.1	0.36

used to create a Poisson distribution. The 20:00 snapshot is plotted with the Poisson fit in Figure 5.14.

Table 5.5: Poisson distribution characteristics for type 2 snapshots on 15-08-2016.

Time snapshot	$\mu$	$\sigma$
08:00	29	29.7
10:00	66	64.6
12:00	82	61.1
14:00	93	65.7
16:00	93	63.6
18:00	90	58.7
20:00	86	57.1
22:00	98	66.9

This can be explained when looking closer to type 2 connections. An aircraft will generate more edges at larger airports (i.e. hubs) as there will be more other aircraft using the runways. But most aircraft will fly from hubs to smaller airports, thus generating an average amount of connections. The highest degree is expected for nodes which have just departed from a large airport, as the last 2 movements at that airport will be included in the snapshot.

The aircraft with the maximum amount of degree in the 20:00 snapshot is N850AS with 316. This aircraft operated by ExpressJet arrived in Atlanta at 18:42 and has just departed from Atlanta at 19:36. Assuming 1.5 minute separation between aircraft (conservative), 20 aircraft can arrive or depart per runway in 30 minutes. The theoretical maximum degree is then equal to  $400 = 20$  (aircraft per runway in 30 minutes) \* 5 (number of runways) \* 2 (edge in both directions) \* 2 (two movements). 316 is below this number, but keeping in mind that international flights are not included in this analysis this is most likely near the maximum.

The average shortest path length is plotted in Figure 5.12b along with the percentage of possible paths. As it can be seen, the average shortest path length goes down throughout the morning and stabilizes just above 4. The number of possible paths increases to above 95%. The diameter of the graph is equal to 9 and this occurs 52 times. One example where this diameter is found from N11544 to N14107. The first aircraft is operated by ExpressJet and is in-flight from Westchester County Airport in New York. The second aircraft is operated by United Airlines and is in-flight from Lihue Airport in Hawaii. This makes sense as both aircraft have just departed from small airports on different sides of the country, making the path as long as possible.

The number of connected components are presented in Figure 5.12c. Both the number of weak and strong components goes down throughout the morning and stabilizes in the afternoon. The number of strong components stays at around 100 while the number of weak components goes down to 1. This means that all of the nodes in the network are connected when directed edges are replaced with undirected edges.

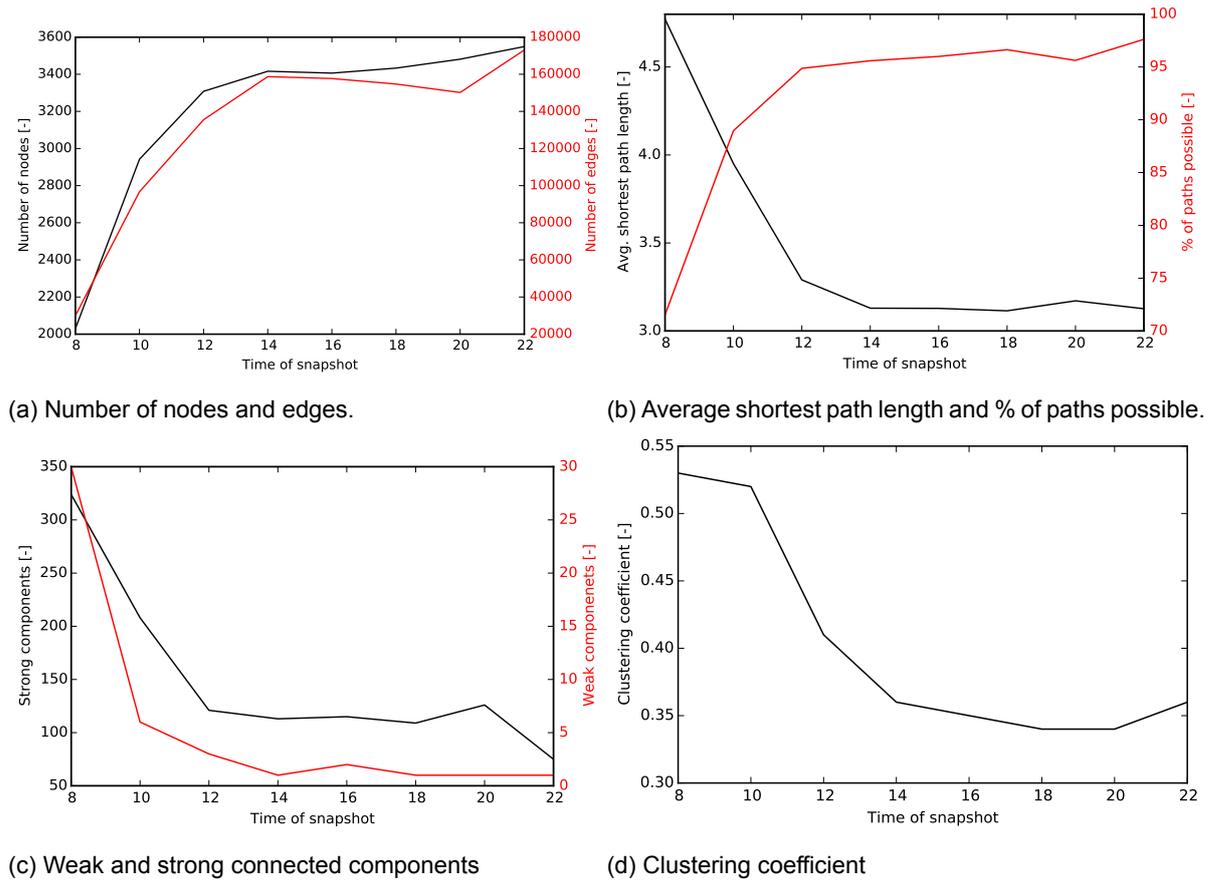


Figure 5.12: Development of network characteristics for snapshots on 15-08-2016.

If the snapshot at 20:00 is considered, the largest strongly connected component is 3,468 nodes. The other strongly connected components all consist of network with less than 3 nodes. This means that there is a large strong core, with small groups (1-3) of aircraft outside of this core. These outside groups are connected in one-way direction (either in or out) with the large core.

To find a stable representation of the network, a snapshot is made for the seven different days in August. The resulting mean, standard deviation and coefficient of variance of the network characteristics are given in Table 5.6. The largest coefficient of variation is found in the number of weakly connected components, but this number varies between 1 and 3. The number of strongly connected components is spread with a coefficient of variation of 22.5%, varying between 42 and 142. Going back to the explanation of this number of strongly connected components, it means that there are a varying number of aircraft outside of the large strongly connected core of the network. In general the coefficient of variation is small as it stays below the threshold of 1.

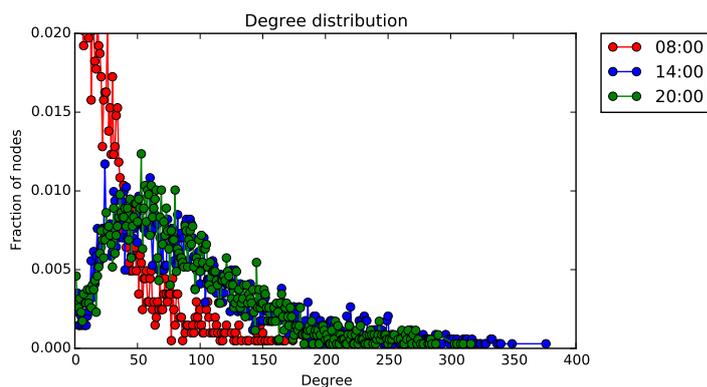


Figure 5.13: Degree distribution for three type 2 snapshots on 15-08-2016.

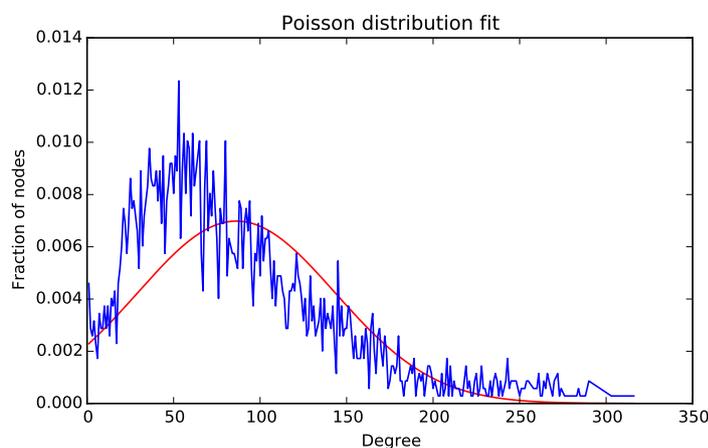


Figure 5.14: Degree distribution for 20:00 snapshot on 15-08-2016 and its Poisson fit.

Table 5.6: Generating type 2 snapshot for several different days in August 2016.

Characteristic	Mean	Standard dev.	Coefficient variation	Min	Max
<b>Nodes</b>	3,455	47	1.4%	3,378	3,564
<b>Edges</b>	159,565	11,761	7.4%	135,436	186,000
<b>Avg. degree</b>	46.2	3.1	6.7%	39.1	53.2
<b>Strong</b>	100	22	22.5%	42	142
<b>Weak</b>	1.25	0.52	41.6%	1	3
<b>% paths possible</b>	96.7%	0.8%	0.8%	95.6%	98.6%
<b>Avg. shortest path</b>	3.1	0.1	1.7%	3.0	3.3
<b>C</b>	0.35	0.01	3.0%	0.33	0.37

### 5.2.3. Type 3 network

The type 3 connections are generated when two aircraft are located at the same airport. The resulting network characteristics for different snapshots are given in Table 5.7 and they are plotted in Figure 5.15. Most of these values show a stabilization towards the end of the afternoon and evening. The node to edge ratio ends up at around 75, which means that nodes are connected to large groups of aircraft.

Figure 5.15a shows the development of the number of nodes and edges when creating a snapshot throughout August 15th. The snapshot at 7 A.M. is again the outlier, as the number of nodes stabilizes when considering the rest of the snapshot. The number of edges do still vary with around 10% in the afternoon snapshots.

Table 5.7: Generating type 3 snapshot for several different times on 15-08-2016.

Time snapshot	Nodes	Edges	Avg. degree	Strong	Weak	% paths possible	Avg. shortest path	C
08:00	3,601	111,786	31	1,967	30	34.0%	3.5	0.68
10:00	3,745	217,811	58.2	1,084	3	67.3%	3.2	0.56
12:00	3,752	266,503	71	685	1	80.3%	2.9	0.37
14:00	3,755	260,659	69.4	724	1	79.3%	2.9	0.34
16:00	3,755	243,369	64.8	741	1	78.9%	3.0	0.41
18:00	3,756	240,309	64	691	1	80.3%	3.0	0.4
20:00	3,756	230,813	61.5	684	1	80.0%	3.0	0.39
22:00	3,755	278,957	74.3	589	1	83.4%	2.9	0.41

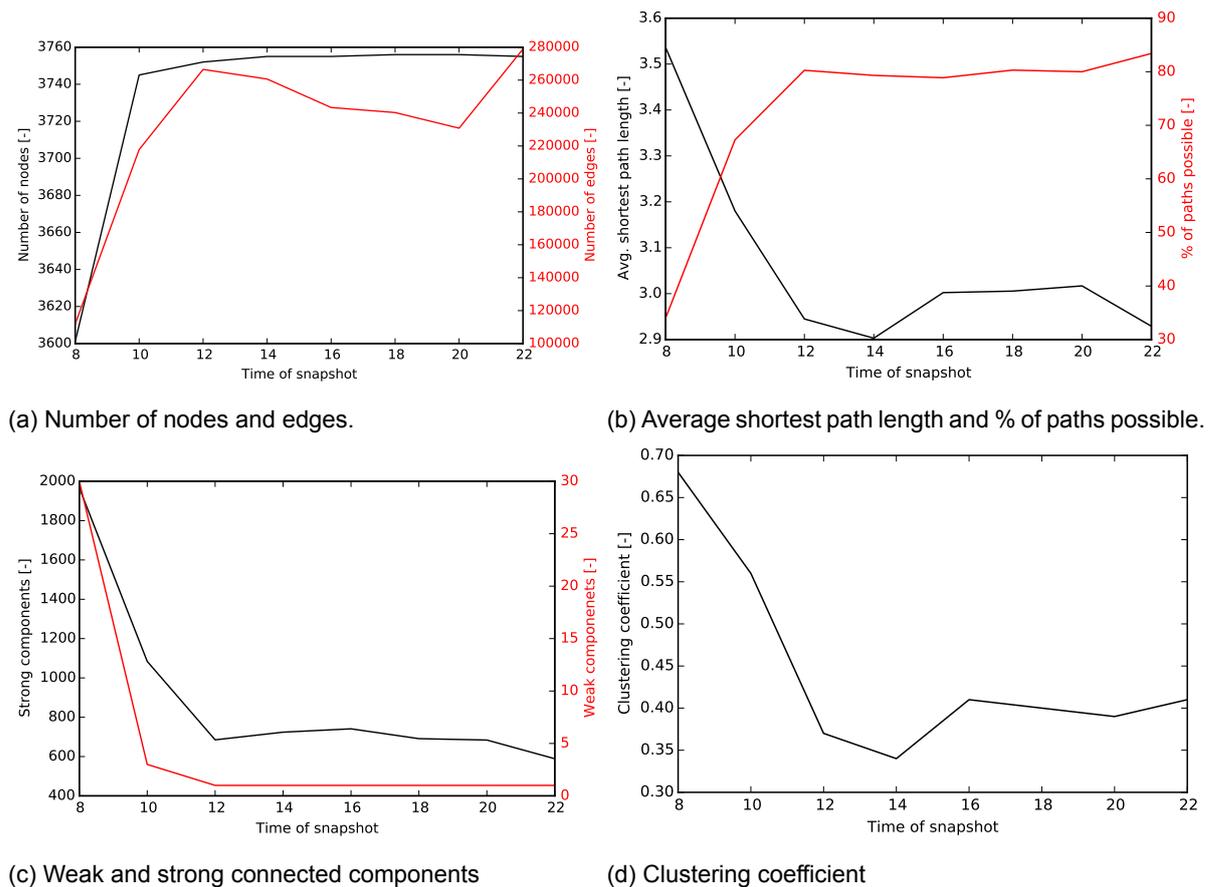


Figure 5.15: Development of network characteristics for type 3 snapshots on 15-08-2016.

Looking at the degree distributions in Figure 5.16, the 08:00 snapshot shows a very erratic pattern. The two degree distribution at 14:00 and 20:00 show similar Poisson shapes. For the 20:00 snapshot, the node with the highest degree is N349NW with 767. This aircraft from Delta Air Lines has just departed

from Atlanta, while not operating any other flights this day. This has provided ample opportunities to connect with other aircraft at the busiest airport of the US and thus seems a likely candidate for the node with the highest degree.

Table 5.8 shows the mean and standard deviation of the degree for each snapshot. These characteristics can be used to approximate a Poisson distribution, which is done for the 20:00 snapshot in Figure 5.17.

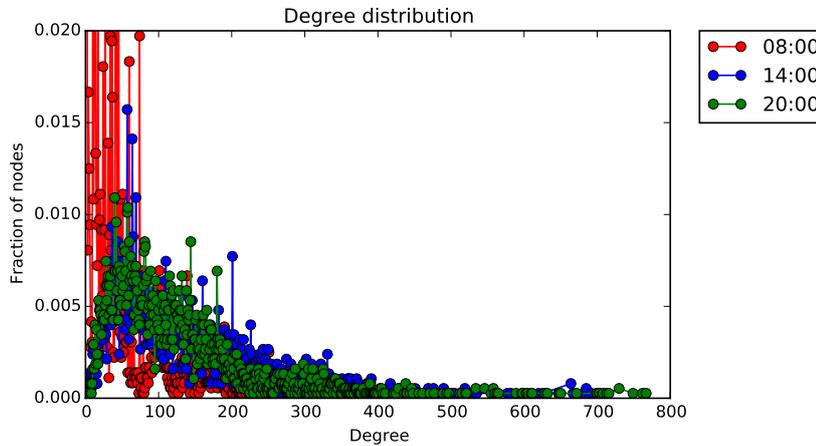


Figure 5.16: Degree distribution for three snapshots of Type 3 on 15-08-2016.

The average shortest path length development throughout the day is shown in Figure 5.15b. The percentage of possible paths goes up throughout the day, which follows from the increased number of edges. The average shortest path length goes down and stabilizes at around 4.

The diameter of the 20:00 snapshot is equal to 9 and this occurs 10 times. One example where this diameter is found is on the path from aircraft N551AS to N513AS. The first aircraft is N551AS, operated by Alaska Airlines, just departed from Omaha airport in Nebraska. N513AS is also operated by Alaska Airlines and has departed from Kahului airport in Hawaii at 16:10. This is a reasonable maximum distance, as these two aircraft have made connections at small airports in different parts of the country.

Table 5.8: Poisson distribution characteristics for type 3 snapshots on 15-08-2016.

Time snapshot	$\mu$	$\sigma$
08:00	62	62.6
10:00	116	94.6
12:00	142	97.5
14:00	139	101.4
16:00	130	97.2
18:00	128	96.8
20:00	123	98.0
22:00	149	117.1

The number of connected components is plotted in Figure 5.15c. The number of weakly connected components quickly reaches 1, which means that all nodes are connected if the direction of edges is neglected. The number of strongly connected components goes down to around 600-700, which means the fully connected core is surrounded by individual groups of aircraft with a one-directional link to the core.

Type 3 snapshots are now computed for different days in August 2016 (at 16:00, 18:00, 20:00 and 22:00). These times should lead to relatively stable networks, as the network characteristics discussed

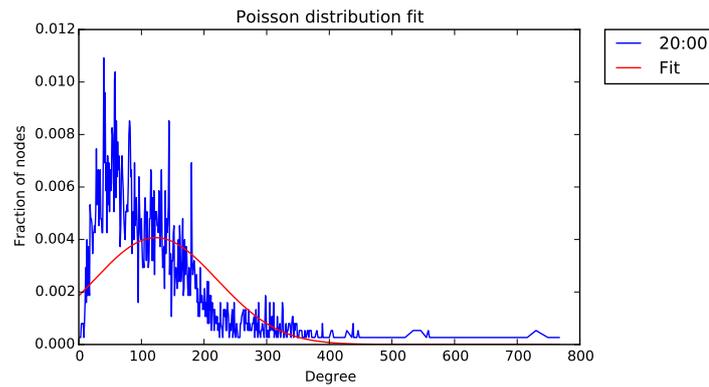


Figure 5.17: Degree distribution for 20:00 snapshot on 15-08-2016 and its Poisson fit.

before showed this behavior. Table 5.9 shows the mean, standard deviation and coefficient of variance of each of the network characteristics for these snapshots. The coefficient of variance is clearly very low, as it stays below 10%. The strongest variance is found in the strongly connected components, but the weakly connected components is always equal to 1. This means that the core of the network is fully connected, and depending on the number of aircraft not unidirectionally connected to this core the number of strongly connected components varies.

Table 5.9: Generating type 3 snapshot for several different days in August 2016.

Characteristic	Mean	Standard dev.	Coefficient variation	Min	Max
<b>Nodes</b>	3,734	29	0.8%	3,692	3,768
<b>Edges</b>	245,432	18,510	7.5%	218,837	284,454
<b>Avg. degree</b>	65.7	4.9	7.5%	58.4	75.5
<b>Strong</b>	661.6	66.6	10.1%	493	765
<b>Weak</b>	1	0	0.0%	1	1
<b>% paths possible</b>	80.8%	2.0%	2.5%	77.9%	86.1%
<b>Avg. shortest path</b>	3.0	0.0	1.7%	2.9	3.1
<b>C</b>	0.39	0.03	7.0%	0.32	0.41

#### 5.2.4. Comparison

The networks created by the three types of connections have been discussed. In this subsection the different networks will be compared, to see if the characteristics show any similarities or major differences.

Table 5.10 shows an overview of the mean value of the network characteristics as they were presented in the previous three subsections. To recap, these values are based on snapshots of seven different days in August 2016 at 16:00, 18:00, 20:00 and 22:00.

When looking at the number of nodes, the difference between type 1 and type 3 is around 700 aircraft. This means that the type 3 network is around 20-25% larger than the type 1 network in terms of nodes. The type 2 network falls a bit in between these two types in terms of number of nodes. The number of edges shows the same behavior as the number of nodes, except the magnitude is much larger. In terms of number of edges, a type 3 network is around 14.5 times larger than a type 1 network. Again the type 2 network falls in between, around 10 times larger than a type 1 network.

In terms of the average shortest path length, the type 2 and 3 networks show similar behavior. This average length is equal to around 4, while for the type 1 network it is more than 9. Also the number of

Table 5.10: Mean value of network characteristics for different connection types.

Characteristic	Type 1	Type 2	Type 3
<b>Nodes</b>	3,050	3,455	3,734
<b>Edges</b>	48,860	159,565	245,432
<b>Avg. degree</b>	16	46	66
<b>Strong</b>	1,292	100	662
<b>Weak</b>	20	1	1
<b>% paths possible</b>	30.4%	96.7%	80.8%
<b>Avg. shortest path</b>	5.3	3.1	3.0
<b>Avg. clustering</b>	0.39	0.35	0.39

paths possible between nodes is significantly larger for type 2 and 3 networks. In type 1 network, only around 20% of the paths are possible.

Finally the number of connected components is also similar for type 2 and 3 networks. Weakly connected components is equal to 1, which means that all of the nodes are connected if the direction of edges is not considered. Type 1 network has an average of 44 weakly connected components, which makes it a fragmented network. This can also be seen in the number of paths which are possible between nodes.

Table 5.11: Shape of degree distribution for the different connection types.

Connection type	Degree	Characteristics
Type 1	Power-law	$\bar{\gamma} =$
Type 2	Poisson	$\bar{\mu} = 92, \bar{\sigma} = 62.4$
Type 3	Poisson	$\bar{\mu} = 134, \bar{\sigma} = 102.1$

Table 5.11 shows an overview of the shape of the degree distribution for each type of connection. A power-law distribution means there are a few nodes which are highly connected, while the majority of nodes only have a few edges. The Poisson distribution is where there are nodes with a small number of edges and high number of edges, but the majority are somewhere in between. The type 1 networks show this power-law distribution for both the in and out degree, while the type 2 networks show the Poisson distribution for the in and out degree. Peculiarly, the the type 3 networks show a different distribution for in and out degree. Most of the aircraft have a low in-degree but a higher out-degree.

### 5.3. Delay propagation

In this section, the delay analysis will be carried out on the networks discussed in Section 5.2. When looking at the delay patterns presented in Figure 5.4, the general pattern is the same but the magnitude is different. For this reason, the analysis is done for a day with average delay (August 15th, 2016) in Section 5.3.1. The results are then compared to other days with different magnitudes of delay in Section 5.3.2.

#### 5.3.1. Neighbor analysis throughout an average day

In this subsection an analysis will be done of the neighbors within the different networks during the time of the day, in this case 15-08-2016. For each individual aircraft in the network, the percentage of infected neighbors is determined. The group is then split into infected and susceptible aircraft to determine if the probability of coming into contact with infected aircraft, increases the probability of being infected. The sampling method that will be used will be random relative sampling with 100 total samples. This process will be repeated until a stable results is achieved in terms of a conclusion.

**Type 1** - For the type 1 snapshots, the number of 1st degree neighbors and the percentage of those neighbors which are infected are given in Table 5.12. As this is calculated for each node separately, a median ( $Md$ ), standard deviation ( $\sigma$ ) and interquartile range (IQR) are given. The IQR is the range within which the middle 50% of the data lies. This data is given for the four snapshots considering all nodes, the susceptible nodes and the infected nodes. The number of infected nodes increases between 16:00 (24% infected) and 22:00 (33% infected).

Table 5.12: Number of 1st degree neighbors and percentage of neighbors infected for type 1 snapshots.

Time	Nodes	Number	Number of neighbors			Percentage infected neighbors		
			$Md$	$\sigma$	IQR	$Md$	$\sigma$	IQR
16:00	All	2,059	19	21.1	7.0, 29.0	16.7%	18.3%	6.7%, 27.5%
	Susceptible	1,566	19	21.5	6.0, 30.0	15.9%	16.8%	6.3%, 25.0%
	Infected	493	19	19.8	8.0, 28.0	22.2%	21.1%	12.5%, 36.4%
18:00	All	2,076	18	21.6	7.0, 29.0	20.8%	19.5%	11.3%, 33.3%
	Susceptible	1,465	17	22.2	6.0, 29.0	19.0%	18.3%	9.5%, 33.3%
	Infected	611	20	20.2	8.0, 28.0	28.6%	20.8%	16.7%, 39.3%
20:00	All	2,012	16	19.6	6.0, 26.3	23.3%	21.5%	10.5%, 37.3%
	Susceptible	1,390	16	20.9	6.0, 28.0	20.0%	19.1%	7.7%, 32.1%
	Infected	622	17	16.4	7.0, 26.0	31.6%	23.9%	18.2%, 50.0%
22:00	All	2,283	18	23.1	8.0, 38.0	25.0%	22.9%	13.6%, 40.0%
	Susceptible	1,522	18	24.8	8.0, 42.0	20.0%	18.3%	11.8%, 33.3%
	Infected	761	18	18.7	7.0, 28.0	35.7%	26.7%	23.1%, 55.9%

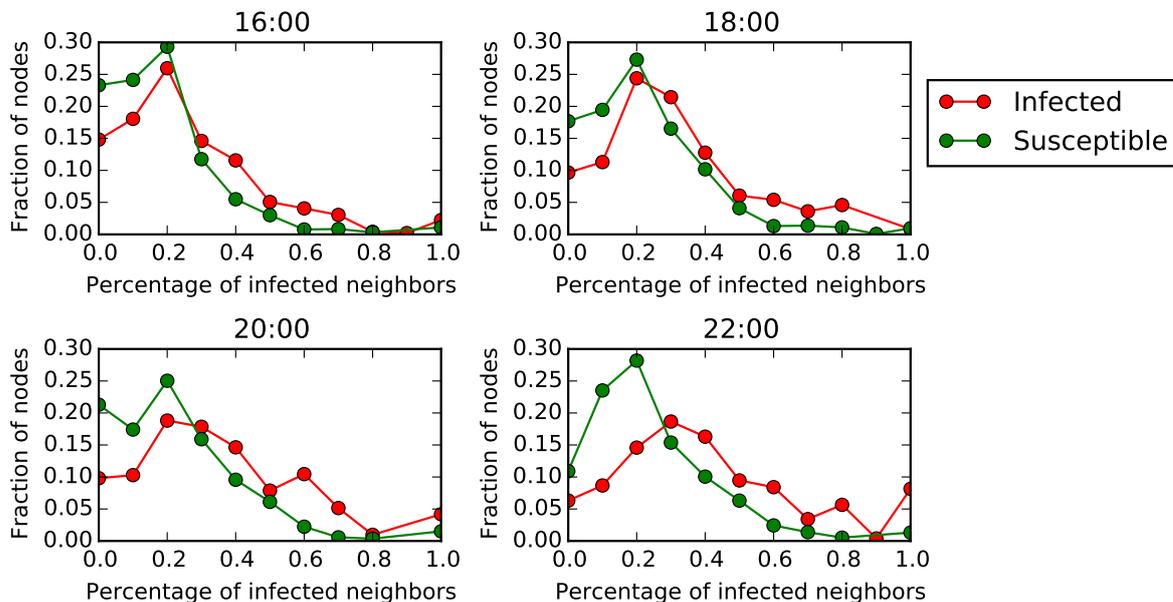


Figure 5.18: The distribution of the percentage of infected 1st degree neighbors of susceptible and infected nodes.

The number of neighbors throughout these snapshots stays relatively constant, with the median varying between 16 and 20. The standard deviation is in the order of the median, which means the coefficient of variation is around 1. This means that the spread of this mean is relatively high. The median lies in the middle of the IQR for the first three snapshots, but moves a bit towards the bottom for the 22:00 snapshot. This means there is more spread in the higher number of neighbors.

The percentage of neighbors which are infected can be seen to increase throughout the time of day,

which is consistent with the increase in number of infected aircraft. The standard deviation of these medians is again in the order of the median, which means that there is a high spread in values.

To visualize the medians of these percentages of infected neighbors, a distribution is made in Figure 5.18. This percentage distribution shows how many nodes have a certain percentage of infected neighbors. The nodes are spread over 10 bins by rounding off all of the percentages to the nearest 10%. This is done for the susceptible and infected nodes individually, which will create a possibility to compare the two distributions. The shape of the two follow the same approximate distribution. The general conclusion is that there are relatively more susceptible nodes with 0-20% infected neighbors and relatively more infected nodes with 50-100%. Along with the fact that there are a lot more susceptible nodes with (rounded) 0% infected neighbors, the median for the infected nodes is higher.

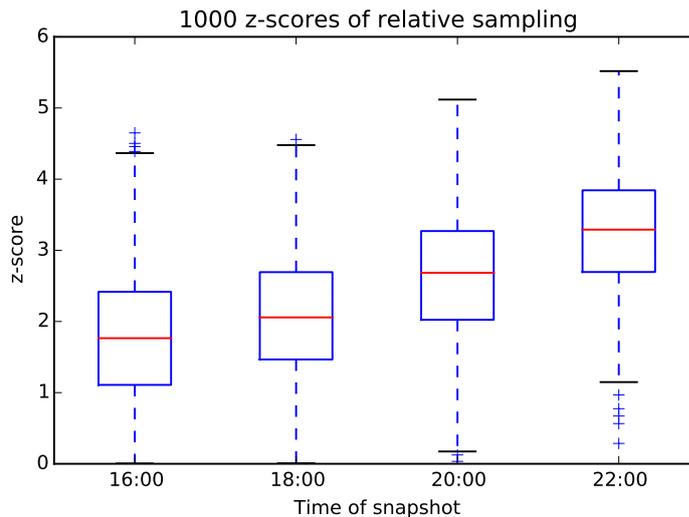


Figure 5.19: Boxplots of the z-score results from the Mann-Whitney U tests for 1st degree type 1 neighbors.

Using the relative random sampling technique, 100 samples will be chosen from the population 1000 times. A Mann-Whitney U test will be performed to find if there is a significant difference between the susceptible population and the infected population. The calculated z-scores are given in Figure 5.19. If a significance level of 95% is assumed and the number of samples is large enough, then the critical z-score is 1.96. This critical z-score defines whether the null hypothesis is accepted or rejected according to the condition given in Equation (4.11).

Figure 5.19 shows that the median z-scores are smaller than this 1.96 for the snapshots at 16:00, 18:00 and 20:00 but is greater than this value for the snapshot at 22:00. It also clear that there are a significant part of the data points which fall outside of the 1.5 times the IQR region (which are plotted as a + above the top whisker). This means that there is a high variability in taking random samples from the population. As the Z-score can not be negative, the bottom of the boxplot lies on the x-axis.

In Table 5.13 the results from the Mann-Whitney U tests are summarized.  $n_i$  and  $n_s$  represent the number of samples from the infected population and the susceptible population. The median values for the percentage of infected ( $Md_i$ ) and susceptible ( $Md_s$ ) first degree neighbors are taken from Table 5.12.  $Md_U$  is the median value for the smallest U value and  $Md_Z$  is the median calculated Z-value. These values are the result of the Mann-Whitney U test and are needed to completely report the result.

In three of the four snapshots, less than 50% of the samples accept the null hypothesis. The formal result for this case is:

**The median percentage of infected first degree neighbors (in the 22:00 snapshot on 15-08-2016) for the infected and susceptible population is 33% and 22% respectively. The infected group is significantly larger than the susceptible group (Mann Whitney,  $U = 813$ ,  $n_i =$**

**33,  $n_s = 67$ ,  $P < 0.05$  one-tailed).**

Table 5.13: Z-scores calculated using the Mann-Whitney U test for first degree neighbors in type 1 snapshots.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
16:00	23	0.22%	77	0.16%	671	1.77	58.2%
18:00	29	0.29%	71	0.19%	760	2.06	45.2%
20:00	30	0.32%	70	0.20%	694	2.68	23.2%
22:00	33	0.36%	67	0.20%	657	3.29	6.3%

This same process can be repeated for the second degree neighbors of individual aircraft. This is done by taking the neighbors of the first degree neighbors, excluding the aircraft that are already a first degree neighbor. The table with median values and figure with the distributions can be found in Appendix B.1 in Table B.1 and Figure B.1 respectively. The resulting Mann Whitney U tests are summarized in Table 5.14, while the boxplot with z-scores is given in Figure B.2. Like the case with the first degree neighbors, there is sufficient evidence to prove that the second degree neighbors are affected in all snapshots. The null hypothesis is accepted in less than 30% of the samples for each snapshot.

Table 5.14: Z-scores calculated using the Mann-Whitney U test for second degree neighbors in type 1 snapshots.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
16:00	23	0.24%	77	0.17%	589	2.44	30.2%
18:00	29	0.29%	71	0.20%	673	2.71	21.1%
20:00	30	0.30%	70	0.20%	652	3	12.1%
22:00	33	0.33%	67	0.24%	675	3.16	8.0%

For the third degree, the same process is repeated. The resulting number of neighbors and percentage of infected neighbors are provided in Table B.2, along with the distributions in Figure B.3. The results of the Mann-Whitney U tests are given in Table 5.15, while the boxplots for the resulting z-scores are given in Figure B.4. The null hypothesis is accepted in every snapshot, along with the similar medians leads to the conclusion that there is no third degree effect.

Table 5.15: Z-scores calculated using the Mann-Whitney U test for third degree neighbors in type 1 snapshots.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
16:00	23	0.22%	77	0.20%	751	1.11	82.6%
18:00	29	0.26%	71	0.23%	886	1.09	81.4%
20:00	30	0.27%	70	0.23%	860	1.43	71.4%
22:00	32	0.30%	68	0.26%	837	1.86	53.8%

In Figure 5.20 the median values are summarized in a bar chart for each snapshot. The infected nodes show a significantly higher median value up until the second degree. This is in line with the results from the performed Mann-Whitney U test.

Each node has an absolute number of infected neighbors. The hypothesis is that more infected neighbors leads to a higher probability of a node being infected. In Figure 5.21 the minimum number of infected first degree neighbors is plotted versus the percentage of nodes infected. The horizontal line represents the percentage of infected nodes in the system, which is the starting point for the graph (i.e., all nodes have at least 0 infected neighbors). This can be translated to a conditional probability: what is the probability a node is infected given it has at least  $n$  infected neighbors?

This conditional probability increasing if the number of infected neighbors is higher, based on the trend

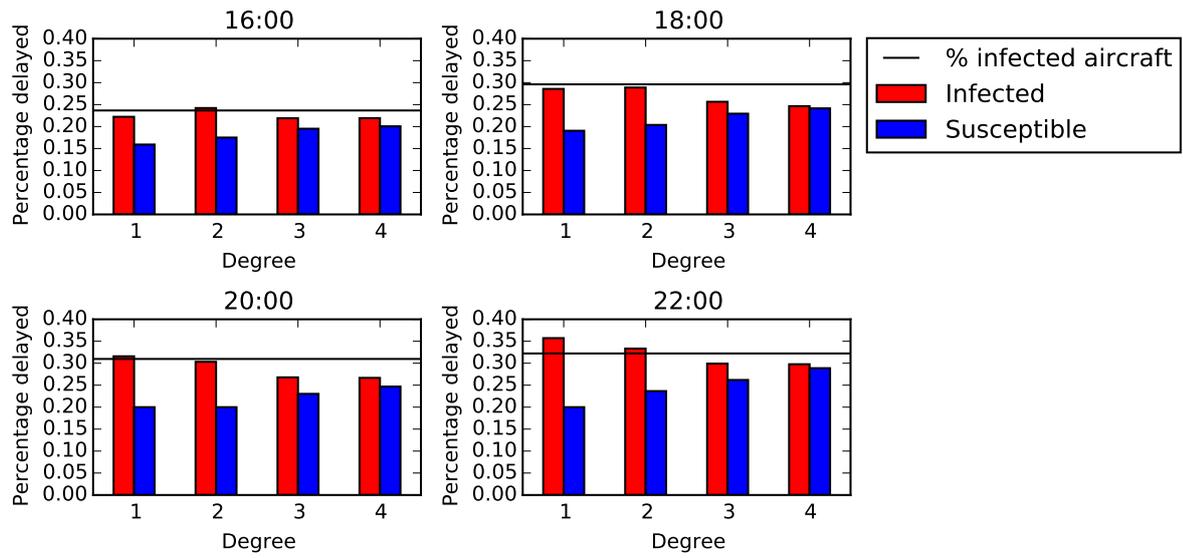


Figure 5.20: Bar chart summarizing median values for percentage of infected neighbors in type 1 networks.

of the line going up. This conclusion can also be made for the second degree effects, but when the graph is made for third degree effects it flattens out. These graphs can be found in Appendix B.2.

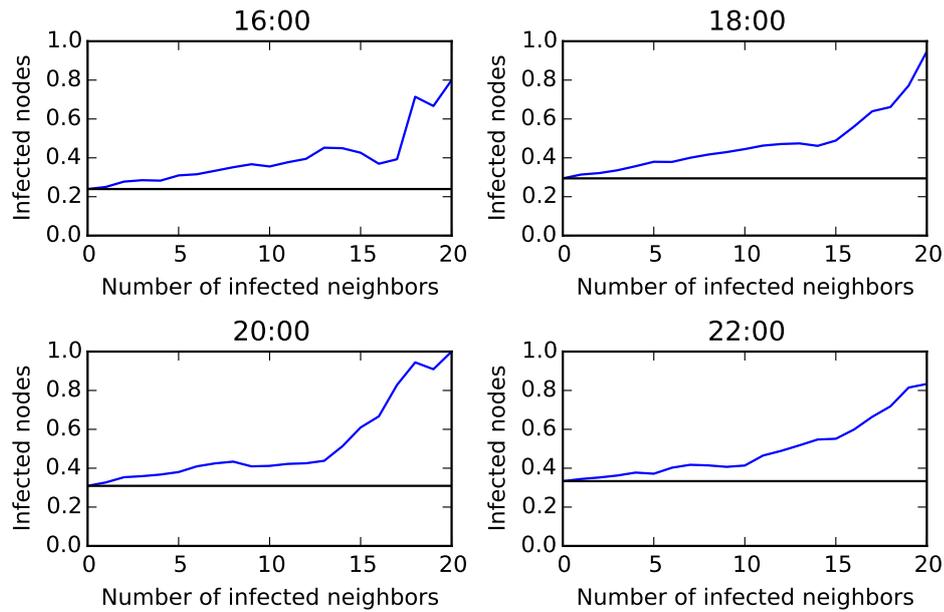


Figure 5.21: Minimum number of infected first degree neighbors versus percentage of infected nodes for type 1 snapshots.

**Type 2** - For the type 2 snapshots the number of neighbors and the percentage of infected neighbors is also calculated. Considering first degree neighbors, the resulting median values can be found in Table 5.16. The total number of nodes is between 3,300 and 3,500, while the percentage infected goes up from 22% at 16:00 to 34% at 22:00. The median number of first degree neighbors lies between 35 and 39, with the IQR varying between 20 and 70. That means about 25% of the nodes have more than 70 first degree neighbors.

The median of the percentage of infected first degree neighbors grows from 19.4% at 16:00 to 28.1% at 22:00. When only considering the susceptible nodes the median declines, while it increases when only the infected nodes are considered. The difference between these two groups can be seen to increase from 16:00 to 22:00.

Table 5.16: Number of 1st degree neighbors and percentage of neighbors infected for type 2 snapshots.

Time	Nodes	Number	Number of neighbors			Percentage infected		
			<i>Md</i>	$\sigma$	IQR	<i>Md</i>	$\sigma$	IQR
16:00	All	3,313	37	33.2	20.0, 60.0	19.4%	15.0%	12.1%, 27.6%
	Susceptible	2,571	38	34.4	20.0, 62.0	18.2%	13.8%	11.1%, 25.6%
	Infected	742	35	28.7	20.0, 56.0	23.7%	17.4%	16.5%, 33.6%
18:00	All	3,339	38	31.7	19.0, 59.0	24.5%	15.2%	16.7%, 33.3%
	Susceptible	2,434	37	32.8	18.0, 59.0	22.8%	14.1%	15.7%, 31.3%
	Infected	905	39	28.4	22.0, 60.0	29.1%	16.8%	20.7%, 40.0%
20:00	All	3,379	35	31.0	18.0, 59.0	27.0%	17.3%	16.7%, 37.5%
	Susceptible	2,366	36	32.4	18.0, 61.0	23.8%	15.3%	14.3%, 33.8%
	Infected	1,013	34	27.2	19.0, 54.0	33.7%	19.3%	23.4%, 46.4%
22:00	All	3,492	39	32.7	21.0, 67.0	28.1%	19.6%	18.2%, 42.2%
	Susceptible	2,305	43	34.3	22.0, 70.0	24.1%	16.1%	16.1%, 35.3%
	Infected	1,187	33	27.8	19.0, 54.0	37.9%	22.2%	25.5%, 54.2%

The percentages of infected first degree neighbors can be translated into a distribution, which is given

in Figure 5.22 for every snapshot. The distribution for the susceptible nodes stays relatively constant, while the distribution for the infected nodes spreads itself across the x-axis. Clearly, both of these percentage distributions follow a skewed Poisson shape.

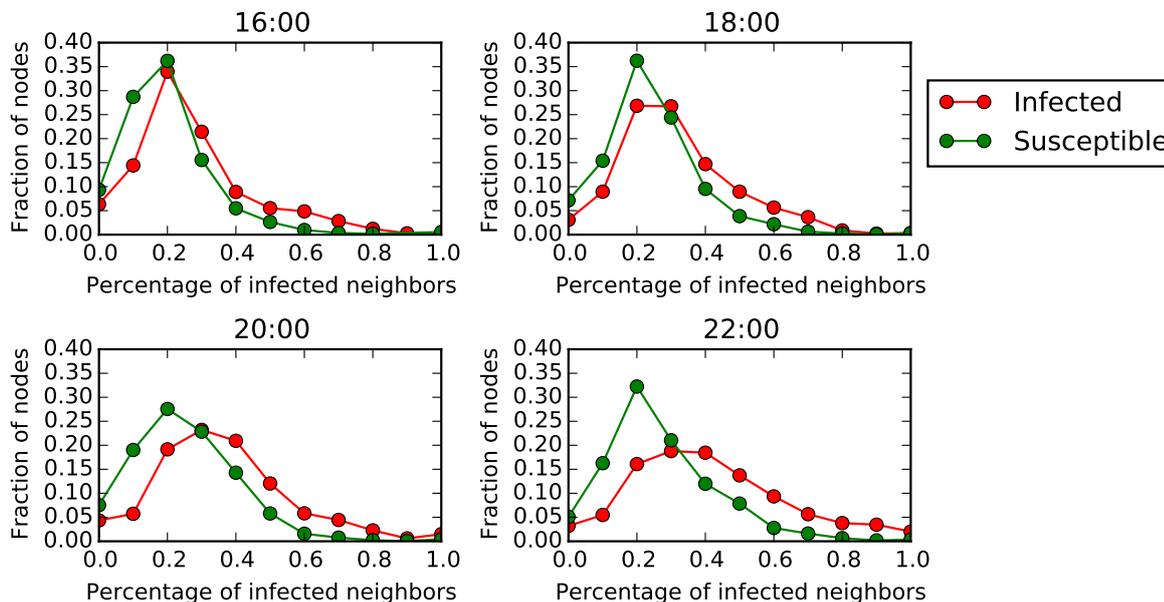


Figure 5.22: The distribution of the percentage of infected 1st degree neighbors of susceptible and infected nodes.

The Mann-Whitney U test is once again applied to the data and the resulting z-scores are given in Figure 5.23. Relative sampling is used and performed 1000 times to achieve a relatively stable result. The difference between the median z-scores is further elaborated in Table 5.17. In this table, the resulting decision on the null hypothesis is also given. For every case, the median percentage of infected first degree neighbors is significantly higher for the infected group when compared to the susceptible group.

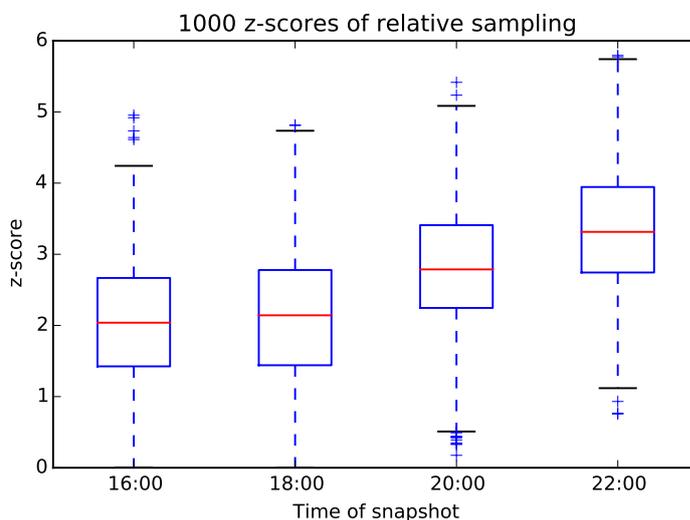


Figure 5.23: Boxplots of the z-score results from the Mann-Whitney U tests for 1st degree type 2 neighbors.

The process is repeated for the second degree neighbors in the type 2 snapshots. The resulting medians, percentage distributions and boxplot of resulting z-scores are given in Appendix B.3. The results of the Mann-Whitney U tests are summarized in Table 5.18. The null hypothesis is rejected for three of the four snapshots and the median values also show little variation between them. Further inspection of the percentage distributions in Figure B.8 leads to the conclusion that there are no significant second

Table 5.17: Z-scores calculated using the Mann-Whitney U test for first degree neighbors in type 2 snapshots.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
16:00	22	0.24%	78	0.18%	613	2.04	45.7%
18:00	27	0.29%	73	0.23%	710	2.14	43.7%
20:00	29	0.34%	71	0.24%	663	2.79	16.6%
22:00	33	0.38%	67	0.24%	654	3.31	6.5%

degree effects in this case.

Table 5.18: Z-scores calculated using the Mann-Whitney U test for second degree neighbors in type 2 snapshots.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
16:00	22	0.22%	78	0.20%	633	1.88	53.1%
18:00	27	0.27%	73	0.26%	785	1.57	65.4%
20:00	29	0.29%	71	0.28%	790	1.82	56.3%
22:00	34	0.33%	66	0.30%	849	1.99	49.4%

In Figure 5.24 the median values are summarized in a bar chart. The first degree effects are clearly visible as the bars from the infected and susceptible nodes show clear differences. From the second degree onward these bars have similar magnitudes which is the same conclusion as from the Mann-Whitney U tests.

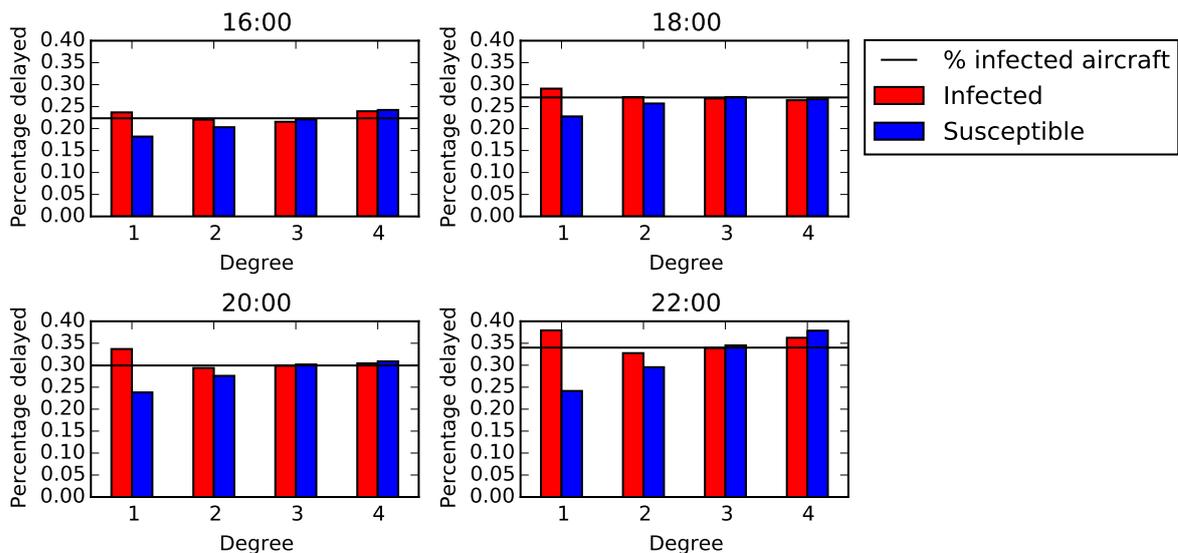


Figure 5.24: Bar chart summarizing median values for percentage of infected neighbors in type 2 networks.

However, in Figure 5.25 the probability of being infected as a function of the percentage of infected first degree neighbors is plotted for the four snapshots. A clear trend can be found moving upwards if the percentage of infected first degree neighbors goes up. In other words, the probability of an aircraft being infected increases if the number neighbors infected increases. This trend can also be found, but less clear, in the second degree neighbors, but disappears when considering the third degree neighbors. These graphs can be found in Appendix B.4.

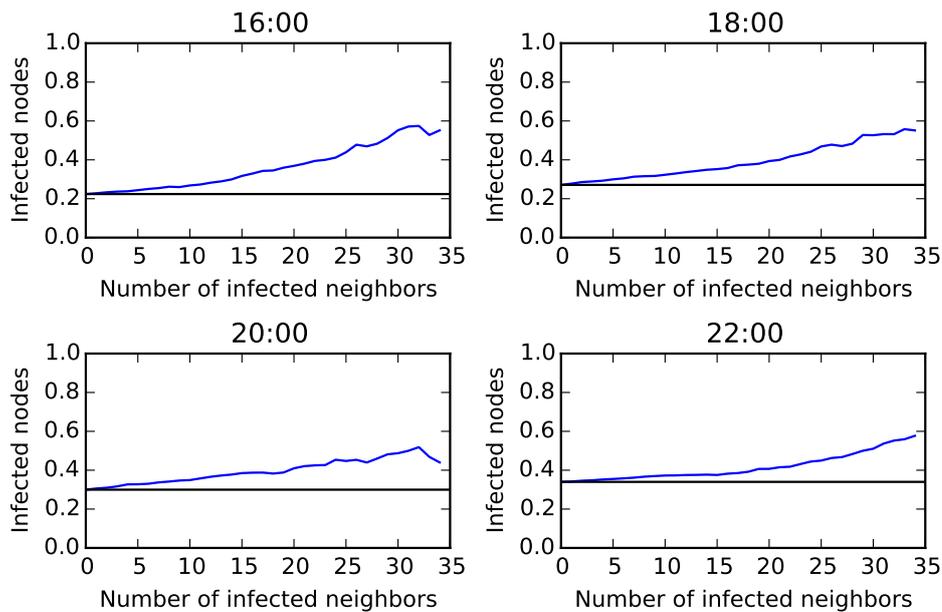


Figure 5.25: Minimum number of infected first degree neighbors versus percentage of infected nodes for type 2 snapshots.

**Type 3 -** For type 3 snapshot, the number of first degree neighbors and the percentage of infected first degree neighbors are given in Table 5.19. The number of nodes included in the network are between 3,000 and 3,200, the percentage of them that are infected increases from 23% at 16:00 to 34% at 22:00. The median number of first degree neighbors varies between 53 and 66 and the median percentage of infected first degree neighbors increases from 18.4% to 28.9%.

Table 5.19: Number of 1st degree neighbors and percentage of neighbors infected for type 3 snapshots.

Time	Nodes	Number	Number of neighbors			Percentage infected		
			<i>Md</i>	$\sigma$	IQR	<i>Md</i>	$\sigma$	IQR
16:00	All	3,061	60	76.3	20.0, 111.0	18.4%	13.3%	12.0%, 25.0%
	Susceptible	2,365	56	72.6	18.0, 107.0	17.5%	13.2%	11.4%, 24.2%
	Infected	696	79	85.8	33.8, 126.3	20.9%	13.3%	14.3%, 28.3%
18:00	All	3,103	58	77.8	19.0, 101.0	21.4%	13.6%	15.7%, 30.0%
	Susceptible	2,247	52	74.1	16.0, 94.0	20.2%	13.2%	15.3%, 28.3%
	Infected	856	80	84.5	32.0, 121.3	24.8%	13.9%	16.7%, 33.4%
20:00	All	3,123	53	81.6	17.0, 98.0	25.0%	16.8%	15.4%, 33.3%
	Susceptible	2,166	49	81.4	15.0, 94.0	23.1%	15.8%	13.3%, 31.6%
	Infected	957	62	81.6	22.0, 106.0	28.6%	17.8%	21.1%, 40.0%
22:00	All	3,191	66	90.1	24.0, 112.0	28.9%	17.4%	18.2%, 39.8%
	Susceptible	2,102	67	91.4	22.0, 112.0	25.8%	16.0%	16.5%, 36.7%
	Infected	1,089	66	87.5	27.0, 111.0	36.0%	18.0%	25.0%, 46.0%

The percentages of infected first degree neighbors can be visualized in a distribution. In Figure 5.26 this is plotted for the different snapshots, grouped by susceptible and infected nodes.

The resulting percentages can then be fed into the Mann-Whitney U test to determine the significance. From these tests, z-score follow which are compared with the critical z-value. They are found in a boxplot in Figure 5.27, with the median z-scores of 20:00 and 22:00 passing this critical value. The results of the Mann-Whitney U test are further specified in Table 5.20. The null hypothesis is rejected

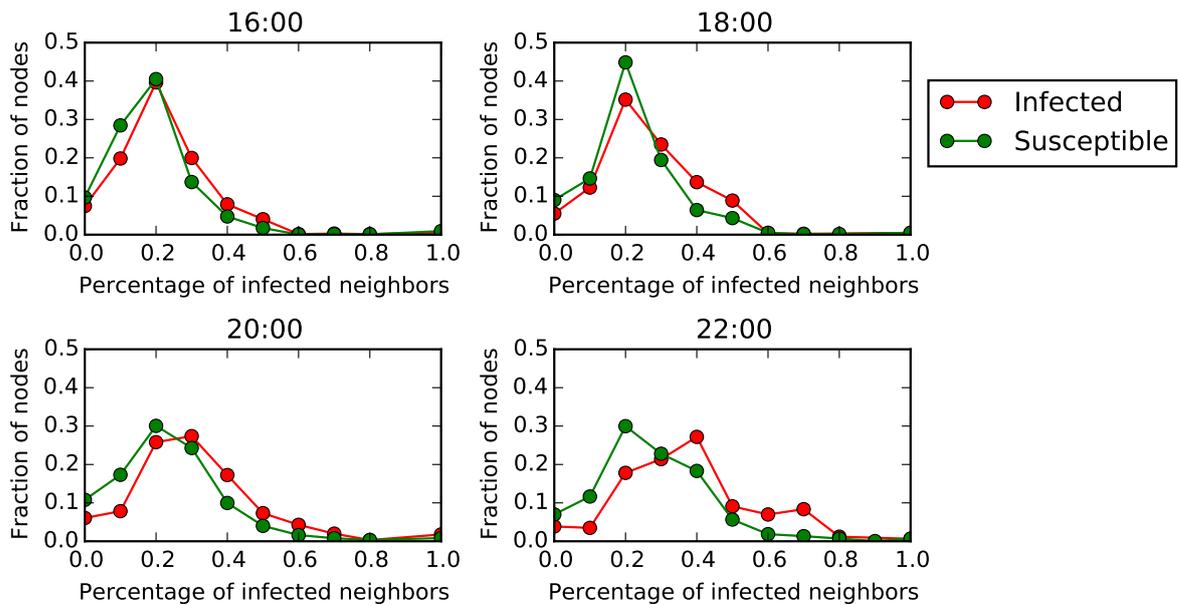


Figure 5.26: The distribution of the percentage of infected 1st degree neighbors of susceptible and infected nodes.

for the snapshots at 20:00 and 22:00.

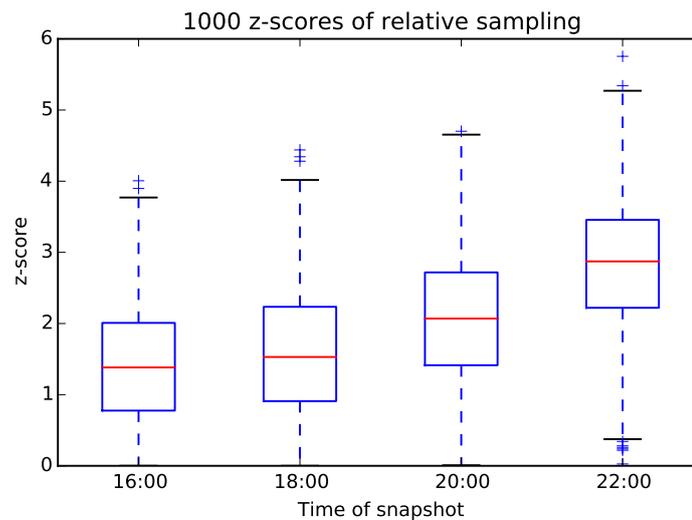


Figure 5.27: Boxplots of the z-score results from the Mann-Whitney U tests for 1st degree type 3 neighbors.

This process is repeated for the percentages of infected second degree neighbors. The specific results are given in Appendix B.5 and the summarized results from the Mann-Whitney U test in Table 5.21. The null hypothesis is accepted for all of the cases, and further research into the distributions (Figure B.12) strengthens the conclusion that there are no second degree effects.

The median values are compared in Figure 5.28 for each snapshot. Consistent with the conclusions from the Mann-Whitney U tests, the bars for susceptible and infected nodes are similar from the second degree onwards. This indicates that there is no difference.

Finally, an analysis is done on the probability of being infected as a function of the percentage of neighbors infected. Figure 5.29 shows this effect for the first degree neighbors. A slight trend can be found in the 20:00 and 22:00 snapshots going upwards. This effect can not be found in the second and

Table 5.20: Z-scores calculated using the Mann-Whitney U test for first degree neighbors in type 3 snapshots.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
16:00	22	0.21%	78	0.18%	692	1.38	73.2%
18:00	27	0.25%	73	0.20%	789	1.53	66.8%
20:00	30	0.29%	70	0.23%	775	2.07	45.9%
22:00	34	0.36%	66	0.26%	728	2.87	17.3%

Table 5.21: Z-scores calculated using the Mann-Whitney U test for second degree neighbors in type 3 snapshots.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
16:00	22	0.20%	78	0.19%	700	1.32	76.3%
18:00	27	0.24%	73	0.23%	817	1.31	75.0%
20:00	30	0.27%	70	0.26%	838	1.6	64.8%
22:00	34	0.31%	66	0.29%	886	1.72	60.5%

third degree neighbors, which can be found in Appendix B.6.

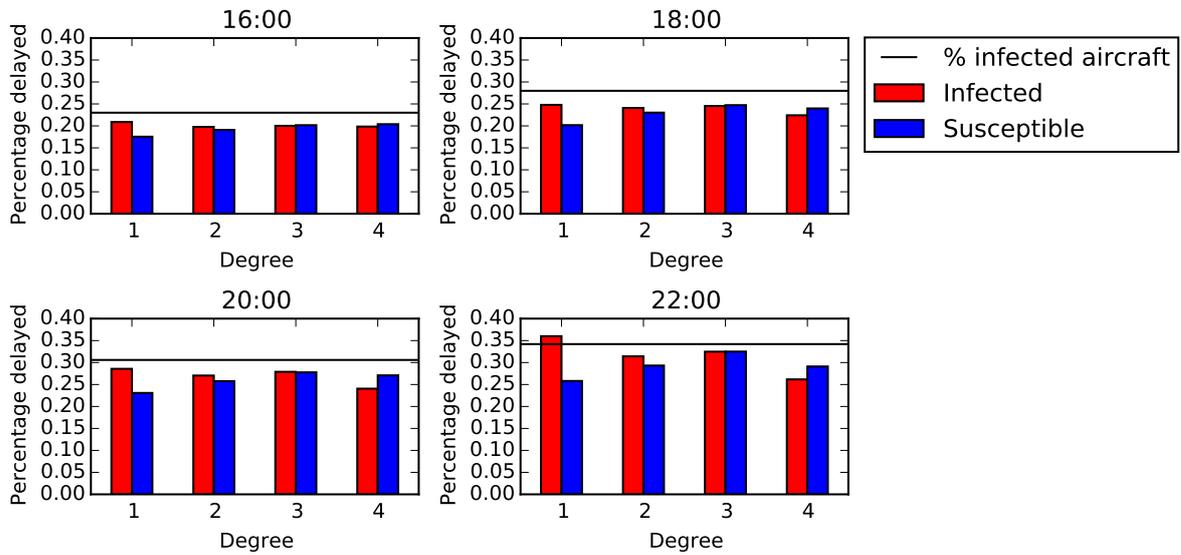


Figure 5.28: Bar chart summarizing median values for percentage of infected neighbors in type 3 networks.

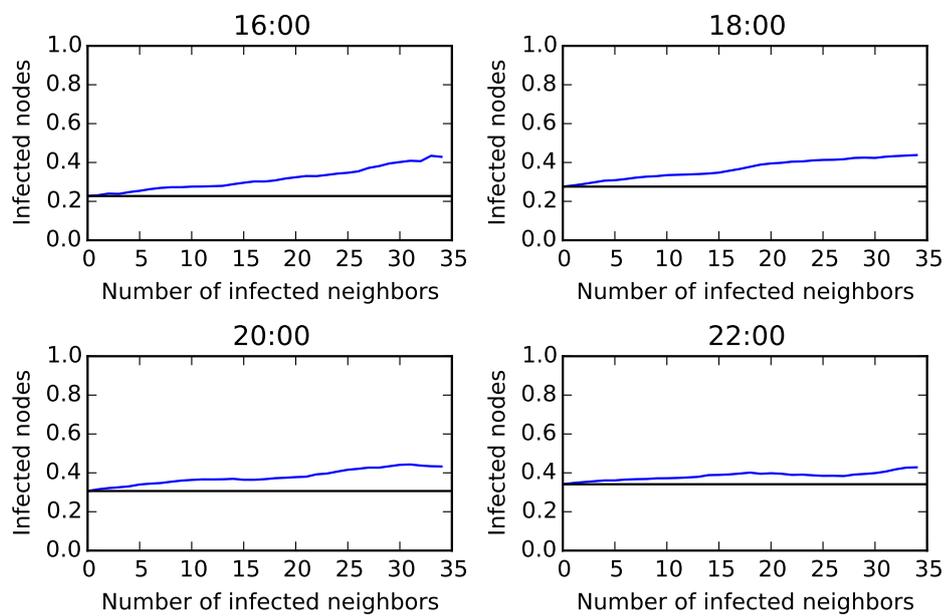


Figure 5.29: Minimum number of infected first degree neighbors versus percentage of infected nodes for type 3 snapshots.

### 5.3.2. Comparison with other days

The previous subsection analyzed one day with average total delay, August 15th. The analysis can also be done for various days with different delay magnitudes. The results for the Mann-Whitney U tests are given in Appendix B.7 to Appendix B.9.

Table 5.22 shows an overview of the results per type of connection and the the different days. As the Mann-Whitney U tests are performed 3000 times, the percentage of passed tests is measured. A check mark in Table 5.22 means that for at least one of the snapshots on that day (16:00, 18:00, 20:00, 22:00), the null hypothesis was rejected at least 50% of the times (i.e., in the majority of the tests the null hypothesis is rejected). If there is no check mark, a majority of the Mann-Whitney U tests had accepted the null hypothesis.

The magnitude of total arrival delay is also given. A few days can be highlighted to have low delay (August 23rd), average delay (August 15th), high delay (August 19th) and abnormal delay caused by special circumstances (August 8th). The effects follow these magnitudes of delay.

For instance, when considering the type 2 connections:

- Low delay: No effect
- Average delay: First degree effects
- High and abnormal delay: Second degree effects

Table 5.22: Summarized results for different days of August.

Type	Day	Arrival delay [min]	1st degree	2nd degree	3rd degree	4th degree
Type 1	02-08	318,529	✓	✓		
	07-08	204,827		✓		
	08-08	605,141	✓	✓	✓	✓
	15-08	265,257	✓	✓		
	19-08	381,823	✓	✓	✓	
	23-08	84,403		✓		
	25-08	131,071		✓		
Type 2	02-08	318,529	✓	✓		
	07-08	204,827	✓			
	08-08	605,141	✓	✓		
	15-08	265,257	✓	✓		
	19-08	381,823	✓	✓		
	23-08	84,403				
	25-08	131,071				
Type 3	02-08	318,529	✓	✓		
	07-08	204,827				
	08-08	605,141	✓			
	15-08	265,257	✓			
	19-08	381,823	✓	✓		
	23-08	84,403				
	25-08	131,071				

A peculiar effect can be noticed when looking at the type 1 snapshots. For certain days, the effect is can not be proven for the first degree but is proven for the second degree. This is attributed to the lower number of infected flights on that day. This means that the majority of aircraft will have 0%

infected neighbor. If random sampling is used on this set of data, the chance that zeros are chosen is relatively high. A lot of zero values will lead to a lot of ties in the Mann-Whitney U test which in turn has a significant effect on the U scores. As can be seen in Table B.5, the median values do show a difference. The statistical evidence to support the difference is slightly lacking on days with below average delay.



# 6

## Verification and validation

This chapter will present two key attributes of found research, the verification and validation. Verifying if the model is doing what it is supposed to do is done in Section 6.1 with the help of a sensitivity analysis. Certain input parameters are changed and the resulting network and delay characteristics are analyzed to see if they behave according to expectation. The sensitivity analysis also aims at showing the effect of certain assumptions on the result.

To validate the model, the results are compared to days from December 2016 and April 2017 in Section 6.2. If these networks and delay characteristics match, the results are assumed to be validated.

### 6.1. Verification and sensitivity analysis

In this section a sensitivity analysis is performed which aims at showing the effect of certain assumptions made and to verify whether the model is working according to expectation. Where applicable the effects on the network representation and on the delay propagation will be analyzed. One snapshot is taken which should provide an average result in terms of effects, August 15th at 20:00. It has been shown that the networks do not vary much in the afternoon and this day has shown an average amount of delay. Section 6.1.1 and Section 6.1.2 will discuss the effects of changing the type 1 and type 2 time windows. When creating the snapshot, a time limit of 6 hours is included to neglect edges made before that time. The effect of this time limit is shown in Section 6.1.3. Next, an aircraft is considered infected when it has a delay of 15 minutes. The effect of changing this definition is shown in Section 6.1.4. Finally, in the analysis predecessors were used to determine the neighbors in a graph. The effect of using successors instead is discussed in Section 6.1.5.

#### 6.1.1. Type 1 time window

Changing the type 1 time windows will effect the (number of) aircraft contacted. For the original definition of type 1, the time window is defined by the scheduled times and actual times. The scheduled arrival time of an arriving aircraft should be between 30-120 minutes before the scheduled departure time of a departing aircraft. The second requirement is that the actual arrival time of the arriving aircraft should be before the actual departure time of the departing aircraft. Both of these time windows will be altered and the effects will be discussed. First, the following scenarios will be covered in terms of changing the scheduled time window:

- Null: 30-120 minutes
- Alternative 1: 30-150 minutes
- Alternative 2: 0-120 minutes

- Alternative 3: 30-90 minutes
- Alternative 4: 30-60 minutes

The null scenario is the one used throughout this report. The first alternative extends the time window 30 minutes at the end, the second alternative extends the time window 30 minutes at the front, the third alternative shortens the time window by 30 minutes and finally the fourth time window is 60 minutes shorter.

Table 6.1 presents the network as described by the original time window and the changes the alternatives cause. Extending the time window by 30 minutes (Alternative 1 and 2) causes increases to the network size. The number of nodes shows a small increase, while the number of edges and average degree show increases between 20 to 30%. The average shortest path length and number of strong components stay relatively constant, while the number of weak components decreases for Alternative 1. This is not the case for Alternative 2

When decreasing the window length to 60 and 30 minutes the network decreases in size accordingly. The number of components increase as the network becomes increasingly less connected. Also the average shortest path length increases which is a reasonable consequence of the decrease in edges.

Table 6.1: Effect on network representation when changing type 1 scheduled time window.

Scenario	Nodes	Edges	Avg. degree	% possible paths	Avg. shortest path	Strong components	Weak components
Null	3,074	41,752	13.6	28.5%	5.34	1,404	11
Alternative 1	2.3%	31.2%	27.9%	0.4%	-12.9%	-0.3%	-18.2%
Alternative 2	1.7%	23.6%	21.3%	0.6%	-10.5%	-0.2%	0.0%
Alternative 3	-2.4%	-28.0%	-26.5%	-4.3%	20.4%	6.0%	109.1%
Alternative 4	-9.4%	-65.6%	-61.8%	-16.1%	46.1%	28.2%	336.4%

The effect of changing the time window on the first degree neighbors is shown in Table 6.2. Again the biggest change in the number of neighbors of a node can be seen in Alternative 3 and 4. The median number of neighbors doubles which is the direct result of doubling the length of the time window. The median values of percentage of infected neighbors does not change significantly. A slight decrease in the standard deviation and IQR can be found in Alternatives 3 and 4, which means the spread of this number is slightly smaller.

When feeding the different alternatives into the Mann-Whitney U tests, the z-scores obtained are given in Table 6.3.

Next, the time window of the actual times is changed. The following scenarios are covered:

- Null: Actual arrival time before actual departure time
- Alternative 1: Actual arrival time at least 15 minutes before actual departure time
- Alternative 2: Actual arrival time no more than 15 minutes after the actual departure time
- Alternative 3: No limit

The null situation is that the actual arrival time of an aircraft needs to be before the actual departure time of the other aircraft, this is used throughout the report. This window is changed to at least 15 minutes before (alternative 1) and at most 15 minutes after (alternative 2) the actual departure time. Finally, the third alternative is that this requirement is not included at all.

The effect on the network representation is shown in Table 6.4. The magnitude of changing the actual time window is relatively small. Alternative 1 and 2 show a slight decrease and increase in network size (number of nodes and edges), which is to be expected when shifting the time window forward and backward. The third alternative shows a slightly higher increase which is also as expected when

Table 6.2: Effect on first degree neighbors when changing type 1 scheduled time window.

Alternative	Nodes	Number	Number of neighbors			Percentage infected		
			<i>Md</i>	$\sigma$	IQR	<i>Md</i>	$\sigma$	IQR
Null	All	2,012	16	19.6	6.0, 26.3	23.3%	21.5%	10.5%, 37.3%
	Susceptible	1,390	16	20.9	6.0, 28.0	20.0%	19.1%	7.7%, 32.1%
	Infected	622	17	16.4	7.0, 26.0	31.6%	23.9%	18.2%, 50.0%
Alternative 1	All	2,031	21	26.0	9.0, 34.0	24.0%	20.1%	12.5%, 35.7%
	Susceptible	1,405	21	27.7	9.0, 36.0	20.0%	18.0%	9.5%, 33.3%
	Infected	626	22	21.4	9.0, 32.0	30.8%	22.1%	20.0%, 50.0%
Alternative 2	All	2,028	20	23.9	8.0, 31.0	22.2%	20.5%	11.1%, 33.3%
	Susceptible	1,401	20	25.5	8.0, 33.0	18.2%	17.6%	7.8%, 30.0%
	Infected	627	22	19.9	9.0, 31.0	30.4%	23.3%	18.8%, 50.0%
Alternative 3	All	1,964	10	14.8	5.0, 20.0	22.2%	23.5%	6.7%, 38.5%
	Susceptible	1,352	10	15.8	5.0, 22.0	18.8%	20.8%	3.6%, 33.3%
	Infected	612	11	12.4	5.0, 18.0	33.3%	26.4%	16.2%, 50.0%
Alternative 4	All	1,818	5	8.2	2.0, 10.0	20.0%	28.2%	0.0%, 39.2%
	Susceptible	1,248	5	8.7	2.0, 11.0	14.3%	25.6%	0.0%, 33.3%
	Infected	570	4	7.0	2.0, 8.0	30.0%	31.3%	0.0%, 50.0%

Table 6.3: Effect on Mann-Whitney U test when changing type 1 scheduled time window.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
Null	30	32%	70	20%	704	2.61	23.8%
Alternative 1	30	31%	70	20%	709	2.57	26.0%
Alternative 2	30	30%	70	18%	662	2.92	13.8%
Alternative 3	31	33%	69	19%	737	2.49	29.8%
Alternative 4	31	30%	69	14%	805	2.02	46.9%

the time window is completely removed. The average shortest path length also decreases in the third alternative which can be attributed to the increase in edges.

Next, the effect on the first degree neighbor characteristics is given in Table 6.5. Again, there is little to no change in these characteristics. All median, standard deviation and IQR values of both the number of first degree neighbors and the percentage infected are relatively unchanged.

Feeding these alternatives in the Mann-Whitney U tests provides the results as given in Table 6.6. The number of accepted null hypotheses seems to go up with the increase in network size. Extra connections included in Alternative 2 and 3 should not have an effect on delay propagation as they are impossible connections in reality. They clearly diminish the effect of delay propagation when looking at the results of the Mann-Whitney U test.

Concluding, the time windows are chosen in order to correctly represent a connection between two aircraft in terms of transferring passengers and crew.

Table 6.4: Effect on network representation when changing type 1 actual time window.

Scenario	Nodes	Edges	Avg. degree	% possible paths	Avg. shortest path	Strong components	Weak components
Null	3074	41752	13.6	28.5%	5.34	1404	11
Alternative 1	-0.2%	-1.6%	-1.5%	0.1%	1.9%	-0.4%	0.0%
Alternative 2	0.3%	1.1%	0.7%	0.0%	-0.7%	0.3%	0.0%
Alternative 3	0.8%	4.7%	3.7%	1.2%	-4.3%	-0.4%	0.0%

Table 6.5: Effect on first degree neighbors when changing type 1 actual time window.

Alternative	Nodes	Number	Number of neighbors			Percentage infected		
			$Md$	$\sigma$	IQR	$Md$	$\sigma$	IQR
Null	All	2,012	16	19.6	6.0, 26.3	23.3%	21.5%	10.5%, 37.3%
	Susceptible	1,390	16	20.9	6.0, 28.0	20.0%	19.1%	7.7%, 32.1%
	Infected	622	17	16.4	7.0, 26.0	31.6%	23.9%	18.2%, 50.0%
Alternative 1	All	2,012	15	19.3	6.0, 26.0	22.2%	21.5%	10.0%, 35.7%
	Susceptible	1,390	15	20.5	6.0, 27.0	19.0%	18.8%	6.3%, 30.4%
	Infected	622	16	16.3	7.0, 26.0	30.9%	23.9%	17.6%, 50.0%
Alternative 2	All	2,012	16	19.8	6.0, 27.0	23.8%	21.6%	11.1%, 37.9%
	Susceptible	1,390	16	21.1	6.0, 28.8	20.0%	19.3%	8.3%, 33.3%
	Infected	622	17	16.5	7.0, 26.0	33.3%	24.0%	18.2%, 50.0%
Alternative 3	All	2,012	17	20.5	7.0, 28.0	25.0%	21.9%	12.0%, 40.0%
	Susceptible	1,390	16	21.8	7.0, 30.0	21.7%	19.8%	10.0%, 35.1%
	Infected	622	17	16.8	7.0, 26.0	33.3%	24.1%	19.1%, 50.0%

Table 6.6: Effect on Mann-Whitney U test when changing type 1 scheduled time window.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
Null	30	32%	70	20%	702	2.62	23.0%
Alternative 1	30	31%	70	19%	687	2.75	21.4%
Alternative 2	30	33%	70	20%	719	2.49	29.3%
Alternative 3	30	33%	70	22%	749	2.27	36.9%

### 6.1.2. Type 2 time window

Type 2 neighbors are made when 2 aircraft make use of the same runway. An assumption was made that edges were made between aircraft who arrive/depart within 15 minutes of each other at the same airport. In this subsection an analysis will be made as to what happens when this time window is changed to the following alternatives:

- Null: +- 15 minutes
- Alternative 1: +- 10 minutes
- Alternative 2: +- 20 minutes
- Alternative 3: - 15 minutes
- Alternative 4: + 15 minutes

For this report, type 2 was defined with a time window of plus and minus 15 minutes around the movement of an aircraft. For this sensitivity analysis this time window is decreased to plus and minus 10 minutes, increased to plus and minus 20 minutes, only include aircraft moving 15 minutes before and only including aircraft moving 15 minutes after.

When changing these time windows, the effect on the resulting network is shown in Table 6.7. One clear effect is the change in length of the time window, which has the same effect on the number of edges included in the network (i.e., a 33% increase in time is a 33% increase in number of edges). The number of nodes stays relatively constant throughout the different alternatives. The number of strongly connected components goes up when the time windows are shortened, which is a direct result of less edges.

Table 6.7: Effect on network representation when changing type 2 time window.

Scenario	Nodes	Edges	Avg. degree	% possible paths	Avg. shortest path	Strong components	Weak components
Null	3,481	150,169	43.1	95.6%	3.17	126	1
Alternative 1	-0.5%	-33.4%	-32.9%	-0.6%	10.4%	19.0%	0.0%
Alternative 2	0.4%	33.2%	32.7%	0.4%	-6.0%	-5.6%	0.0%
Alternative 3	-0.6%	-50.7%	-50.3%	-1.7%	18.9%	62.7%	0.0%
Alternative 4	-0.3%	-49.3%	-49.2%	-1.6%	18.9%	64.3%	0.0%

The effect of changing the type 2 time window on the first degree neighbors is shown in Table 6.8. The number of first degree neighbors changes along with the number of edges in the network. The median percentage of those first degree neighbors that are infected is not affected.

The effect on the results of the Mann-Whitney U test is given in Table 6.9. Changing the type 2 time window does not have a major effect on the results. The median values for infected and susceptible aircraft stays constant and there is slight decrease in z-scores if the time window is changed to one side (alternative 3 and 4).

Table 6.8: Effect on first degree neighbors when changing type 2 time window.

Alternative	Nodes	Number	Number of neighbors			Percentage infected		
			$Md$	$\sigma$	IQR	$Md$	$\sigma$	IQR
Null	All	3,379	35	31.0	18.0, 59.0	27.0%	17.3%	0.17, 0.38
	Susceptible	2,366	36	32.4	18.0, 61.0	23.8%	15.3%	0.14, 0.34
	Infected	1,013	34	27.2	19.0, 54.0	33.7%	19.3%	0.23, 0.46
Alternative 1	All	3,346	24	21.3	13.0, 40.0	26.5%	18.5%	0.15, 0.38
	Susceptible	2,344	24.5	22.3	12.0, 42.0	23.6%	16.6%	0.13, 0.34
	Infected	1,002	22	18.7	13.0, 37.0	33.3%	20.5%	0.23, 0.47
Alternative 2	All	3,394	46	40.2	24.0, 77.0	27.1%	16.6%	0.17, 0.38
	Susceptible	2,377	47	42.1	23.0, 79.0	24.1%	14.7%	0.15, 0.34
	Infected	1,017	44	35.1	25.0, 73.0	33.7%	18.3%	0.24, 0.44
Alternative 3	All	3,325	18	16.1	9.0, 30.0	26.3%	19.7%	0.15, 0.38
	Susceptible	2,329	18	16.9	9.0, 31.0	23.1%	17.8%	0.13, 0.35
	Infected	996	17	14.2	10.0, 28.0	33.3%	21.7%	0.22, 0.48
Alternative 4	All	3,325	19	16.4	10.0, 31.0	26.7%	19.7%	0.15, 0.39
	Susceptible	2,329	19	17.2	10.0, 32.0	24.0%	18.1%	0.13, 0.36
	Infected	996	17	14.2	10.0, 29.0	33.3%	21.3%	0.22, 0.5

Table 6.9: Effect on Mann-Whitney U test when changing type 2 time window.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
Null	29	34%	71	24%	664	2.78	18.9%
Alternative 1	29	34%	71	24%	663	2.79	20.3%
Alternative 2	29	34%	71	24%	668	2.75	19.6%
Alternative 3	29	33%	71	23%	711	2.42	31.0%
Alternative 4	29	33%	71	24%	699	2.52	29.6%

### 6.1.3. 6 hours snapshot

As a snapshot of the network is needed, edges before a certain time limit are excluded. This time limit is assumed to be 6 hours to include aircraft performing regular cross-country flights (flight time from New York to Los Angeles is 5.5 hours).

- Null: 6 hours
- Alternative 1: No limit
- Alternative 2: 8 hours
- Alternative 3: 4 hours
- Alternative 4: 2 hours
- Alternative 5: 1 hour

So the null time limit is equal to this 6 hours. In the alternatives, this time limit is completely removed, increased to 8 hours and decreased to 4, 2 and 1 hours. The previous two changes were type specific, this change can be applied to all three different types of networks. The changes on the network representation of type 1 is given in Table 6.10. Increasing the time limit leads to larger network, while decreasing the time limit leads to significantly smaller network.

Table 6.10: Effect on network representation when changing type 1 snapshot time limit.

Scenario	Nodes	Edges	Avg. degree	% possible paths	Avg. shortest path	Strong components	Weak components
Null	3,074	41,752	13.6	28.5%	5.34	1,404	11
Alternative 1	7.4%	7.5%	0.0%	2.7%	-9.9%	1.9%	-18.2%
Alternative 2	2.9%	3.8%	0.7%	1.2%	-8.2%	0.3%	0.0%
Alternative 3	-6.8%	-10.3%	-3.7%	-5.0%	19.1%	-0.3%	109.1%
Alternative 4	-22.3%	-40.3%	-23.5%	-16.9%	72.8%	7.9%	200.0%
Alternative 5	-37.7%	-64.5%	-43.4%	-27.2%	-59.4%	6.2%	427.3%

The changes on the results from the Mann-Whitney U tests for type 1 snapshots are given in Table 6.12. The result is only really affected once the time limit is down to 1 hour. As the network has shrunk in size by over 50%, an effect on the result was expected. The difference between imposing a limit of 2 hours and a limit of 8 hours has almost no effect on the delay results. Also the limit itself is disputed, as the first alternative with no limit shows no major differences.

Table 6.11: Effect on first degree neighbors when changing type 1 snapshot time limit.

Alternative	Nodes	Number	Number of neighbors			Percentage infected		
			<i>Md</i>	$\sigma$	IQR	<i>Md</i>	$\sigma$	IQR
Null	All	2,012	16	19.6	6.0, 26.3	23.3%	21.5%	10.5%, 37.3%
	Susceptible	1,390	16	20.9	6.0, 28.0	20.0%	19.1%	7.7%, 32.1%
	Infected	622	17	16.4	7.0, 26.0	31.6%	23.9%	18.2%, 50.0%
Alternative 1	All	2,189	16	19.5	6.0, 26.0	23.5%	21.6%	11.1%, 37.5%
	Susceptible	1,521	15	20.7	6.0, 27.0	20.0%	19.3%	7.7%, 32.3%
	Infected	668	17	16.4	6.0, 26.0	31.6%	23.9%	18.2%, 50.0%
Alternative 2	All	2,085	16	19.5	6.0, 27.0	23.5%	21.4%	11.1%, 37.5%
	Susceptible	1,442	16	20.7	6.0, 28.0	20.0%	19.0%	7.8%, 32.1%
	Infected	643	17	16.4	7.0, 26.0	31.8%	23.7%	18.3%, 50.0%
Alternative 3	All	1,780	16	20.0	6.0, 27.0	22.2%	21.3%	9.1%, 34.5%
	Susceptible	1,245	16	21.4	6.0, 29.0	18.8%	18.7%	6.3%, 30.8%
	Infected	535	16	16.2	7.0, 26.0	30.0%	23.9%	17.6%, 50.0%
Alternative 4	All	1,149	16	20.1	7.0, 28.0	18.8%	20.1%	7.7%, 30.0%
	Susceptible	816	16	21.3	7.0, 32.0	16.7%	17.6%	5.9%, 26.8%
	Infected	333	16	16.6	7.0, 26.0	25.0%	23.6%	12.5%, 42.9%
Alternative 5	All	637	17	21.9	7.0, 29.0	12.5%	16.3%	6.3%, 25.0%
	Susceptible	446	17	23.4	7.0, 34.0	11.6%	14.4%	5.9%, 23.1%
	Infected	191	17	17.5	6.0, 26.0	14.9%	19.6%	8.7%, 28.6%

This analysis can also be done for type 2 and type 3 networks. The specific results can be found in Appendix C.1, but the main conclusions are similar to the type 1 network.

Table 6.12: Effect on Mann-Whitney U test when changing type 1 snapshot time limit.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
Null	30	32%	70	20%	702	2.62	25.2%
Alternative 1	30	32%	70	20%	696	2.67	23.1%
Alternative 2	30	32%	70	20%	695	2.68	24.1%
Alternative 3	30	30%	70	19%	672	2.85	19.1%
Alternative 4	28	25%	72	17%	745	2.02	46.3%
Alternative 5	29	15%	71	12%	886	1.09	81.8%

#### 6.1.4. 15 minutes delay

As described by the FAA, a delay in air transportation is usually considered to be at least 15 minutes. This definition has been used throughout this report, but what effect does it have on the main results? In this subsection the following alternative definitions of a delay are analyzed:

- Null: 15 minutes
- Alternative 1: 10 minutes
- Alternative 2: 20 minutes

The changes in the network representation should not be affected by this change in delay definition. This is shown in Table 6.13 for type 1 networks and also holds for the other types of networks.

Table 6.13: Effect on network representation when changing type 1 definition of infected neighbors.

Scenario	Nodes	Edges	Avg. degree	% possible paths	Avg. shortest path	Strong components	Weak components
Null	3,074	41,752	13.6	0.285	5.34	1,404	11
Alternative 1	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Alternative 2	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Moving to the delay analysis, the percentage of infected neighbors is now defined as the number of neighbors with a delay larger than 10, 15 or 20 minutes depending on the alternative. The resulting effect on the neighbors is shown in Table 6.14. The number of neighbors is unchanged, as just the definition of a infected contact is changed. The median values for percentage of infected first degree neighbors go down when the definition is changed to 20 minutes and increase when the definition is changed to 10 minutes. This is a direct result of more aircraft being infected in the first alternative and less aircraft being infected in the second alternative.

Table 6.14: Effect on first degree neighbors when changing type 1 definition of infected neighbors.

Alternative	Nodes	Number	Number of neighbors			Percentage infected		
			$Md$	$\sigma$	IQR	$Md$	$\sigma$	IQR
Null	All	2,012	16	19.6	6.0, 26.3	23.3%	21.5%	10.5%, 37.3%
	Susceptible	1,390	16	20.9	6.0, 28.0	20.0%	19.1%	7.7%, 32.1%
	Infected	622	17	16.4	7.0, 26.0	31.6%	23.9%	18.2%, 50.0%
Alternative 1	All	2,012	16	19.6	6.0, 26.3	31.5%	23.0%	16.7%, 44.4%
	Susceptible	1,390	16	20.9	6.0, 28.0	27.8%	21.4%	12.5%, 40.0%
	Infected	622	17	16.4	7.0, 26.0	38.3%	24.1%	25.0%, 58.3%
Alternative 2	All	2,012	16	19.6	6.0, 26.3	20.0%	20.2%	7.1%, 32.4%
	Susceptible	1,390	16	20.9	6.0, 28.0	16.7%	17.4%	4.5%, 25.9%
	Infected	622	17	16.4	7.0, 26.0	26.5%	23.1%	14.3%, 45.2%

Now, the definition of an infected node is also changed. The nodes are now grouped by susceptible and infected according to the definition in each alternative and the results are given in Table 6.15. Interestingly, comparing this to the results in Table 6.14 gives almost no difference. The number of nodes in each category shifts a bit, but the median values for the percentage infected first degree neighbors stays the same.

Table 6.15: Effect on first degree neighbors when changing type 1 definition of infected aircraft and infected neighbors.

Alternative	Nodes	Number	Number of neighbors			Percentage infected		
			<i>Md</i>	$\sigma$	IQR	<i>Md</i>	$\sigma$	IQR
Null	All	2,012	16	19.6	6.0, 26.3	23.3%	21.5%	10.5%, 37.3%
	Susceptible	1,390	16	20.9	6.0, 28.0	20.0%	19.1%	7.7%, 32.1%
	Infected	622	17	16.4	7.0, 26.0	31.6%	23.9%	18.2%, 50.0%
Alternative 1	All	2,012	16	19.6	6.0, 26.3	31.5%	23.0%	16.7%, 44.4%
	Susceptible	1,255	16	21.3	6.0, 29.5	26.7%	21.4%	11.8%, 40.0%
	Infected	757	16	16.2	7.0, 25.0	37.5%	23.6%	25.0%, 57.1%
Alternative 2	All	2,012	16	19.6	6.0, 26.3	20.0%	20.2%	7.1%, 32.4%
	Susceptible	1,469	15	20.9	6.0, 28.0	16.7%	18.0%	4.7%, 26.8%
	Infected	543	17	15.6	7.0, 26.0	27.3%	22.9%	15.2%, 47.9%

The effect on the result of the Mann-Whitney U test is given in Table 6.16. Changing the definition to 10 minutes does not have a large effect, as the number of infected aircraft increase all over the network. Changing the definition to 20 minutes leads to less infected aircraft in the network and also decreases the z-score for the first degree effects on type 1 networks. The same effects can be found in type 2 and type 3 networks, with the specific results presented in Appendix C.2.

Table 6.16: Effect on Mann-Whitney U test when changing type 1 delay definition.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_z$	H0?
Null	30	32%	70	20%	707	2.58	24.1%
Alternative 1	37	38%	63	27%	793	2.66	24.2%
Alternative 2	26	27%	74	17%	622	2.69	22.2%

### 6.1.5. Predecessors vs. successors

In this report, when referring to the neighbors of an aircraft the predecessors are used. These aircraft have a directed connection, so when analyzing the neighbors of an aircraft the predecessors are used as they have had an effect on that aircraft. The effect of using successors will be shown in this subsection. As type 2 and type 3 networks usually lead to bi-directional edges between nodes, only the effects on type 1 networks will be investigated.

- Null: Predecessors
- Alternative 1: Successors

This change has no effect on the network representation, as shown in Table 6.17.

Table 6.17: Effect on network representation using successors instead of predecessors in type 1 networks.

Scenario	Nodes	Edges	Avg. degree	% possible paths	Avg. shortest path	Strong components	Weak components
Null	3074	41752	13.6	0.285	5.34	1404	11
Alternative 1	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

In Table 6.18 the effect of using successors instead of predecessors is presented. The median number of first degree neighbors goes down when using successors, this means that on average each aircraft has more incoming edges than outgoing edges. The number of aircraft included in the network is also higher when using successors. As aircraft are removed from the analysis when they have zero neighbors, the previous observation is immediately proven. As far as the median values for the percentage of infected first degree neighbors, these are not affected by this change.

Table 6.18: Effect on first degree neighbors in type 1 networks when using successors instead of predecessors.

Alternative	Nodes	Number	Number of neighbors			Percentage infected		
			$Md$	$\sigma$	IQR	$Md$	$\sigma$	IQR
Null	All	2,012	16	19.6	6.0, 26.3	23.3%	21.5%	10.5%, 37.3%
	Susceptible	1,390	16	20.9	6.0, 28.0	20.0%	19.1%	7.7%, 32.1%
	Infected	622	17	16.4	7.0, 26.0	31.6%	23.9%	18.2%, 50.0%
Alternative 1	All	2,767	10	14.1	5.0, 19.0	28.6%	27.9%	14.2%, 50.0%
	Susceptible	1,992	11	14.7	6.0, 21.0	25.0%	25.7%	11.3%, 42.9%
	Infected	775	9	11.8	5.0, 17.0	42.9%	29.9%	22.2%, 66.7%

When these values are put into the Mann-Whitney U test, the resulting z-scores can be compared (as shown in Table 6.19). The z-score decreases significantly, decreasing the effect of delay propagation.

Table 6.19: Effect on Mann-Whitney U test when using successors instead of predecessors.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	H0?
Null	30	32%	70	20%	710	2.56	23.7%
Alternative 1	28	43%	72	25%	689	2.45	29.2%

## 6.2. Validation

To validate the results from this research, they are compared to days in different parts of the year. Two days (December 11th, 2016 and April 2nd, 2017) are chosen which have a similar total arrival delay to August 15th as shown in Figure 6.1. The arrival delay develops in a similar manner on each day, although the peak of April 2nd is slightly lower than the other two days.

The characteristics of the 20:00 snapshot built on each of these days is given in Table 6.20 for each different connection type. Clearly the network is smaller on December 11th (in terms of number of nodes and edges), which can be explained by the fact that less flights are operated in the winter schedule. The rest of the network characteristics are relatively similar which provides support that the 20:00 snapshot characteristics presented in Section 5.2 are valid for days with similar total delay.

Table 6.20: 20:00 snapshots for all connection types on 3 similar days.

	Date	Nodes	Edges	Avg. degree	% paths possible	Avg. shortest path	Strong	Weak	$C$
Type 1	2016-8-15	3,074	41,752	13.6	28.5%	5.34	1404	11	0.36
	2016-12-11	2,701	34,759	12.9	25.3%	6.28	1274	22	0.38
	2017-4-2	3,020	38,084	12.6	24.7%	5.34	1456	22	0.36
Type 2	2016-8-15	3,481	150,169	43.1	95.6%	3.17	126	1	0.34
	2016-12-11	3,245	123,993	38.2	92.6%	3.27	156	1	0.34
	2017-4-2	3,447	140,130	40.7	96.0%	3.23	123	1	0.35
Type 3	2016-8-15	3,756	230,813	61.5	80.0%	3.02	684	1	0.39
	2016-12-11	3,552	208,960	58.8	78.5%	3.05	718	1	0.39
	2017-4-2	3,771	227,814	60.4	78.7%	3.06	727	1	0.40

Next the first degree neighbor data is compared and this is given in Table 6.21 for all connection types.

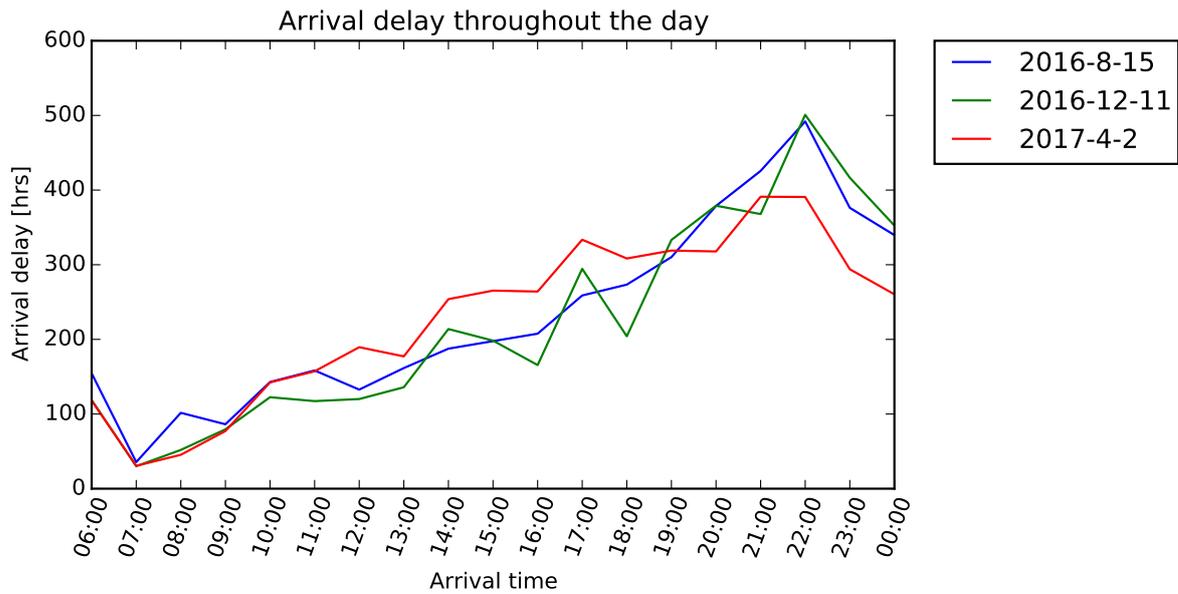


Figure 6.1: The development of arrival delays throughout similar days used in validation.

The number of first degree neighbors of each node is slightly lower for the December and April snapshots. This is in accordance with the slightly lower average degree observed in Table 6.20. What stays relatively constant is the percentage of infected aircraft in the entire network, which varies between 27.6% and 30.9% for the different scenarios. The number of infected first degree neighbors seems to vary along with this number, with the median value for all nodes varying between 18.4% and 27.0%.

When comparing the percentage of infected first degree neighbors for the infected nodes to the susceptible nodes, the median value of the first is generally higher. Feeding the data into Mann-Whitney U tests leads to the results given in Table 6.22. Again the median z-scores and % of accepted null hypotheses are in the same order of magnitude. In Table C.15 and Table C.16 the Z-scores are given for the second and third degree Mann-Whitney U tests. These also show similar behavior when comparing the different days.

Table 6.21: Results for first degree neighbors on 3 similar days.

Connection	Date	Status	Number	Number of neighbors			Percentage infected		
				$Md$	$\sigma$	IQR	$Md$	$\sigma$	IQR
Type 1	2016-8-15	All	2,012	16	19.6	6.0, 26.25	23.3%	21.5%	10.5%, 37.3%
		Susceptible	1,390	16	20.9	6.0, 28.0	20.0%	19.1%	7.7%, 32.1%
		Infected	622	17	16.4	7.0, 26.0	31.6%	23.9%	18.2%, 50.0%
	2016-12-11	All	1,796	13	20.3	6.0, 24.0	19.3%	23.1%	11.1%, 33.3%
		Susceptible	1,263	14	21.2	6.0, 25.0	17.0%	19.4%	10.7%, 27.3%
		Infected	533	12	17.9	6.0, 20.0	25.0%	28.3%	14.1%, 50.0%
	2017-4-2	All	1,964	13	19.5	6.0, 25.0	18.4%	24.7%	8.5%, 37.6%
		Susceptible	1,397	13	20.3	6.0, 25.0	15.4%	21.1%	7.8%, 33.3%
		Infected	567	13	17.3	6.0, 26.0	33.3%	28.6%	14.3%, 62.0%
Type 2	2016-8-15	All	3,379	35	31.0	18.0, 59.0	27.0%	17.3%	16.7%, 37.5%
		Susceptible	2,366	36	32.4	18.0, 61.0	23.8%	15.3%	14.3%, 33.8%
		Infected	1,013	34	27.2	19.0, 54.0	33.7%	19.3%	23.4%, 46.4%
	2016-12-11	All	3,113	29	31.4	15.0, 48.0	23.3%	17.7%	16.3%, 33.9%
		Susceptible	2,222	30	32.6	15.0, 51.0	21.9%	14.3%	15.6%, 30.8%
		Infected	891	27	28.0	15.0, 42.0	29.4%	22.6%	18.8%, 45.5%
	2017-4-2	All	3,346	32	30.0	17.0, 52.0	22.4%	18.4%	15.0%, 34.6%
		Susceptible	2,423	33	30.9	17.0, 54.0	20.6%	16.2%	14.0%, 30.0%
		Infected	923	31	27.5	18.0, 47.0	30.4%	21.2%	18.8%, 50.4%
Type 3	2016-8-15	All	3,123	53	81.6	17.0, 98.0	25.0%	16.8%	15.4%, 33.3%
		Susceptible	2,166	49	81.4	15.0, 94.0	23.1%	15.8%	13.3%, 31.6%
		Infected	957	62	81.6	22.0, 106.0	28.6%	17.8%	21.1%, 40.0%
	2016-12-11	All	2,875	52	77.6	17.0, 93.0	21.6%	15.4%	15.9%, 30.5%
		Susceptible	2,062	50	78.2	16.0, 93.0	20.5%	14.1%	15.4%, 28.6%
		Infected	813	56	76.0	24.0, 93.0	25.4%	17.4%	17.5%, 34.6%
	2017-4-2	All	3,093	50	81.4	17.0, 92.0	20.0%	17.0%	13.6%, 30.0%
		Susceptible	2,234	49	82.8	16.0, 89.0	18.8%	16.2%	12.9%, 27.5%
		Infected	859	56	77.8	18.0, 97.0	24.5%	18.1%	16.2%, 42.9%

Table 6.22: Mann-Whitney U test results for first degree neighbors of similar days.

Connection	Date	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	H0?
Type 1	2016-8-15	30	0%	70	0%	697	2.66	24%
	2016-12-11	29	0%	71	0%	766	2.01	49%
	2017-4-2	28	0%	72	0%	656	2.71	22%
Type 2	2016-8-15	29	0%	71	0%	663	2.79	18%
	2016-12-11	28	0%	72	0%	728	2.16	42%
	2017-4-2	27	0%	73	0%	673	2.43	32%
Type 3	2016-8-15	30	0%	70	0%	772	2.1	44%
	2016-12-11	28	0%	72	0%	769	1.84	55%
	2017-4-2	27	0%	73	0%	740	1.91	52%

# 7

## Conclusion and recommendations

This chapter will present the most important conclusions and recommendations following from this research. It is split up into two main sections which will cover the conclusions (Section 7.1) and recommendations (Section 7.2) respectively.

### 7.1. Conclusion

The conclusions of the research will be presented by answering the research questions discussed in Section 3.2. By answering these questions the research objective will be achieved.

#### **Q1 General: Is the probability of an aircraft being delayed influenced when it comes into contact with other delayed aircraft?**

Yes the probability of being delayed increases when aircraft come into contact with other delayed aircraft. This is proven by analyzing the neighbors of aircraft in different types of network. These effects are dependent on the magnitude of delay and the type of connection. In figures such as Figure 5.21 this effect is shown, an increase in the number of infected neighbors leads to an increase in the probability of an aircraft being delayed.

Of course this effect is to be expected because aircraft interact at the same airport. If this airport is experiencing bad weather or congestion, the probability of an aircraft being delayed will inherently increase. The most interesting effects are thus when second, third and fourth degree effects are proven. The probability of an aircraft being delayed increases when aircraft multiple degrees away are delayed, depending on the day and connection type. This means that the probability of an aircraft being delayed increases when other aircraft are delayed which have no direct connection.

When considering the second degree connections, these aircraft could have still operated at the same airport. This is where a careful analysis of the network characteristics is needed in order to get to the right conclusion. Table 5.10 shows the mean values of the network characteristics, and in particular the average shortest path length and clustering coefficient are important. The clustering coefficient can be seen to be in the range of 0.35-0.40, so about 35 to 40 % of the possible 'triangles' between three nodes are present. This means that a node shares this percentage of neighbors with one of its specific neighbors. The average shortest path length can be seen to be around 3 for Type 2 and Type 3, while it is around 5 for Type 1. This means that on average it would take 3 steps to get to another node in the network (i.e., another aircraft performing a domestic flight in the US).

So when third degree effects are proven to exist within a Type 2 network, it means that the average shortest path length has been reached. This means that on average you would be able to reach every other node in the network by the third degree. Making the assumption that this means around 50% of

the nodes, that is half of the total amount of aircraft performing a domestic flight on that particular day. The local effects of being at the same airport have disappeared when considering such a big subset of aircraft. In the case of Type 2 and Type 3 however, only second degree effects are proven. Still this would suggest that within 2 degrees a significant portion (around 30%) of the nodes in the network has been reached.

So to which degree is the probability of an aircraft being delayed affected? The degree to which effects can be proven per connection type and total delay in the system can be summarized as follows:

- Type 1:
  - Low delay: 2nd degree
  - Average delay: 2nd degree
  - High delay: 3rd degree
  - Extreme delay: 4th degree
- Type 2:
  - Low delay: No effect
  - Average delay: 1st degree
  - High delay: 2nd degree
  - Extreme delay: 2nd degree
- Type 3:
  - Low delay: No effect
  - Average delay: 1st degree
  - High delay: 2nd degree

### **Q2 General: Can delay propagation be found every day in the ATS?**

No, the amount of delay on a specific day along with the total number of aircraft infected will have an effect on the delay propagation. Certain effects that were observed during days with high delay were not found on days with low delay.

Going to the analogy with a disease spreading across a network of individual people, more people will become infected if more people are infected (i.e., higher the chance of exposure to the disease). The same can be said for the chance for aircraft of becoming delayed. Directly this is caused by things like bad weather and congestion, as more aircraft will be delayed in a bad storm. But the indirect causes which is ultimately delay propagation are proven to exist in this research. However, the amount of degrees this effect travels will ultimately be defined by the total aircraft infected at a given point in time.

### **Q3 General: Is the spreading of delays in the ATS comparable to the spreading of social behavior?**

In papers such as Christakis and Fowler [2007] the spread of obesity throughout a social network is analyzed. In this case it was observed that if contacts in your network up to three degrees away had obesity, it would increase your chance of becoming obese.

In this research, a very similar conclusion can be made. Given the right circumstances, aircraft up to three degrees away can impact the probability of an aircraft being delayed. This means that aircraft are affecting each other without making direct contact with each other, much like the obesity spread.

In this case the aircraft interacting with each other can be compared to individuals interacting in a social environment. The biggest difference is in defining the network and in particular the set of edges between nodes. Defining a social tie is relatively easy (i.e., a good friend, family, colleague) and is usually long lasting. In this research, defining an edge between two aircraft was done in an abstract

and temporary manner (i.e., an edge between two aircraft will disappear after a few hours). This is done in order to represent two aircraft which have interacted in a certain way, but after performing another set of flights this edge would no longer exist. This way, the network of aircraft is a dynamic and ever-evolving network but it has been proven to show relatively stable network characteristics.

When looking at the different networks created by the connection types, only Type 1 network shows small-world characteristics. A power-law fit can be made for the higher degree values with  $\gamma$  varying between 2.5 and 3.5 for the majority of the networks. For the other two types a random degree distribution is found, which does not correlate with small-world networks.

## 7.2. Limitations and recommendations

The methodology used in this research has some limitations which will be explained in this section. Also, some recommendations can be made which can be used in further research.

### 7.2.1. Limitations

The first limitation is related to defining the network. The process of creating the snapshots was used in order to create a static representation of the network. Adding time to the equation might provide an added bonus to the network.

When looking at better representing edges between nodes, more data is needed. In the approach to currently representing the networks, more information could be used to more accurately represent when aircraft come into contact. In short for each type of connection the following data can be collected:

- Type 1: extra data on transferring passengers and crew.
- Type 2: more precise information on runways used (i.e., same runway or parallel runways).
- Type 3: more precise information on location of gate or pier used.

Adding this information to generating the network will narrow down the number of edges in each network and provide a more accurate representation of the aircraft that have interacted. This will determine whether the assumptions made in this research are too generic or are fairly accurate.

Another limitation of the model is geographical scope of the data. Only flights between two US airports are included in the dataset, which means that international flights are out of the scope. However, these flights should be included in all three types of connections in order to gain a full picture. Adding these flights will increase the number of nodes and edges in the network. Adding the international flights to the current model should enhance the results, as the impact these larger aircraft have on each type of connection is greater (i.e., more passenger to transfer, longer runway usage, more resources needed at the airport).

Finally, no direct causality can be proven using this model. It is not possible to pinpoint which flight directly affected another flight. It is only proven that there is correlation between the probability of an aircraft being delayed and the  $n$ th degree neighbors of that aircraft.

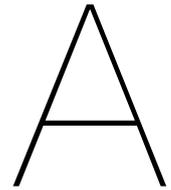
### 7.2.2. Recommendations for future work

A few recommendations can be made which (if implemented) will increase the quality of the research.

- Data quality. A few errors were manually removed from the data. These errors had to do with the wrong flight date in the data. This led to situations where one aircraft was performing two flights at the same time. When running the model, this error leads to a break in the calculations. On average every day has around 2 flights that have this error, which were corrected by comparing the dataset to FlightRadar24 data. However, the fact that these errors existed in the dataset might point to other discrepancies in the data. Some simple constraints were considered but not

implemented due to lack of data. For example, a simple continuity constraint for each aircraft at each airport could be included. The problem in this case was that there was no available data on 'deadhead' flights, where airlines reposition aircraft by flying it without passengers.

- Broaden geographical scope. This step would validate conclusions made in this research. Can the same patterns be found in places such as Europe and Asia. Due to the lack of data available in this research the geographical scope was limited to the US, but comparing it to different places or including all of the data in one model can add valuable information.
- Simulate delay patterns. Attempts can also be made to simulate the delays by using epidemic contagion and recovery rates.
- Decrease computational time. No attempt was made to optimize the computational time of the model. Adding a computer science perspective to the problem, the computational time could probably be decreased. At this moment it takes about 15 minutes to analyze one day of data and another 5 minutes are needed to build a single snapshot and save the data to an external source. This is all done in Python software which provides a useful data analysis environment with Pandas.
- Aircraft size. As mentioned, the aircraft size might have an impact on the delay propagation. Larger effects are expected as larger aircraft use more resources, have longer runway times and transfer more passengers. Including this as a variable in the model would increase the quality of the model.
- Airport capacity. The capacity of an airport will have an influence on the delay propagation. If an airport is operating to its full capacity, a disturbance such as a delay will have a larger effect on the delay propagation. Including this as a variable in the model will also increase the quality.



## BTS data profile

Table A.1: Detailed definitions of BTS dataset.

<b>Term</b>	<b>Definition</b>
Actual Arrival Times	Gate arrival time is the instance when the pilot sets the aircraft parking brake after arriving at the airport gate or passenger unloading area. If the parking brake is not set, record the time for the opening of the passenger door. Also, carriers using a Docking Guidance System (DGS) may record the official gate-arrival time when the aircraft is stopped at the appropriate parking mark.
Actual Departure Times	Gate departure time is the instance when the pilot releases the aircraft parking brake after passengers have loaded and aircraft doors have been closed. In cases where the flight returned to the departure gate before wheels-off time and departed a second time, report the last gate departure time before wheels-off time. In cases of an air return, report the last gate departure time before the gate return. If passengers were boarded without the parking brake being set, record the time that the passenger door was closed. Also, carriers using a Docking Guidance System may record the official gate-departure time based on aircraft movement. For example, one DGS records gate departure time when the aircraft moves more than 1 meter from the appropriate parking mark within 15 seconds. Fifteen seconds is then subtracted from the recorded time to obtain the appropriate out time.
Airline ID	An identification number assigned by US DOT to identify a unique airline (carrier). A unique airline (carrier) is defined as one holding and reporting under the same DOT certificate regardless of its Code, Name, or holding company/corporation. Use this field for analysis across a range of years.
Airport Code	A three character alpha-numeric code issued by the U.S. Department of Transportation which is the official designation of the airport. The airport code is not always unique to a specific airport because airport codes can change or can be reused.

*Continued on next page*

Table A.1 – *Continued from previous page*

<b>Term</b>	<b>Definition</b>
Airport ID	An identification number assigned by US DOT to identify a unique airport. Use this field for airport analysis across a range of years because an airport can change its airport code and airport codes can be reused.
Arrival Delay	Arrival delay equals the difference of the actual arrival time minus the scheduled arrival time. A flight is considered on-time when it arrives less than 15 minutes after its published arrival time.
CRS	Computer Reservation System. CRS provide information on airline schedules, fares and seat availability to travel agencies and allow agents to book seats and issue tickets.
Cancelled Flight	A flight that was listed in a carrier's computer reservation system during the seven calendar days prior to scheduled departure but was not operated.
Carrier Code	Code assigned by IATA and commonly used to identify a carrier. As the same code may have been assigned to different carriers over time, the code is not always unique.
Certificate Of Public Convenience And Necessity	A certificate issued to an air carrier under 49 U.S.C. 41102, by the Department of Transportation authorizing the carrier to engage in air transportation.
Certificated Air Carrier	An air carrier holding a Certificate of Public Convenience and Necessity issued by DOT to conduct scheduled services interstate. Nonscheduled or charter operations may also be conducted by these carriers. (same as Certified Air Carrier)
Certified Air Carrier	An air carrier holding a Certificate of Public Convenience and Necessity issued by DOT to conduct scheduled services interstate. Nonscheduled or charter operations may also be conducted by these carriers. (same as Certificated Air Carrier)
City Market ID	An identification number assigned by US DOT to identify a city market. Use this field to consolidate airports serving the same city market.
Departure Delay	The difference between the scheduled departure time and the actual departure time from the origin airport gate.
Diverted Flight	A flight that is required to land at a destination other than the original scheduled destination for reasons beyond the control of the pilot/company.
Domestic Operations	All air carrier operations having destinations within the 50 United States, the District of Columbia, the Commonwealth of Puerto Rico, and the U.S. Virgin Islands.
Elapsed Time	The time computed from gate departure time to gate arrival time.
FIPS	Federal Information Processing Standards. Usually referring to a code assigned to any of a variety of geographic entities (e.g. counties, states, metropolitan areas, etc). FIPS codes are intended to simplify the collection, processing, and dissemination of data and resources of the Federal Government.
Flight Number	A one to four character alpha-numeric code for a particular flight.
In-Flight Time	The total time an aircraft is in the air between an origin-destination airport pair, i.e. from wheels-off at the origin airport to wheels-down at the destination airport.

*Continued on next page*

Table A.1 – *Continued from previous page*

<b>Term</b>	<b>Definition</b>
Late Flight	A flight arriving or departing 15 minutes or more after the scheduled time.
Passenger Revenues	Revenues from the air transportation of passengers.
Scheduled Departure Time	The scheduled time that an aircraft should lift off from the origin airport.
Scheduled Time Of Arrival	The scheduled time that an aircraft should cross a certain point (landing or metering fix).
Taxi-In Time	The time elapsed between wheels down and arrival at the destination airport gate.
Taxi-Out Time	The time elapsed between departure from the origin airport gate and wheels off.
Unique Carrier	Unique Carrier Code. It is the Carrier Code most recently used by a carrier. A numeric suffix is used to distinguish duplicate codes, for example, PA, PA (1), PA (2). Use this field to perform analysis of data reported by one and only one carrier.
World Area Code (WAC)	Numeric codes used to identify geopolitical areas such as countries, states (U.S.), provinces (Canada), and territories or possessions of certain countries. The codes are used within the various data banks maintained by the Office of Airline Information (OAI) and are created by OAI.



# B

## Neighbor results

### B.1. Neighbor analysis 15-08-2016, Type 1, Degree 2-3

Table B.1: Number of 2nd degree neighbors and percentage of neighbors infected for type 1 snapshots.

Time	Nodes	Number	Number of neighbors			Percentage infected		
			<i>Md</i>	$\sigma$	IQR	<i>Md</i>	$\sigma$	IQR
16:00	All	2,006	50	56.2	21.0, 100.0	19.0%	14.7%	11.1%, 27.2%
	Susceptible	1,526	49	56.6	21.0, 100.0	17.5%	12.7%	9.7%, 25.0%
	Infected	480	54	54.8	22.0, 105.0	24.2%	18.2%	15.8%, 32.5%
18:00	All	2,031	51	65.0	19.0, 101.0	22.6%	15.0%	15.5%, 31.8%
	Susceptible	1,431	46	65.1	18.0, 95.0	20.4%	13.2%	14.1%, 28.9%
	Infected	600	61	64.7	23.0, 106.3	28.9%	17.0%	19.6%, 37.8%
20:00	All	1,974	41	61.4	16.0, 100.0	23.8%	17.7%	14.3%, 33.3%
	Susceptible	1,365	37	61.0	16.0, 93.0	20.0%	15.7%	12.3%, 31.2%
	Infected	609	47	62.1	17.0, 113.0	30.3%	19.2%	21.5%, 43.7%
22:00	All	2,253	58	75.5	21.0, 110.0	27.5%	18.6%	17.3%, 38.2%
	Susceptible	1,507	58	75.1	23.0, 109.0	23.7%	16.3%	15.4%, 33.6%
	Infected	746	55	76.3	19.0, 112.8	33.3%	20.3%	25.2%, 48.4%

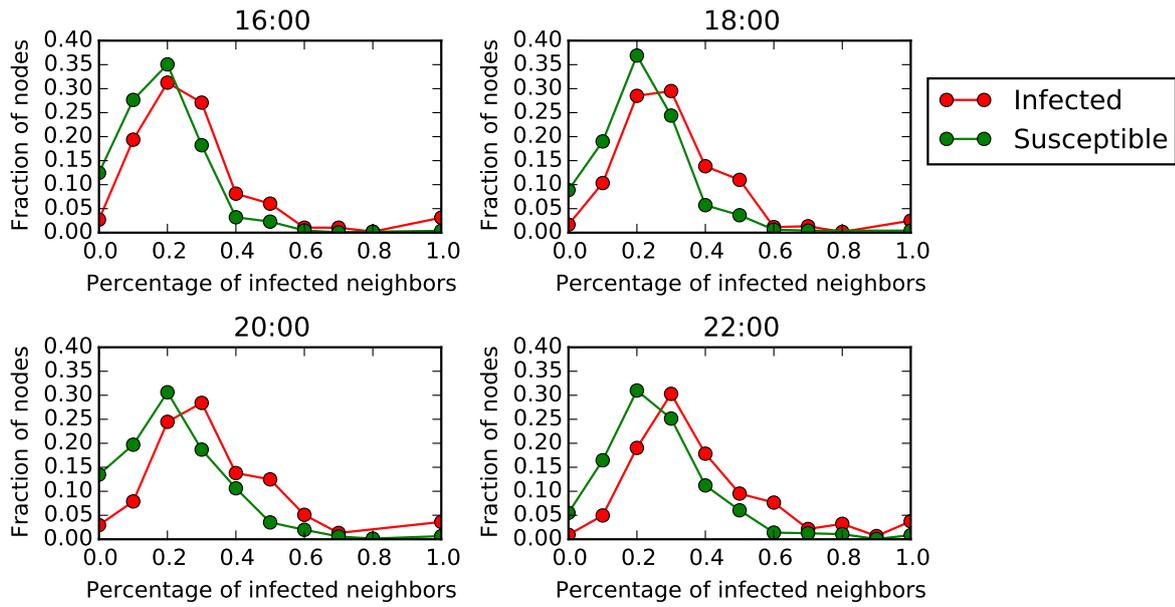


Figure B.1: The distribution of the percentage of infected 2nd degree neighbors of susceptible and infected nodes.

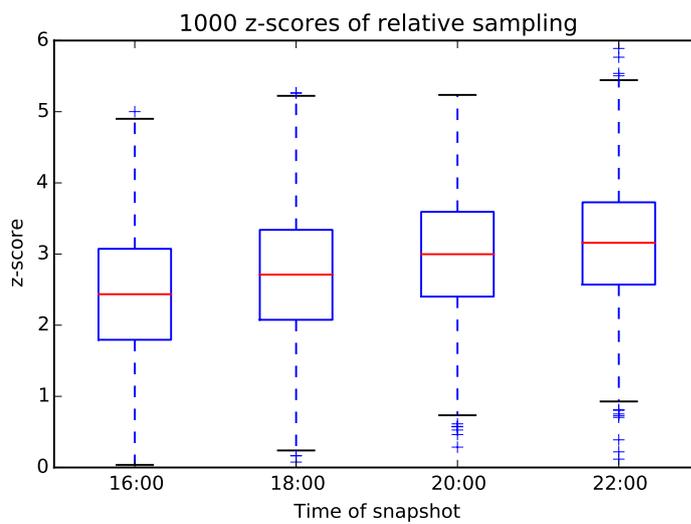


Figure B.2: Boxplots of the z-score results from the Mann-Whitney U tests for 2nd degree type 1 neighbors.

Table B.2: Number of 3rd degree neighbors and percentage of neighbors infected for type 1 snapshots.

Time	Nodes	Number	Number of neighbors			Percentage infected		
			<i>Md</i>	$\sigma$	IQR	<i>Md</i>	$\sigma$	IQR
16:00	All	1,935	128	110.2	44.5, 209.0	20.0%	10.3%	14.3%, 23.8%
	Susceptible	1,476	126	110.1	44.0, 202.0	19.5%	9.9%	13.8%, 23.6%
	Infected	459	148	110.0	52.0, 242.0	21.9%	11.1%	17.2%, 24.6%
18:00	All	1,947	146	116.9	44.0, 221.5	23.5%	10.2%	18.6%, 28.9%
	Susceptible	1,371	130	114.5	41.0, 203.0	22.9%	10.5%	18.0%, 28.3%
	Infected	576	174	121.2	59.8, 249.3	25.6%	9.4%	20.5%, 30.1%
20:00	All	1,900	104	128.7	45.0, 209.3	24.6%	12.8%	18.5%, 30.6%
	Susceptible	1,314	101	125.0	43.0, 185.8	23.0%	12.0%	17.9%, 29.9%
	Infected	586	124	135.0	59.3, 254.0	26.8%	14.2%	20.9%, 32.2%
22:00	All	2,194	165	141.4	64.0, 251.0	27.1%	13.0%	22.6%, 33.6%
	Susceptible	1,477	162	137.1	65.0, 251.0	26.2%	12.4%	21.4%, 32.4%
	Infected	717	168	149.6	62.0, 272.0	29.9%	13.9%	24.7%, 35.1%

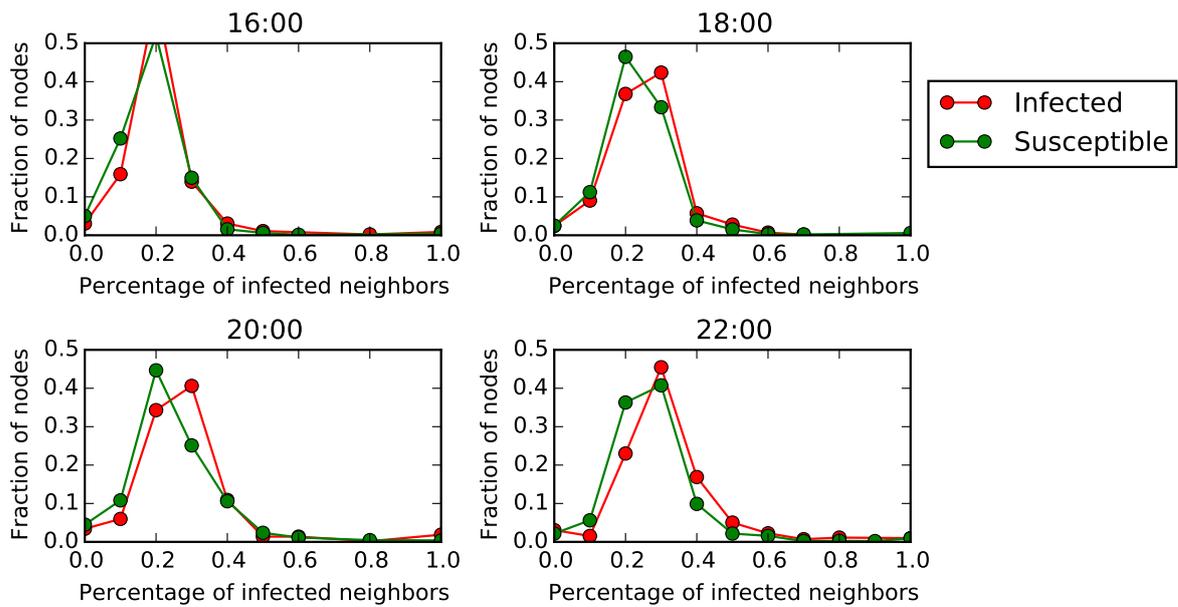


Figure B.3: The distribution of the percentage of infected 3rd degree neighbors of susceptible and infected nodes.

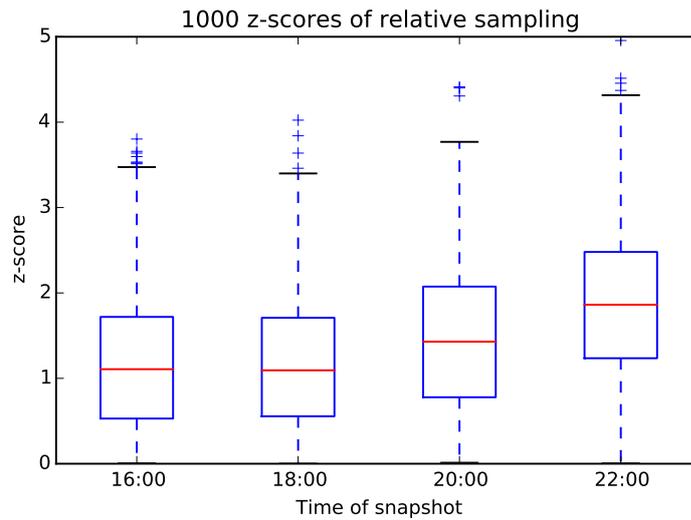


Figure B.4: Boxplots of the z-score results from the Mann-Whitney U tests for 3rd degree type 1 neighbors.

## B.2. Probability of being infected 15-08-2016, Type 1, Degree 2-4

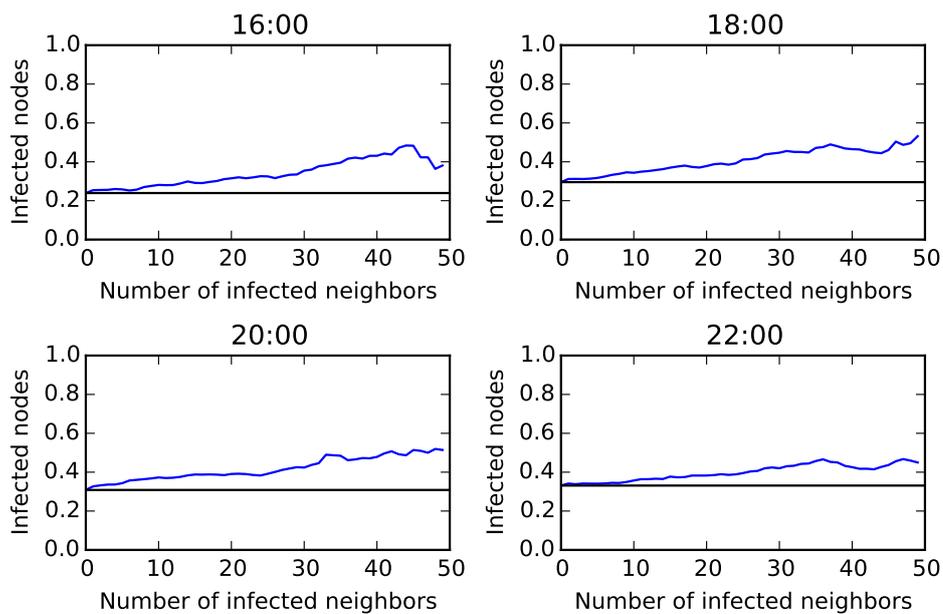


Figure B.5: Minimum number of infected second degree neighbors versus percentage of infected nodes for type 1 snapshots.

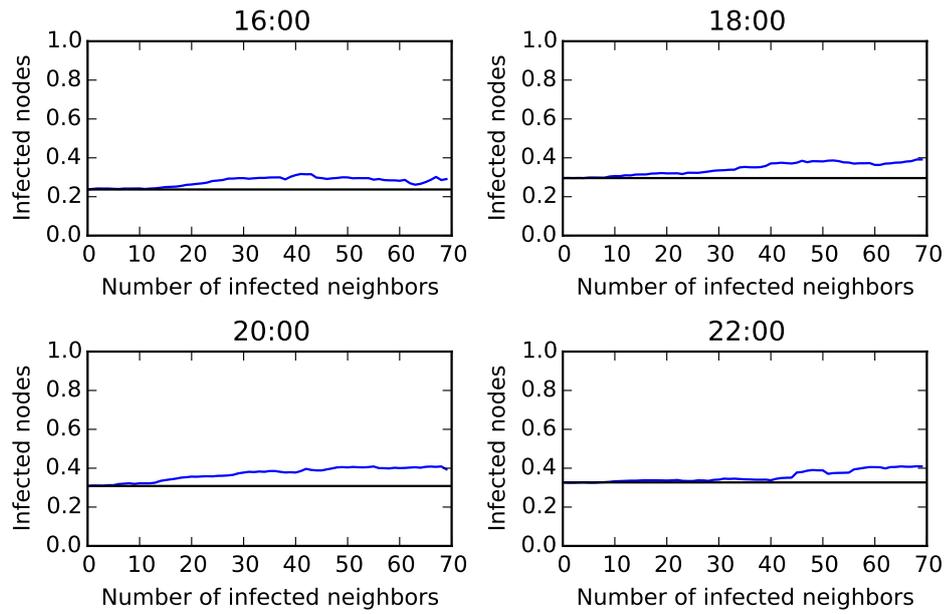


Figure B.6: Minimum number of infected third degree neighbors versus percentage of infected nodes for type 1 snapshots.

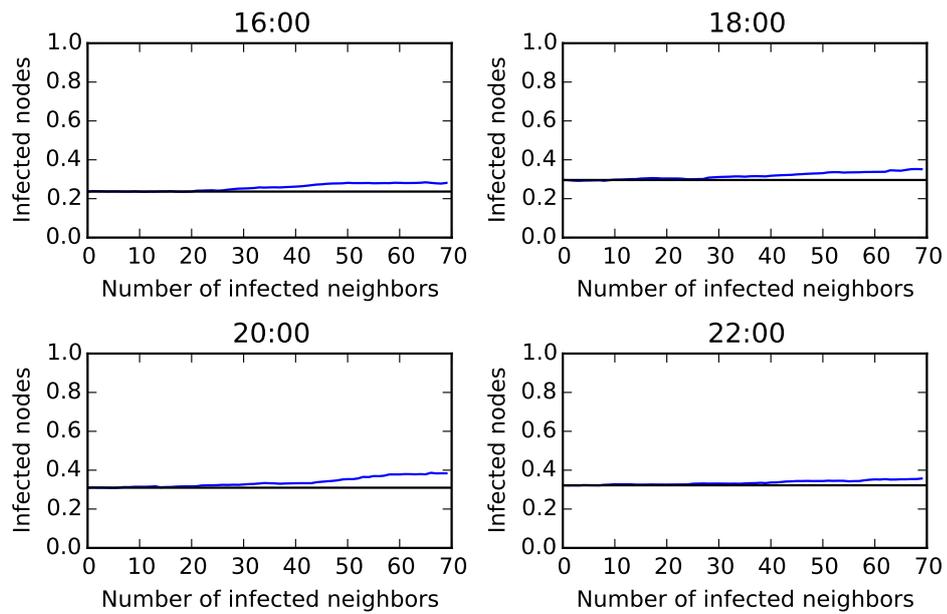


Figure B.7: Minimum number of infected fourth degree neighbors versus percentage of infected nodes for type 1 snapshots.

### B.3. Neighbor analysis 15-08-2016, Type 2, Degree 2

Table B.3: Number of 2nd degree neighbors and percentage of neighbors infected for type 2 snapshots.

Time	Nodes	Number	Number of neighbors			Percentage infected		
			<i>Md</i>	$\sigma$	IQR	<i>Md</i>	$\sigma$	IQR
16:00	All	3,312	510	326.3	284.0, 758.0	20.7%	6.7%	18.1%, 23.2%
	Susceptible	2,570	516	331.1	289.0, 773.8	20.3%	5.8%	17.7%, 22.6%
	Infected	742	485	307.6	260.0, 723.3	22.0%	8.8%	19.5%, 25.1%
18:00	All	3,338	512	338.4	287.0, 751.0	26.1%	6.5%	23.3%, 28.7%
	Susceptible	2,433	509	347.5	268.0, 749.0	25.7%	6.0%	23.0%, 28.2%
	Infected	905	523	312.4	319.0, 756.0	27.2%	7.3%	24.4%, 30.0%
20:00	All	3,379	495	337.5	249.5, 765.0	28.1%	8.3%	24.7%, 31.3%
	Susceptible	2,366	499	344.1	240.0, 777.0	27.6%	7.3%	23.9%, 30.5%
	Infected	1,013	490	321.5	269.0, 740.0	29.4%	9.8%	26.1%, 32.9%
22:00	All	3,490	558	325.9	330.0, 807.0	30.5%	8.5%	25.8%, 35.2%
	Susceptible	2,303	587	333.1	342.0, 829.5	29.6%	7.7%	25.3%, 33.9%
	Infected	1,187	512	308.3	312.5, 754.0	32.8%	9.4%	27.1%, 38.0%

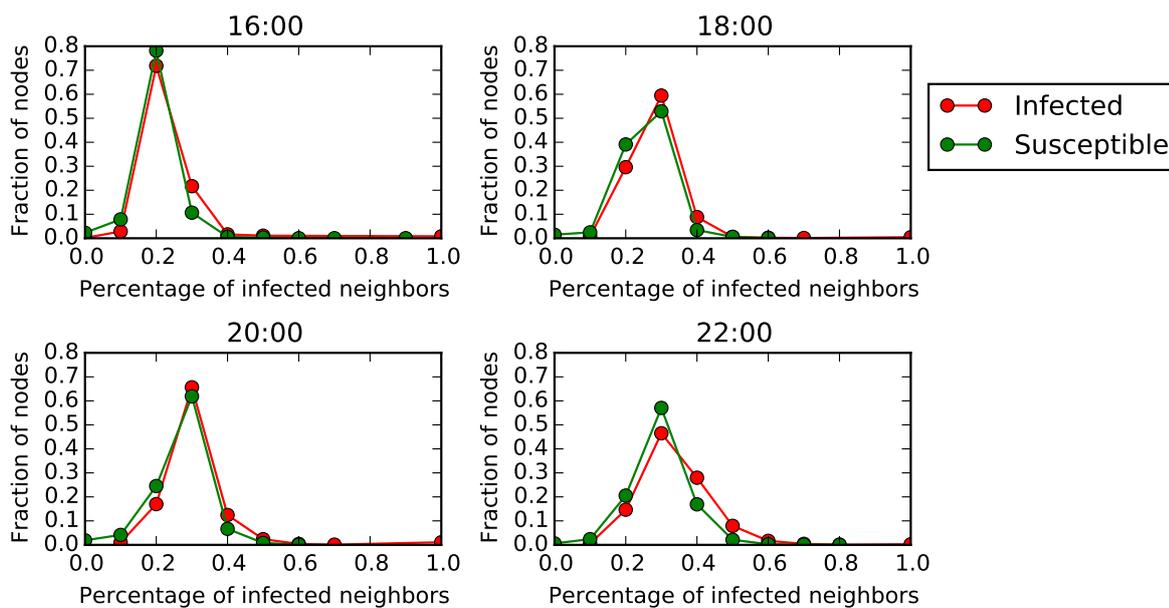


Figure B.8: The distribution of the percentage of infected 2nd degree neighbors of susceptible and infected nodes.

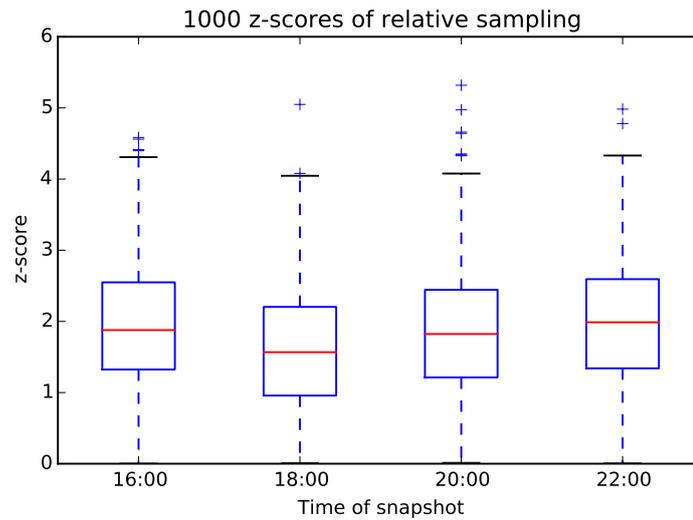


Figure B.9: Boxplots of the z-score results from the Mann-Whitney U tests for 2nd degree type 2 neighbors.

## B.4. Probability of being infected 15-08-2016, Type 2, Degree 2-3

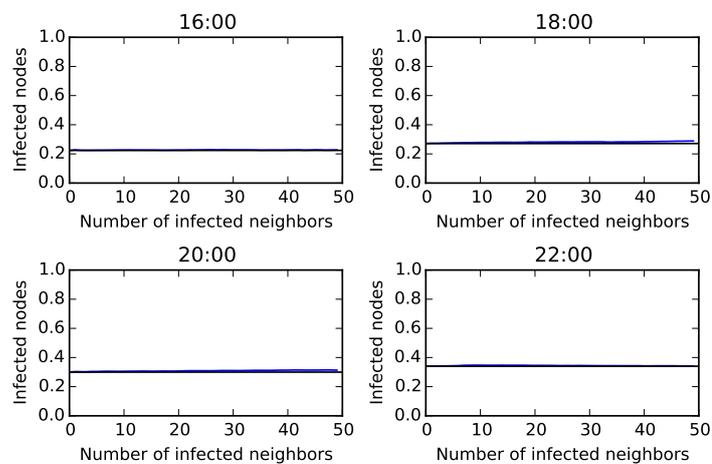


Figure B.10: Minimum number of infected second degree neighbors versus percentage of infected nodes for type 2 snapshots.

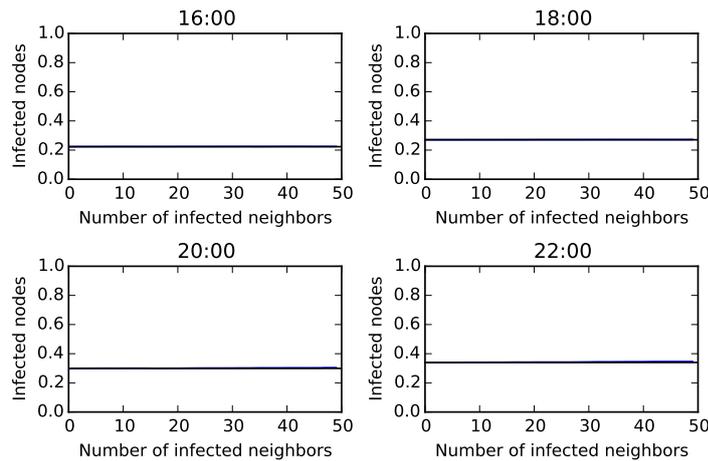


Figure B.11: Minimum number of infected third degree neighbors versus percentage of infected nodes for type 2 snapshots.

## B.5. Neighbor analysis 15-08-2016, Type 3, Degree 2

Table B.4: Number of 2nd degree neighbors and percentage of neighbors infected for type 3 snapshots.

Time	Nodes	Number	Number of neighbors			Percentage infected		
			<i>Md</i>	$\sigma$	IQR	<i>Md</i>	$\sigma$	IQR
16:00	All	3,053	564	738.8	221.0, 1055.0	19.3%	7.5%	17.0%, 21.4%
	Susceptible	2,357	535	706.9	208.0, 1010.0	19.1%	6.4%	16.5%, 21.3%
	Infected	696	650	823.2	298.8, 1309.8	19.8%	10.1%	18.4%, 22.2%
18:00	All	3,098	572	698.2	247.3, 1030.0	23.4%	8.0%	20.5%, 26.0%
	Susceptible	2,245	524	670.7	218.0, 961.0	23.0%	7.3%	20.0%, 25.7%
	Infected	853	744	745.8	368.0, 1201.0	24.1%	9.2%	21.9%, 26.8%
20:00	All	3,117	564	715.0	214.0, 1054.0	26.2%	11.2%	23.3%, 29.3%
	Susceptible	2,162	506	692.3	180.0, 1000.0	25.8%	9.0%	22.6%, 28.6%
	Infected	955	687	754.3	313.5, 1195.0	27.0%	14.3%	24.4%, 30.3%
22:00	All	3,188	663	725.5	331.0, 1204.5	30.0%	9.9%	25.4%, 33.7%
	Susceptible	2,099	646	710.8	327.5, 1168.0	29.3%	8.6%	24.6%, 33.0%
	Infected	1,089	693	751.5	345.0, 1269.0	31.4%	11.4%	27.0%, 34.8%

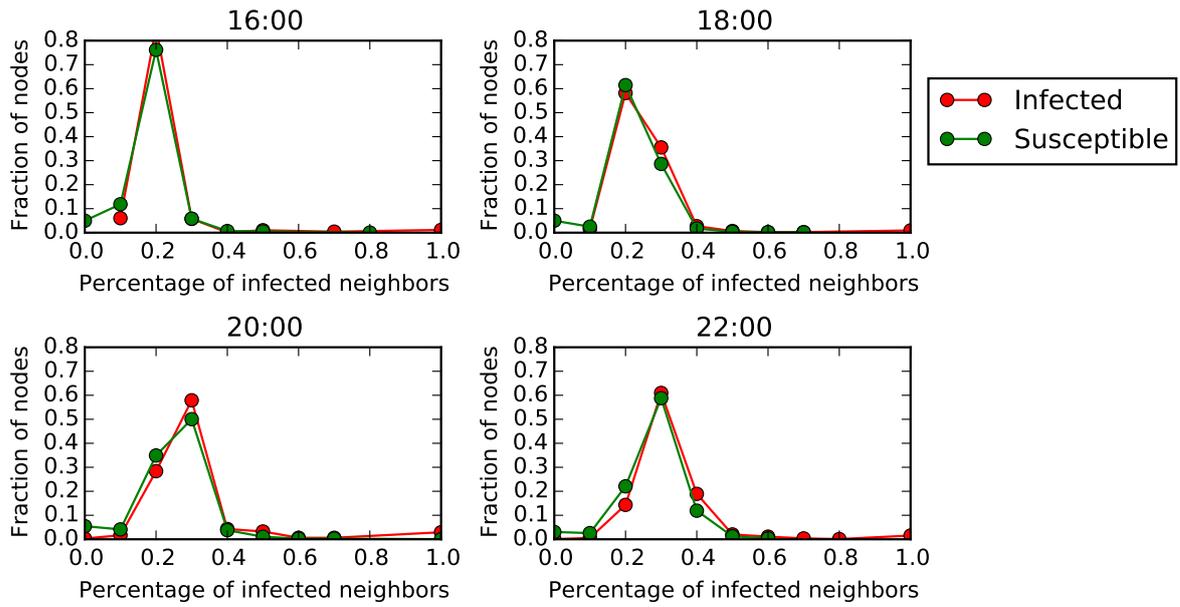


Figure B.12: The distribution of the percentage of infected 2nd degree neighbors of susceptible and infected nodes.

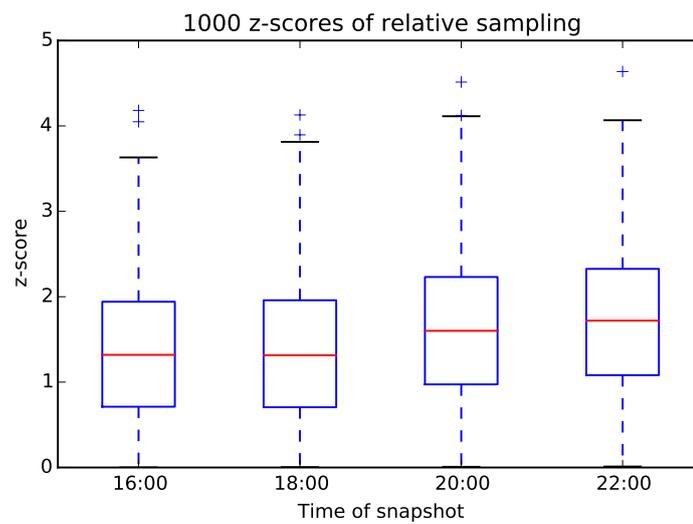


Figure B.13: Boxplots of the z-score results from the Mann-Whitney U tests for 2nd degree type 3 neighbors.

## B.6. Probability of being infected 15-08-2016, Type 3, Degree 2-3

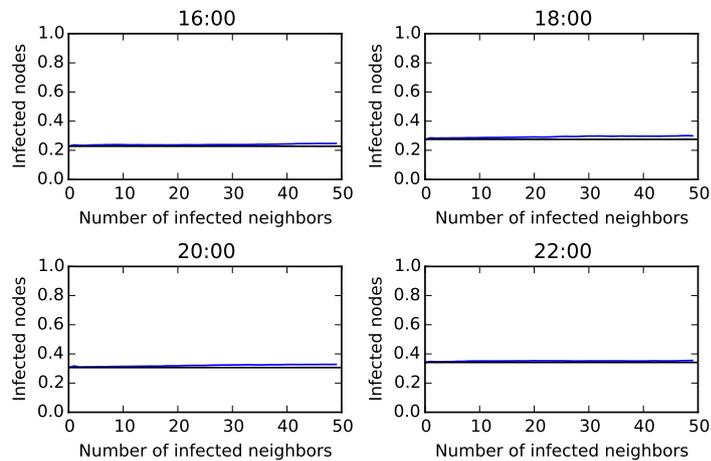


Figure B.14: Minimum number of infected second degree neighbors versus percentage of infected nodes for type 3 snapshots.

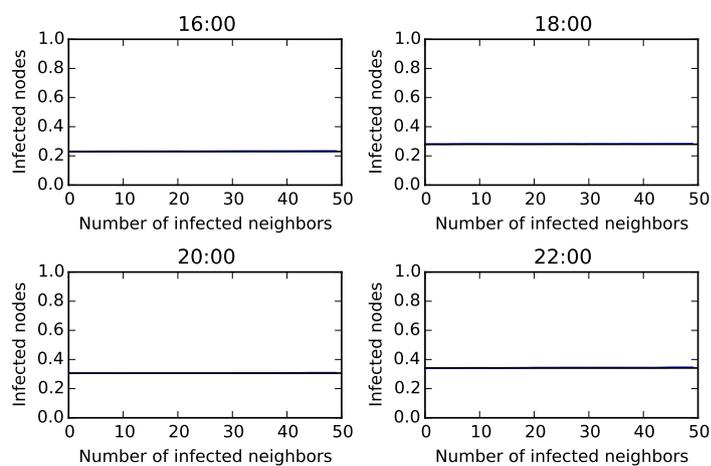


Figure B.15: Minimum number of infected third degree neighbors versus percentage of infected nodes for type 3 snapshots.

## B.7. Multiple days, Type 1, Degree 1-4

Table B.5: Z-scores calculated using the Mann-Whitney U test for first degree neighbors in type 1 snapshots on multiple days.

Day	Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
02-08	16:00	21	0.24%	79	0.11%	625	1.78	58%
	18:00	28	0.29%	72	0.17%	774	1.83	55%
	20:00	30	0.33%	70	0.18%	793	1.97	50%
	22:00	32	0.38%	68	0.20%	766	2.41	31%
07-08	16:00	14	0.14%	86	0.06%	492	1.15	80%
	18:00	22	0.20%	78	0.17%	720	1.17	78%
	20:00	27	0.25%	73	0.17%	811	1.38	75%
	22:00	28	0.27%	72	0.20%	833	1.35	73%
08-08	16:00	31	0.50%	69	0.11%	528	4.13	1%
	18:00	34	0.44%	66	0.15%	634	3.62	4%
	20:00	33	0.43%	67	0.17%	661	3.32	8%
	22:00	33	0.44%	67	0.19%	706	2.95	17%
15-08	16:00	24	0.22%	76	0.14%	760	1.26	78%
	18:00	30	0.27%	70	0.17%	837	1.63	62%
	20:00	31	0.33%	69	0.20%	813	1.95	51%
	22:00	33	0.33%	67	0.22%	808	2.19	42%
19-08	16:00	23	0.27%	77	0.13%	674	1.77	57%
	18:00	30	0.38%	70	0.17%	781	2.05	44%
	20:00	39	0.50%	61	0.25%	878	2.22	38%
	22:00	44	0.50%	56	0.33%	891	2.38	34%
23-08	16:00	10	0.05%	90	0.00%	375	0.97	86%
	18:00	11	0.00%	89	0.00%	431	0.71	95%
	20:00	11	0.00%	89	0.00%	430	0.71	92%
	22:00	11	0.00%	89	0.06%	425	0.75	94%
25-08	16:00	11	0.05%	89	0.00%	409	0.96	87%
	18:00	15	0.12%	85	0.06%	508	1.32	73%
	20:00	17	0.13%	83	0.09%	628	0.74	93%
	22:00	19	0.20%	81	0.14%	681	0.79	92%

Table B.6: Z-scores calculated using the Mann-Whitney U test for second degree neighbors in type 1 snapshots on multiple days.

Day	Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
02-08	16:00	20	0.23%	80	0.15%	543	2.23	40%
	18:00	28	0.31%	72	0.17%	604	3.12	8%
	20:00	30	0.37%	70	0.19%	621	3.24	9%
	22:00	31	0.35%	69	0.22%	663	3.04	14%
07-08	16:00	14	0.16%	86	0.10%	393	2.11	42%
	18:00	22	0.25%	78	0.17%	533	2.73	20%
	20:00	26	0.25%	74	0.17%	633	2.59	23%
	22:00	28	0.31%	72	0.19%	612	3.05	13%
08-08	16:00	31	0.50%	69	0.15%	402	5	0%
	18:00	35	0.50%	65	0.17%	515	4.53	0%
	20:00	33	0.43%	67	0.20%	488	4.55	0%
	22:00	33	0.38%	67	0.22%	684	3.09	13%
15-08	16:00	24	0.25%	76	0.15%	571	2.78	19%
	18:00	30	0.31%	70	0.17%	614	3.3	8%
	20:00	31	0.33%	69	0.18%	674	2.96	13%
	22:00	32	0.35%	68	0.23%	730	2.65	25%
19-08	16:00	23	0.40%	77	0.14%	453	3.56	5%
	18:00	30	0.47%	70	0.20%	599	3.42	5%
	20:00	39	0.48%	61	0.26%	723	3.31	7%
	22:00	44	0.50%	56	0.33%	819	2.87	17%
23-08	16:00	10	0.14%	90	0.03%	274	2.09	46%
	18:00	11	0.15%	89	0.05%	280	2.38	33%
	20:00	11	0.14%	89	0.06%	330	1.8	56%
	22:00	10	0.11%	90	0.07%	342	1.26	77%
25-08	16:00	11	0.11%	89	0.04%	306	2.07	45%
	18:00	15	0.18%	85	0.08%	400	2.35	34%
	20:00	17	0.18%	83	0.10%	462	2.27	38%
	22:00	19	0.23%	81	0.13%	481	2.55	27%

Table B.7: Z-scores calculated using the Mann-Whitney U test for third degree neighbors in type 1 snapshots on multiple days.

Day	Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
02-08	16:00	20	0.18%	80	0.18%	730	0.61	97%
	18:00	29	0.25%	71	0.19%	773	1.96	50%
	20:00	31	0.27%	69	0.20%	912	1.18	80%
	22:00	31	0.25%	69	0.23%	953	0.87	89%
07-08	16:00	14	0.14%	86	0.12%	495	1.07	81%
	18:00	22	0.21%	78	0.19%	720	1.16	81%
	20:00	26	0.20%	74	0.18%	840	0.96	85%
	22:00	27	0.25%	73	0.20%	763	1.73	58%
08-08	16:00	32	0.41%	68	0.17%	541	4.05	1%
	18:00	36	0.43%	64	0.22%	631	3.75	3%
	20:00	34	0.34%	66	0.23%	678	3.25	8%
	22:00	33	0.29%	67	0.23%	843	1.93	51%
15-08	16:00	24	0.23%	76	0.17%	728	1.49	69%
	18:00	30	0.25%	70	0.19%	814	1.78	56%
	20:00	31	0.30%	69	0.23%	865	1.53	68%
	22:00	32	0.31%	68	0.26%	869	1.62	64%
19-08	16:00	23	0.33%	77	0.18%	610	2.27	36%
	18:00	30	0.39%	70	0.25%	817	1.75	57%
	20:00	40	0.43%	60	0.33%	975	1.59	65%
	22:00	44	0.44%	56	0.33%	951	1.96	50%
23-08	16:00	10	0.09%	90	0.05%	373	0.9	87%
	18:00	11	0.09%	89	0.07%	422	0.75	93%
	20:00	11	0.10%	89	0.08%	412	0.86	87%
	22:00	10	0.09%	90	0.08%	380	0.82	92%
25-08	16:00	11	0.07%	89	0.06%	416	0.83	90%
	18:00	14	0.09%	86	0.09%	535	0.67	95%
	20:00	18	0.13%	82	0.12%	666	0.65	96%
	22:00	19	0.18%	81	0.13%	607	1.44	72%

Table B.8: Z-scores calculated using the Mann-Whitney U test for fourth degree neighbors in type 1 snapshots on multiple days.

Day	Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
02-08	16:00	20	0.17%	80	0.17%	700	0.86	91%
	18:00	29	0.23%	71	0.22%	923	0.81	91%
	20:00	31	0.20%	69	0.20%	978	0.68	96%
	22:00	31	0.22%	69	0.21%	964	0.79	93%
07-08	16:00	14	0.13%	86	0.12%	535	0.67	94%
	18:00	22	0.20%	78	0.18%	746	0.93	89%
	20:00	26	0.20%	74	0.18%	867	0.75	93%
	22:00	27	0.23%	73	0.20%	786	1.55	68%
08-08	16:00	33	0.37%	67	0.20%	586	3.81	2%
	18:00	37	0.40%	63	0.21%	665	3.58	3%
	20:00	36	0.38%	64	0.26%	768	2.76	20%
	22:00	33	0.32%	67	0.25%	764	2.5	31%
15-08	16:00	24	0.23%	76	0.19%	752	1.3	75%
	18:00	30	0.25%	70	0.22%	867	1.38	72%
	20:00	31	0.28%	69	0.27%	960	0.82	91%
	22:00	32	0.29%	68	0.27%	978	0.82	90%
19-08	16:00	23	0.29%	77	0.19%	629	2.1	43%
	18:00	31	0.38%	69	0.29%	909	1.2	81%
	20:00	40	0.43%	60	0.38%	1042	1.11	82%
	22:00	44	0.37%	56	0.33%	1082	1.04	84%
23-08	16:00	9	0.07%	91	0.07%	349	0.74	93%
	18:00	10	0.09%	90	0.09%	387	0.73	94%
	20:00	11	0.10%	89	0.09%	423	0.74	95%
	22:00	10	0.09%	90	0.09%	388	0.72	94%
25-08	16:00	11	0.08%	89	0.07%	426	0.71	94%
	18:00	14	0.08%	86	0.09%	528	0.74	94%
	20:00	18	0.14%	82	0.12%	652	0.77	92%
	22:00	19	0.16%	81	0.14%	663	0.94	87%

## B.8. Multiple days, Type 2, Degree 1-4

Table B.9: Z-scores calculated using the Mann-Whitney U test for first degree contacts in type 2 snapshots on multiple days.

Day	Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
02-08	16:00	19	0.22%	81	0.17%	542	2	48%
	18:00	26	0.27%	74	0.20%	657	2.4	33%
	20:00	31	0.35%	69	0.22%	676	2.94	15%
	22:00	35	0.42%	65	0.24%	626	3.7	4%
07-08	16:00	13	0.14%	87	0.12%	445	1.24	75%
	18:00	21	0.23%	79	0.19%	644	1.58	64%
	20:00	24	0.26%	76	0.22%	676	1.9	53%
	22:00	27	0.29%	73	0.23%	693	2.27	37%
08-08	16:00	31	0.44%	69	0.19%	497	4.27	1%
	18:00	34	0.40%	66	0.24%	601	3.79	2%
	20:00	35	0.39%	65	0.27%	715	3.06	11%
	22:00	36	0.40%	64	0.27%	723	3.08	14%
15-08	16:00	22	0.24%	78	0.18%	604	2.12	43%
	18:00	27	0.29%	73	0.23%	713	2.12	43%
	20:00	29	0.34%	71	0.24%	664	2.78	18%
	22:00	33	0.38%	67	0.24%	651	3.33	8%
19-08	16:00	23	0.26%	77	0.17%	580	2.51	30%
	18:00	29	0.33%	71	0.22%	652	2.87	18%
	20:00	38	0.43%	62	0.31%	748	3.06	13%
	22:00	44	0.50%	56	0.35%	770	3.21	9%
23-08	16:00	10	0.12%	90	0.08%	331	1.37	75%
	18:00	11	0.11%	89	0.08%	381	1.2	80%
	20:00	11	0.11%	89	0.09%	379	1.22	77%
	22:00	11	0.10%	89	0.09%	404	0.95	86%
25-08	16:00	10	0.11%	90	0.08%	341	1.26	75%
	18:00	14	0.14%	86	0.11%	452	1.5	68%
	20:00	17	0.20%	83	0.16%	554	1.39	75%
	22:00	19	0.22%	81	0.17%	569	1.76	58%

Table B.10: Z-scores calculated using the Mann-Whitney U test for second degree contacts in type 2 snapshots on multiple days.

Day	Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
02-08	16:00	19	0.21%	81	0.19%	634	1.2	78%
	18:00	26	0.27%	74	0.24%	722	1.89	53%
	20:00	31	0.34%	69	0.28%	758	2.32	34%
	22:00	35	0.38%	65	0.32%	781	2.58	24%
07-08	16:00	13	0.15%	87	0.14%	467	1.01	86%
	18:00	21	0.22%	79	0.21%	651	1.53	67%
	20:00	24	0.25%	76	0.24%	713	1.62	63%
	22:00	27	0.27%	73	0.26%	818	1.31	75%
08-08	16:00	31	0.36%	69	0.26%	667	3	13%
	18:00	34	0.37%	66	0.30%	748	2.72	22%
	20:00	35	0.40%	65	0.33%	782	2.57	28%
	22:00	36	0.38%	64	0.35%	963	1.36	72%
15-08	16:00	22	0.22%	78	0.20%	641	1.81	57%
	18:00	27	0.27%	73	0.26%	777	1.62	64%
	20:00	29	0.29%	71	0.28%	802	1.73	59%
	22:00	34	0.33%	66	0.30%	851	1.98	49%
19-08	16:00	23	0.24%	77	0.22%	672	1.75	58%
	18:00	29	0.31%	71	0.29%	787	1.84	56%
	20:00	38	0.42%	62	0.38%	878	2.13	43%
	22:00	44	0.48%	56	0.45%	987	1.71	60%
23-08	16:00	10	0.12%	90	0.08%	306	1.66	64%
	18:00	11	0.10%	89	0.09%	386	1.15	79%
	20:00	11	0.10%	89	0.10%	395	1.05	84%
	22:00	11	0.10%	89	0.10%	423	0.75	92%
25-08	16:00	10	0.11%	90	0.10%	351	1.14	78%
	18:00	14	0.14%	86	0.13%	449	1.53	68%
	20:00	17	0.17%	83	0.17%	597	1	84%
	22:00	19	0.19%	81	0.19%	675	0.84	90%

Table B.11: Z-scores calculated using the Mann-Whitney U test for third degree contacts in type 2 snapshots on multiple days.

Day	Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
02-08	16:00	19	0.19%	81	0.20%	633	1.24	78%
	18:00	25	0.26%	75	0.27%	803	1.08	83%
	20:00	31	0.31%	69	0.32%	965	0.78	91%
	22:00	35	0.35%	65	0.35%	1045	0.68	95%
07-08	16:00	13	0.14%	87	0.14%	469	1.03	85%
	18:00	21	0.21%	79	0.21%	737	0.81	89%
	20:00	24	0.25%	76	0.25%	830	0.69	95%
	22:00	27	0.28%	73	0.28%	896	0.71	95%
08-08	16:00	31	0.31%	69	0.33%	774	2.22	40%
	18:00	34	0.34%	66	0.35%	907	1.58	66%
	20:00	35	0.36%	65	0.37%	937	1.46	70%
	22:00	36	0.36%	64	0.37%	943	1.52	67%
15-08	16:00	22	0.22%	78	0.22%	714	1.23	80%
	18:00	27	0.27%	73	0.27%	864	0.96	84%
	20:00	29	0.30%	71	0.30%	939	0.7	95%
	22:00	34	0.34%	66	0.35%	976	1.08	85%
19-08	16:00	23	0.24%	77	0.24%	685	1.68	62%
	18:00	29	0.29%	71	0.30%	787	1.89	52%
	20:00	37	0.39%	63	0.40%	985	1.3	74%
	22:00	44	0.45%	56	0.45%	1100	0.92	89%
23-08	16:00	10	0.09%	90	0.10%	351	1.16	81%
	18:00	11	0.10%	89	0.11%	383	1.22	82%
	20:00	11	0.11%	89	0.11%	400	1.02	86%
	22:00	11	0.11%	89	0.11%	421	0.83	89%
25-08	16:00	10	0.10%	90	0.11%	362	1.05	85%
	18:00	14	0.14%	86	0.14%	483	1.23	79%
	20:00	17	0.18%	83	0.18%	617	0.84	91%
	22:00	19	0.20%	81	0.20%	689	0.74	94%

Table B.12: Z-scores calculated using the Mann-Whitney U test for fourth degree contacts in type 2 snapshots on multiple days.

Day	Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
02-08	16:00	19	0.19%	81	0.19%	688	0.72	93%
	18:00	25	0.24%	75	0.25%	810	1.02	85%
	20:00	31	0.29%	69	0.32%	827	1.81	57%
	22:00	35	0.35%	65	0.38%	888	1.8	57%
07-08	16:00	13	0.13%	87	0.13%	503	0.65	95%
	18:00	21	0.20%	79	0.21%	727	0.88	87%
	20:00	24	0.25%	76	0.25%	834	0.63	95%
	22:00	27	0.27%	73	0.27%	899	0.68	95%
08-08	16:00	31	0.27%	69	0.28%	950	0.9	87%
	18:00	34	0.28%	66	0.31%	892	1.68	61%
	20:00	35	0.31%	65	0.33%	904	1.69	60%
	22:00	36	0.35%	64	0.35%	1059	0.67	95%
15-08	16:00	22	0.24%	78	0.24%	777	0.68	94%
	18:00	27	0.26%	73	0.27%	890	0.75	93%
	20:00	29	0.30%	71	0.31%	946	0.64	96%
	22:00	34	0.36%	66	0.38%	988	0.98	85%
19-08	16:00	23	0.20%	77	0.21%	782	0.85	91%
	18:00	29	0.25%	71	0.25%	933	0.73	92%
	20:00	37	0.33%	63	0.34%	946	1.57	66%
	22:00	44	0.42%	56	0.43%	1095	0.95	87%
23-08	16:00	10	0.11%	90	0.12%	350	1.16	82%
	18:00	11	0.12%	89	0.13%	435	0.61	97%
	20:00	11	0.13%	89	0.13%	432	0.65	96%
	22:00	11	0.13%	89	0.13%	433	0.64	94%
25-08	16:00	10	0.11%	90	0.11%	389	0.72	94%
	18:00	14	0.14%	86	0.14%	534	0.69	94%
	20:00	17	0.18%	83	0.18%	632	0.68	95%
	22:00	19	0.20%	81	0.20%	696	0.65	96%

## B.9. Multiple days, Type 3, Degree 1-4

Table B.13: Z-scores calculated using the Mann-Whitney U test for first degree contacts in type 3 snapshots on multiple days.

Day	Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
02-08	16:00	20	0.18%	80	0.15%	629	1.47	68%
	18:00	26	0.24%	74	0.19%	727	1.85	55%
	20:00	31	0.30%	69	0.21%	748	2.4	34%
	22:00	35	0.36%	65	0.23%	729	2.96	17%
07-08	16:00	14	0.12%	86	0.11%	484	1.18	75%
	18:00	22	0.21%	78	0.18%	692	1.38	75%
	20:00	25	0.22%	75	0.19%	769	1.35	70%
	22:00	28	0.25%	72	0.22%	810	1.52	66%
08-08	16:00	31	0.32%	69	0.18%	623	3.33	8%
	18:00	34	0.32%	66	0.22%	745	2.74	22%
	20:00	35	0.32%	65	0.26%	849	2.09	45%
	22:00	36	0.34%	64	0.27%	901	1.8	56%
15-08	16:00	22	0.21%	78	0.18%	702	1.3	75%
	18:00	27	0.25%	73	0.20%	794	1.49	69%
	20:00	30	0.29%	70	0.23%	767	2.13	43%
	22:00	34	0.36%	66	0.26%	736	2.81	18%
19-08	16:00	24	0.20%	76	0.17%	728	1.49	67%
	18:00	29	0.26%	71	0.20%	789	1.83	56%
	20:00	38	0.36%	62	0.29%	869	2.2	40%
	22:00	45	0.42%	55	0.33%	893	2.39	32%
23-08	16:00	10	0.09%	90	0.07%	357	1.08	82%
	18:00	10	0.10%	90	0.08%	362	1.01	84%
	20:00	11	0.10%	89	0.09%	404	0.95	88%
	22:00	11	0.10%	89	0.09%	418	0.79	91%
25-08	16:00	11	0.10%	89	0.08%	383	1.18	79%
	18:00	14	0.14%	86	0.10%	443	1.58	65%
	20:00	17	0.17%	83	0.15%	590	1.06	82%
	22:00	19	0.20%	81	0.17%	637	1.17	80%

Table B.14: Z-scores calculated using the Mann-Whitney U test for second degree contacts in type 3 snapshots on multiple days.

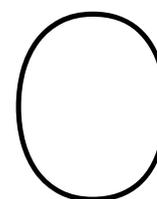
Day	Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
02-08	16:00	20	0.18%	80	0.18%	694	0.92	89%
	18:00	26	0.24%	74	0.22%	759	1.6	66%
	20:00	31	0.30%	69	0.24%	726	2.56	25%
	22:00	35	0.34%	65	0.29%	752	2.79	18%
07-08	16:00	14	0.13%	86	0.13%	518	0.85	91%
	18:00	22	0.20%	78	0.19%	682	1.47	70%
	20:00	25	0.23%	75	0.22%	723	1.72	61%
	22:00	28	0.26%	72	0.25%	872	1.05	86%
08-08	16:00	31	0.27%	69	0.24%	841	1.71	61%
	18:00	34	0.32%	66	0.28%	881	1.76	58%
	20:00	35	0.34%	65	0.30%	872	1.92	53%
	22:00	36	0.35%	64	0.33%	967	1.33	73%
15-08	16:00	22	0.20%	78	0.19%	690	1.41	74%
	18:00	27	0.24%	73	0.23%	805	1.41	73%
	20:00	30	0.27%	70	0.26%	839	1.59	64%
	22:00	34	0.31%	66	0.29%	878	1.78	59%
19-08	16:00	24	0.22%	76	0.21%	727	1.5	69%
	18:00	29	0.27%	71	0.26%	892	1.05	83%
	20:00	38	0.38%	62	0.35%	874	2.16	41%
	22:00	45	0.43%	55	0.41%	1007	1.6	66%
23-08	16:00	10	0.10%	90	0.09%	327	1.43	70%
	18:00	11	0.11%	89	0.09%	347	1.59	64%
	20:00	11	0.11%	89	0.10%	385	1.16	81%
	22:00	11	0.10%	89	0.10%	418	0.8	90%
25-08	16:00	11	0.10%	89	0.09%	384	1.17	81%
	18:00	14	0.13%	86	0.12%	462	1.4	72%
	20:00	17	0.16%	83	0.16%	552	1.43	71%
	22:00	19	0.18%	81	0.18%	673	0.86	90%

Table B.15: Z-scores calculated using the Mann-Whitney U test for third degree contacts in type 3 snapshots on multiple days.

Day	Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
02-08	16:00	20	0.17%	80	0.18%	679	1.08	82%
	18:00	26	0.24%	74	0.24%	819	1.14	79%
	20:00	31	0.29%	69	0.30%	926	1.08	82%
	22:00	35	0.32%	65	0.33%	965	1.26	78%
07-08	16:00	14	0.12%	86	0.13%	528	0.76	92%
	18:00	22	0.19%	78	0.19%	732	1.1	82%
	20:00	25	0.23%	75	0.23%	861	0.62	97%
	22:00	28	0.26%	72	0.26%	906	0.81	91%
08-08	16:00	32	0.28%	68	0.30%	877	1.57	65%
	18:00	34	0.31%	66	0.32%	903	1.6	63%
	20:00	35	0.32%	65	0.33%	1002	0.98	86%
	22:00	36	0.33%	64	0.34%	971	1.31	74%
15-08	16:00	22	0.20%	78	0.20%	760	0.84	92%
	18:00	27	0.25%	73	0.25%	886	0.79	92%
	20:00	30	0.28%	70	0.28%	961	0.68	94%
	22:00	34	0.32%	66	0.32%	1030	0.68	96%
19-08	16:00	24	0.21%	76	0.22%	733	1.47	69%
	18:00	29	0.27%	71	0.27%	869	1.24	75%
	20:00	38	0.35%	62	0.36%	1001	1.27	75%
	22:00	45	0.41%	55	0.42%	1060	1.24	78%
23-08	16:00	10	0.09%	90	0.09%	385	0.78	92%
	18:00	11	0.10%	89	0.10%	406	0.95	87%
	20:00	11	0.11%	89	0.11%	426	0.74	95%
	22:00	11	0.10%	89	0.11%	430	0.69	95%
25-08	16:00	11	0.10%	89	0.10%	428	0.73	94%
	18:00	14	0.13%	86	0.13%	489	1.18	79%
	20:00	17	0.17%	83	0.17%	627	0.75	95%
	22:00	19	0.19%	81	0.19%	691	0.72	92%

Table B.16: Z-scores calculated using the Mann-Whitney U test for fourth degree contacts in type 3 snapshots on multiple days.

Day	Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
02-08	16:00	20	0.14%	80	0.15%	626	1.51	69%
	18:00	26	0.19%	74	0.20%	816	1.15	81%
	20:00	31	0.23%	69	0.26%	861	1.56	66%
	22:00	35	0.26%	65	0.30%	839	2.16	41%
07-08	16:00	14	0.09%	86	0.10%	522	0.8	94%
	18:00	22	0.16%	78	0.16%	754	0.87	89%
	20:00	25	0.19%	75	0.21%	815	0.98	86%
	22:00	28	0.19%	72	0.21%	907	0.78	91%
08-08	16:00	32	0.22%	68	0.25%	849	1.77	58%
	18:00	34	0.22%	66	0.25%	915	1.51	70%
	20:00	35	0.25%	65	0.27%	951	1.35	74%
	22:00	36	0.26%	64	0.27%	1012	1.01	84%
15-08	16:00	23	0.20%	77	0.20%	783	0.85	90%
	18:00	27	0.22%	73	0.24%	872	0.89	89%
	20:00	30	0.24%	70	0.27%	921	0.98	86%
	22:00	34	0.26%	66	0.29%	991	0.96	86%
19-08	16:00	24	0.17%	76	0.17%	814	0.8	91%
	18:00	29	0.20%	71	0.22%	895	1.03	85%
	20:00	38	0.25%	62	0.27%	1024	1.1	83%
	22:00	45	0.33%	55	0.35%	1009	1.59	65%
23-08	16:00	10	0.09%	90	0.10%	377	0.85	89%
	18:00	11	0.10%	89	0.10%	425	0.72	94%
	20:00	11	0.12%	89	0.12%	425	0.72	95%
	22:00	11	0.12%	89	0.12%	429	0.68	95%
25-08	16:00	11	0.08%	89	0.09%	387	1.14	82%
	18:00	14	0.10%	86	0.11%	511	0.91	89%
	20:00	17	0.15%	83	0.15%	633	0.67	94%
	22:00	19	0.15%	81	0.16%	683	0.76	92%



# Validation results

## C.1. Snapshot time limit

Table C.1: Effect on network representation when changing type 2 snapshot time limit.

Scenario	Nodes	Edges	Avg. degree	% possible paths	Avg. shortest path	Strong components	Weak components
Null	3,481	150,169	43.1	95.6%	3.17	126	1
Alternative 1	6.3%	15.7%	9.0%	0.6%	-3.8%	-7.9%	0.0%
Alternative 2	2.1%	7.1%	5.1%	0.0%	-3.2%	0.8%	0.0%
Alternative 3	-2.0%	-14.1%	-12.3%	-5.0%	11.7%	65.1%	0.0%
Alternative 4	-11.8%	-48.0%	-41.1%	-22.9%	73.2%	378.6%	500.0%
Alternative 5	-24.1%	-71.8%	-62.9%	-78.7%	208.8%	761.9%	1800.0%

Table C.2: Effect on first degree neighbors when changing type 2 snapshot time limit.

Alternative	Nodes	Number	Number of neighbors			Percentage infected		
			$Md$	$\sigma$	IQR	$Md$	$\sigma$	IQR
Null	All	3,379	35	31.0	18.0, 59.0	27.0%	17.3%	16.7%, 37.5%
	Susceptible	2,366	36	32.4	18.0, 61.0	23.8%	15.3%	14.3%, 33.8%
	Infected	1,013	34	27.2	19.0, 54.0	33.7%	19.3%	23.4%, 46.4%
Alternative 1	All	3,627	39	32.0	21.0, 62.0	27.1%	16.5%	17.1%, 37.0%
	Susceptible	2,565	39	33.2	20.0, 64.0	24.7%	14.7%	15.6%, 33.3%
	Infected	1,062	37	28.7	21.0, 59.0	33.3%	18.4%	23.2%, 44.2%
Alternative 2	All	3,460	37	31.5	20.0, 61.0	27.3%	16.9%	16.7%, 37.5%
	Susceptible	2,431	38	32.7	19.0, 63.0	24.3%	15.0%	14.9%, 33.7%
	Infected	1,029	36	28.2	21.0, 57.0	33.3%	18.8%	23.5%, 45.3%
Alternative 3	All	3,231	30	29.7	15.0, 52.0	25.8%	18.2%	15.4%, 37.5%
	Susceptible	2,263	31	31.0	16.0, 56.0	23.1%	15.9%	13.5%, 33.3%
	Infected	968	28	26.2	15.0, 46.0	33.3%	20.6%	22.2%, 46.4%
Alternative 4	All	2,511	22	26.0	12.0, 40.0	24.1%	19.3%	12.6%, 34.6%
	Susceptible	1,772	23	27.0	11.0, 43.0	21.1%	17.0%	10.5%, 32.0%
	Infected	739	21	23.1	12.0, 35.0	30.4%	21.9%	20.0%, 43.8%
Alternative 5	All	1,655	19	21.8	10.0, 35.0	19.4%	17.3%	10.3%, 30.8%
	Susceptible	1,158	19	23.0	10.0, 37.0	17.3%	15.7%	9.7%, 27.3%
	Infected	497	19	18.5	10.0, 31.0	25.0%	19.4%	13.6%, 38.5%

Table C.3: Effect on Mann-Whitney U test when changing type 2 snapshot time limit.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
Null	29	0.34%	71	0.24%	662	2.79	20.0%
Alternative 1	29	0.33%	71	0.25%	679	2.67	22.9%
Alternative 2	29	0.33%	71	0.24%	662	2.8	19.5%
Alternative 3	29	0.33%	71	0.23%	680	2.66	22.5%
Alternative 4	29	0.30%	71	0.21%	684	2.63	25.1%
Alternative 5	30	0.25%	70	0.17%	788	1.97	49.4%

Table C.4: Effect on network representation when changing type 3 snapshot time limit.

Scenario	Nodes	Edges	Avg. degree	% possible paths	Avg. shortest path	Strong components	Weak components
Null	3,756	230,813	61.5	80.0%	3.02	684	1
Alternative 1	0.0%	14.9%	14.8%	10.9%	-8.6%	-59.2%	0.0%
Alternative 2	0.0%	4.9%	4.7%	3.6%	-3.3%	-19.6%	0.0%
Alternative 3	-0.1%	-13.3%	-13.3%	-11.0%	7.0%	56.9%	0.0%
Alternative 4	-0.5%	-45.3%	-45.2%	-38.9%	23.2%	204.4%	0.0%
Alternative 5	-1.7%	-68.3%	-67.8%	-62.3%	24.2%	314.6%	0.0%

Table C.5: Effect on first degree neighbors when changing type 3 snapshot time limit.

Alternative	Nodes	Number	Number of neighbors			Percentage infected		
			<i>Md</i>	$\sigma$	IQR	<i>Md</i>	$\sigma$	IQR
Null	All	3,123	53	81.6	17.0, 98.0	25.0%	16.8%	15.4%, 33.3%
	Susceptible	2,166	49	81.4	15.0, 94.0	23.1%	15.8%	13.3%, 31.6%
	Infected	957	62	81.6	22.0, 106.0	28.6%	17.8%	21.1%, 40.0%
Alternative 1	All	3,523	54	81.0	17.0, 101.0	27.0%	16.7%	17.0%, 35.7%
	Susceptible	2,475	51	80.5	16.0, 96.0	25.4%	15.9%	15.4%, 33.3%
	Infected	1,048	65	81.8	22.0, 110.0	30.9%	17.5%	22.9%, 42.5%
Alternative 2	All	3,255	53	81.4	17.0, 98.0	25.6%	16.8%	16.1%, 34.4%
	Susceptible	2,264	49	81.0	15.0, 94.0	24.3%	15.9%	14.0%, 32.6%
	Infected	991	63	81.8	22.0, 108.0	29.2%	17.6%	21.6%, 40.4%
Alternative 3	All	2,735	51	82.0	16.0, 97.0	23.2%	17.1%	14.6%, 32.7%
	Susceptible	1,903	48	83.3	15.0, 94.5	21.4%	15.9%	12.5%, 29.8%
	Infected	832	58	78.9	20.8, 105.0	27.3%	18.4%	20.0%, 39.8%
Alternative 4	All	1,727	48	85.2	16.0, 97.0	19.9%	16.2%	11.5%, 26.7%
	Susceptible	1,202	47	86.7	15.0, 96.0	18.5%	14.9%	10.3%, 24.6%
	Infected	525	55	81.5	20.0, 98.0	22.4%	18.0%	16.5%, 31.0%
Alternative 5	All	935	50	89.9	17.0, 102.0	14.3%	14.8%	8.3%, 21.2%
	Susceptible	644	48	94.3	16.0, 101.0	13.3%	13.2%	7.7%, 20.0%
	Infected	291	55	79.2	21.0, 103.5	16.9%	17.0%	11.2%, 26.7%

Table C.6: Effect on Mann-Whitney U test when changing type 3 snapshot time limit.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	H0?
Null	30	0.29%	70	0.23%	763	2.16	41.2%
Alternative 1	29	0.31%	71	0.25%	764	2.02	47.7%
Alternative 2	30	0.29%	70	0.24%	784	2	48.3%
Alternative 3	30	0.27%	70	0.21%	755	2.22	40.7%
Alternative 4	30	0.22%	70	0.19%	768	2.13	42.2%
Alternative 5	31	0.17%	69	0.13%	825	1.83	55.9%

## C.2. Delay definition

Table C.7: Effect on network representation when changing type 2 delay definition.

Scenario	Nodes	Edges	Avg. degree	% possible paths	Avg. shortest path	Strong components	Weak components
Null	3481	150169	43.1	0.956	3.17	126	1
Alternative 1	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Alternative 2	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Table C.8: Effect on first degree neighbors when changing type 2 definition of infected neighbors.

Alternative	Nodes	Number	Number of neighbors			Percentage infected		
			$Md$	$\sigma$	IQR	$Md$	$\sigma$	IQR
Null	All	3,379	35	31.0	18.0, 59.0	27.0%	17.3%	16.7%, 37.5%
	Susceptible	2,366	36	32.4	18.0, 61.0	23.8%	15.3%	14.3%, 33.8%
	Infected	1,013	34	27.2	19.0, 54.0	33.7%	19.3%	23.4%, 46.4%
Alternative 1	All	3,379	35	31.0	18.0, 59.0	34.2%	18.3%	23.1%, 45.6%
	Susceptible	2,366	36	32.4	18.0, 61.0	31.5%	16.5%	20.5%, 41.7%
	Infected	1,013	34	27.2	19.0, 54.0	42.1%	19.6%	31.3%, 54.7%
Alternative 2	All	3,379	35	31.0	18.0, 59.0	23.1%	16.2%	13.9%, 33.3%
	Susceptible	2,366	36	32.4	18.0, 61.0	20.4%	14.3%	12.1%, 29.8%
	Infected	1,013	34	27.2	19.0, 54.0	29.4%	18.2%	20.0%, 40.9%

Table C.9: Effect on first degree neighbors when changing type 2 definition of infected aircraft and infected neighbors .

Alternative	Nodes	Number	Number of neighbors			Percentage infected		
			$Md$	$\sigma$	IQR	$Md$	$\sigma$	IQR
Null	All	3,379	35	31.0	18.0, 59.0	27.0%	17.3%	16.7%, 37.5%
	Susceptible	2,366	36	32.4	18.0, 61.0	23.8%	15.3%	14.3%, 33.8%
	Infected	1,013	34	27.2	19.0, 54.0	33.7%	19.3%	23.4%, 46.4%
Alternative 1	All	3,379	35	31.0	18.0, 59.0	34.2%	18.3%	23.1%, 45.6%
	Susceptible	2,147	36	32.7	18.0, 62.0	31.0%	16.2%	20.0%, 41.3%
Alternative 2	Infected	1,232	34	27.5	19.0, 55.0	40.5%	19.6%	30.0%, 53.6%
	All	3,379	35	31.0	18.0, 59.0	23.1%	16.2%	13.9%, 33.3%
	Susceptible	2,494	36	32.3	18.0, 61.0	20.8%	14.5%	12.5%, 30.3%
	Infected	885	34	26.7	19.0, 54.0	29.4%	18.4%	20.0%, 41.3%

Table C.10: Effect on Mann-Whitney U test when changing type 2 delay definition.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
Null	29	0.34%	71	0.24%	661	2.8	21.5%
Alternative 1	36	0.41%	64	0.31%	763	2.8	16.8%
Alternative 2	26	0.29%	74	0.21%	639	2.54	26.0%

Table C.11: Effect on network representation when changing type 3 delay definition.

Scenario	Nodes	Edges	Avg. degree	% possible paths	Avg. shortest path	Strong components	Weak components
Null	3,756	230,813	61.5	0.8	3.02	684	1
Alternative 1	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Alternative 2	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Table C.12: Effect on first degree neighbors when changing type 3 definition of infected neighbors.

Alternative	Nodes	Number	Number of neighbors			Percentage infected		
			$Md$	$\sigma$	IQR	$Md$	$\sigma$	IQR
Null	All	3,123	53	81.6	17.0, 98.0	25.0%	16.8%	15.4%, 33.3%
	Susceptible	2,166	49	81.4	15.0, 94.0	23.1%	15.8%	13.3%, 31.6%
	Infected	957	62	81.6	22.0, 106.0	28.6%	17.8%	21.1%, 40.0%
Alternative 1	All	3,123	53	81.6	17.0, 98.0	31.3%	18.0%	20.4%, 41.0%
	Susceptible	2,166	49	81.4	15.0, 94.0	29.7%	17.3%	18.2%, 37.5%
	Infected	957	62	81.6	22.0, 106.0	34.6%	18.5%	26.7%, 45.2%
Alternative 2	All	3,123	53	81.6	17.0, 98.0	21.5%	15.8%	13.1%, 29.4%
	Susceptible	2,166	49	81.4	15.0, 94.0	20.0%	14.7%	11.1%, 27.6%
	Infected	957	62	81.6	22.0, 106.0	24.6%	17.1%	17.9%, 35.6%

Table C.13: Effect on first degree neighbors when changing type 3 definition of infected aircraft and infected neighbors .

Alternative	Nodes	Number	Number of neighbors			Percentage infected		
			$Md$	$\sigma$	IQR	$Md$	$\sigma$	IQR
Null	All	3,123	53	81.6	17.0, 98.0	25.0%	16.8%	15.4%, 33.3%
	Susceptible	2,166	49	81.4	15.0, 94.0	23.1%	15.8%	13.3%, 31.6%
	Infected	957	62	81.6	22.0, 106.0	28.6%	17.8%	21.1%, 40.0%
Alternative 1	All	3,123	53	81.6	17.0, 98.0	31.3%	18.0%	20.4%, 41.0%
	Susceptible	1,954	49	82.3	15.0, 94.0	28.9%	17.2%	17.6%, 37.5%
	Infected	1,169	60	80.1	21.0, 104.0	34.0%	18.5%	26.5%, 44.4%
Alternative 2	All	3,123	53	81.6	17.0, 98.0	21.5%	15.8%	13.1%, 29.4%
	Susceptible	2,287	49	82.6	15.0, 94.0	20.0%	14.9%	11.5%, 27.9%
	Infected	836	63	78.3	22.0, 106.3	25.0%	17.0%	18.0%, 36.5%

Table C.14: Effect on Mann-Whitney U test when changing type 3 delay definition.

Time	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	% accepted H0
Null	30	0.29%	70	0.23%	767	2.13	43.0%
Alternative 1	37	0.34%	63	0.29%	879	2.05	45.2%
Alternative 2	26	0.25%	74	0.20%	722	1.89	52.6%

### C.3. Validation similar days

Table C.15: Mann-Whitney U test results for second degree neighbors of similar days.

Connection	Date	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	H0?
Type 1	2016-8-15	30	30.3%	70	20.0%	652	3	11%
	2016-12-11	29	26.8%	71	19.1%	729	2.29	37%
	2017-4-2	28	33.3%	72	19.5%	594	3.18	10%
Type 2	2016-8-15	29	29.4%	71	27.6%	799	1.75	58%
	2016-12-11	28	26.0%	72	24.2%	749	2	49%
	2017-4-2	27	27.4%	73	24.3%	713	2.12	44%
Type 3	2016-8-15	30	27.0%	70	25.8%	851	1.5	67%
	2016-12-11	28	24.2%	72	22.9%	809	1.54	68%
	2017-4-2	27	23.9%	73	21.7%	756	1.79	57%

Table C.16: Mann-Whitney U test results for first degree neighbors of similar days.

Connection	Date	$n_i$	$Md_i$	$n_s$	$Md_s$	$Md_U$	$Md_Z$	H0?
Type 1	2016-8-15	30	26.8%	70	23.0%	861	1.42	71%
	2016-12-11	29	23.3%	71	22.0%	929	0.76	93%
	2017-4-2	29	28.6%	71	23.0%	820	1.59	65%
Type 2	2016-8-15	29	29.9%	71	30.2%	937	0.71	95%
	2016-12-11	28	26.9%	72	27.3%	902	0.83	91%
	2017-4-2	27	27.4%	73	28.0%	850	1.07	81%
Type 3	2016-8-15	30	27.9%	70	27.8%	965	0.65	97%
	2016-12-11	28	25.7%	72	26.0%	909	0.77	93%
	2017-4-2	27	24.5%	73	25.3%	860	0.98	85%

# Bibliography

- Aguirregabiria, V. and Ho, C. Y. A dynamic game of airline network competition: Hub-and-spoke networks and entry deterrence. *International Journal of Industrial Organization*, 28(4):377–382, 2010. ISSN 01677187. doi: 10.1016/j.ijindorg.2010.03.003.
- AhmadBeygi, S., Cohn, A., Guan, Y., and Belobaba, P. Analysis of the potential for delay propagation in passenger airline networks. *Journal of Air Transport Management*, 14(5):221–236, 2008. ISSN 09696997. doi: 10.1016/j.jairtraman.2008.04.010.
- Albert, R. and Barabasi, A. L. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97, 2002. ISSN 1478-3967. doi: 10.1088/1478-3967/1/3/006.
- Albert, R., Jeong, H., and Barabasi, A. L. Error and attack tolerance of complex networks. *Nature*, 406(6819):378–382, 2000.
- Baden, W., Dearmon, J., Kee, J., and Smith, L. Assessing Schedule Delay Propagation in the National Airspace System. 2003.
- Bagler, G. Analysis of the airport network of India as a complex weighted network. *Physica A: Statistical Mechanics and its Applications*, 387(12):2972–2980, 2008. ISSN 03784371. doi: 10.1016/j.physa.2008.01.077.
- Ball, M. ., Barnhart, C., Dresner, M., Neels, K., Odoni, A., Peterson, E., Sherry, L., Trani, A., Zou, B., and Hansen, M. Total Delay Impact Study – A comprehensive assessment of the cost and impacts of flight delay in the United States. page 91, 2010.
- Barabasi, A. L. and Albert, R. Emergence of Scaling in Random Networks. *Science*, 286(October): 509–512, 1999. ISSN 00368075. doi: 10.1126/science.286.5439.509.
- Barrat, A., Barthélemy, M., Pastor-Satorras, R., and Vespignani, A. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3747–3752, 2004. ISSN 0027-8424. doi: 10.1073/pnas.0400087101.
- Baspinar, B. and Koyuncu, E. A Data-Driven Air Transportation Delay Propagation Model Using Epidemic Process Models. *International Journal of Aerospace Engineering*, 2016, 2016. doi: 10.1155/2016/4836260.
- Bavelas, A. Communication Patterns in Task-Oriented Groups. *The Journal of the Acoustical Society of America*, 22(6):725–730, 1950. doi: 10.1121/1.1906679.
- Beatty, R., Hsu, R., Berry, L., and Rome, J. Preliminary evaluation of flight delay propagation through an airline schedule. *Air Track Control Quarterly*, 7(December):259–270, 1999. URL <http://tinyurl.com/c8o8wsv>.
- Belkoura, S., Cook, A., Peña, J. M., and Zanin, M. On the multi-dimensionality and sampling of air transport networks. *Transportation Research Part E: Logistics and Transportation Review*, 94:95–109, 2016. ISSN 13665545. doi: 10.1016/j.tre.2016.07.013. URL <http://dx.doi.org/10.1016/j.tre.2016.07.013>.
- Belobaba, P., Odoni, A., and Barnhart, C. *The Global Airline Industry*. John Wiley & Sons, Ltd, first edition, 2009. ISBN 9780470740774.
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D. U. Complex networks: Structure and dynamics. *Physics Reports*, 424(4-5):175–308, 2006. ISSN 03701573. doi: 10.1016/j.physrep.2005.10.009.

- Bonacich, P. Power and Centrality: A Family of Measures. *American Journal of Sociology*, 92(5): 1170–1182, 1987.
- Boswell, S. and Evans, J. Analysis of Downstream Impacts of Air Traffic Delay. Technical report, 1997.
- Campanelli, B., Fleurquin, P., Arranz, A., Etxebarria, I., Ciruelos, C., Eguíluz, V. M., and Ramasco, J. J. Comparing the modeling of delay propagation in the US and European air traffic networks. *Journal of Air Transport Management*, 56:1–7, 2016. ISSN 09696997. doi: 10.1016/j.jairtraman.2016.03.017. URL <http://www.sciencedirect.com/science/article/pii/S0969699716301016>.
- Cardillo, A., Gómez-Gardeñes, J., Zanin, M., Romance, M., Papo, D., del Pozo, F., and Boccaletti, S. Emergence of network features from multiplexity. *Scientific Reports*, 3:1–6, 2013. ISSN 2045-2322. doi: 10.1038/srep01344. URL <http://www.nature.com/doifinder/10.1038/srep01344>.
- Chen, D. B., Gao, H., Lü, L., and Zhou, T. Identifying influential nodes in large-scale directed networks: The role of clustering. *PLoS ONE*, 8(10):1–10, 2013. ISSN 19326203. doi: 10.1371/journal.pone.0077455.
- Chen, D., Lu, L., Shang, M.-S., Zhang, Y.-C., and Zhou, T. Identifying influential nodes in complex networks. *Physica A: Statistical ...*, 391(4):1777–1787, 2012. ISSN 03784371. doi: 10.1016/j.physa.2011.09.017. URL <http://www.sciencedirect.com/science/article/pii/S0378437111007333>.
- Cheung, D. P. and Gunes, M. H. A Complex Network Analysis of the United States Air Transportation. *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 699–701, 2012. doi: 10.1109/ASONAM.2012.116. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6425687>.
- Christakis, N. a. and Fowler, J. H. The spread of obesity in a large social network over 32 years. *The New England journal of medicine*, 357(4):370–9, 2007. ISSN 1533-4406. doi: 10.1056/NEJMsa066082. URL <http://www.ncbi.nlm.nih.gov/pubmed/17652652>.
- Ciruelos, C., Arranz, A., Etxebarria, I., and Peces, S. Modelling Delay Propagation Trees for Scheduled Flights. *11th USA/Europe Air Traffic Management Research and Development Seminar*, 2015.
- Cohen, R., Erez, K., Ben-Avraham, D., and Havlin, S. Resilience of the Internet to random breakdowns. *Physical Review Letters*, 85(21):4626–4628, 2000. ISSN 00319007. doi: 10.1103/PhysRevLett.85.4626.
- Colizza, V. and Vespignani, A. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations. *Journal of Theoretical Biology*, 251(3):450–467, 2008. ISSN 00225193. doi: 10.1016/j.jtbi.2007.11.028.
- Colizza, V., Barrat, A., Barthelemy, M., Valleron, A. J., and Vespignani, A. Modeling the worldwide spread of pandemic influenza: Baseline case and containment interventions. *PLoS Medicine*, 4(1): 0095–0110, 2007. ISSN 15491277. doi: 10.1371/journal.pmed.0040013.
- Cook, G. N. A Review of History , Structure , and Competition in the U . S . Airline Industry. *Journal of Aviation/Aerospace Education & Research*, 7(1):33–42, 1996.
- Cook, G. N. and Goodwin, J. Airline Networks : A Comparison of Hub-and- Spoke and Point-to-Point Systems Airline Networks : A Comparison of Hub-and-Spoke and Point-to-Point Systems. *The Journal of Aviation/Aerospace Education & Research*, 17(2):51–60, 2008. URL <http://commons.erau.edu/jaaer/vol17/iss2/1>.
- Crucitti, P., Latora, V., and Porta, S. Centrality in networks of urban streets. *Chaos*, 16(1):1–4, 2006. ISSN 10541500. doi: 10.1063/1.2150162.
- Daley, D. and Kendall, D. Stochastic Rumours. *IMA Journal of Applied Mathematics*, 1(1):42–55, 1965.

- Daniel, J. I. and Harback, K. T. (When) Do hub airlines internalize their self-imposed congestion delays? *Journal of Urban Economics*, 63(2):583–612, 2008. ISSN 00941190. doi: 10.1016/j.jue.2007.04.001.
- Du, W.-B., Zhou, X.-L., Lordan, O., Wang, Z., Zhao, C., and Zhu, Y.-B. Analysis of the Chinese Airline Network as multi-layer networks. *Transportation Research Part E: Logistics and Transportation Review*, 89:108–116, 2016. ISSN 13665545. doi: 10.1016/j.tre.2016.03.009. URL <http://www.sciencedirect.com/science/article/pii/S1366554515300521>.
- Erdős, P. and Rényi, A. On random graphs. *Publicationes Mathematicae*, 6:290–297, 1959. ISSN 00029947. doi: 10.2307/1999405.
- Fleurquin, P., Ramasco, J. J., and Eguiluz, V. M. Systemic delay propagation in the US airport network. *Scientific reports*, 3:1159, 2013. ISSN 2045-2322. doi: 10.1038/srep01159.
- Freeman, L. C. A Set of Measures of Centrality Based on Betweenness, 1977. ISSN 00380431. URL <http://www.jstor.org/stable/3033543?origin=crossref>.
- Garas, A., Schweitzer, F., and Havlin, S. A k-shell Decomposition Method for Weighted Networks. *New Journal of Physics*, 14:083030, 2012. ISSN 1367-2630. doi: 10.1088/1367-2630/14/8/083030.
- Gilbert, E. N. Random Graphs. *Annals of Mathematical Statistics*, 30(4):1141–1144, 1959. ISSN 0003-4851. doi: 10.1214/aoms/1177706098.
- Goh, K.-I., Oh, E., Kahng, B., and Kim, D. Betweenness centrality correlation in social networks. *Physical Review E*, 67(1):017101, 2003. ISSN 1063-651X. doi: 10.1103/PhysRevE.67.017101. URL <http://link.aps.org/doi/10.1103/PhysRevE.67.017101>.
- Guida, M. and Maria, F. Topology of the Italian airport network: A scale-free small-world network with a fractal structure? *Chaos, Solitons and Fractals*, 31(3):527–536, 2007. ISSN 09600779. doi: 10.1016/j.chaos.2006.02.007.
- Guimerà, R., Mossa, S., Turtschi, A., and Amaral, L. The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proceedings of the National Academy of Sciences of the United States of America*, 102(22):7794–7799, 2005. ISSN 0027-8424. doi: 10.1073/pnas.0407994102. URL <http://www.pnas.org/content/102/22/7794>.
- Han, D. D., Qian, J. H., and Liu, J. G. Network topology and correlation features affiliated with European airline companies. *Physica A: Statistical Mechanics and its Applications*, 388(1):71–81, 2009. ISSN 03784371. doi: 10.1016/j.physa.2008.09.021. URL <http://dx.doi.org/10.1016/j.physa.2008.09.021>.
- ICAO. Chicago Convention on International Civil Aviation. 1944.
- Jetzki, M. The propagation of air transport delays in Europe. pages 1–113, 2009.
- Jia, T., Qin, K., and Shan, J. An exploratory analysis on the evolution of the US airport network. *Physica A: Statistical Mechanics and its Applications*, 413:266–279, 2014. ISSN 03784371. doi: 10.1016/j.physa.2014.06.067. URL <http://dx.doi.org/10.1016/j.physa.2014.06.067>.
- Kim, D. H., Eisenberg, D. A., Chun, Y. H., and Park, J. Network topology and resilience analysis of South Korean power grid. *Physica A: Statistical Mechanics and its Applications*, 465:Published online, 2016. ISSN 03784371. doi: 10.1016/j.physa.2016.08.002. URL <http://linkinghub.elsevier.com/retrieve/pii/S0378437116305180>.
- Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E., and Makse, H. A. Identification of influential spreaders in complex networks. *Nature Physics*, 6(11):888–893, 2010. ISSN 1745-2473. doi: 10.1038/nphys1746. URL <http://arxiv.org/abs/1001.5285>.
- Kondo, A. Impacts of delay propagation on airline operations: Network vs. point-to-point carriers. *ICNS 2011 - Integrated Communications, Navigation and Surveillance Conference: Renovating the Global Air Transportation System, Proceedings*, pages 1–7, 2011. doi: 10.1109/ICNSURV.2011.5935340.

- Latora, V. and Marchiori, M. Efficient behavior of small-world networks. *Physical review letters*, 87(19): 198701, 2001. ISSN 0031-9007. doi: 10.1103/PhysRevLett.87.198701.
- Latora, V. and Marchiori, M. Economic small-world behavior in weighted networks. *European Physical Journal B*, 32(2):249–263, 2003. ISSN 14346028. doi: 10.1140/epjb/e2003-00095-5.
- Li, W. and Cai, X. Statistical Analysis of Airport Network of China. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 69(4 Pt 2):6, 2003. ISSN 1539-3755. doi: 10.1103/PhysRevE.69.046106.
- Liu, H. K., Zhang, X. L., and Zhou, T. Structure and external factors of Chinese city airline network. *Physics Procedia*, 3(5):1781–1789, 2010. ISSN 18753884. doi: 10.1016/j.phpro.2010.07.019. URL <http://dx.doi.org/10.1016/j.phpro.2010.07.019>.
- Lordan, O. *Airline Route Networks: A Complex Network Approach*. PhD thesis, 2014a.
- Lordan, O. Study of the full-service and low-cost carriers network configuration. *Journal of Industrial Engineering and Management*, 7(5):1112–1123, 2014b. ISSN 20130953. doi: 10.3926/jiem.1191.
- Mann, H. and Whitney, D. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, 18(1):50–60, 1947.
- Mao, G. and Zhang, N. Analysis of Average Shortest-Path Length of Scale-Free Network. *Journal of Applied Mathematics*, 2013, 2013.
- Moreno, Y., Nekovee, M., and Vespignani, A. Efficiency and reliability of epidemic data dissemination in complex networks. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 69(5):69–72, 2004. ISSN 1539-3755. doi: 10.1103/PhysRevE.69.055101. URL <http://link.aps.org/doi/10.1103/PhysRevE.69.055101>.
- Newman, M. E. J. Random graphs as models of networks. In *Handbook of Graphs and Networks: From the Genome to the Internet*, number 1, pages 34–68. 2002. ISBN 9783527602759. doi: 10.1002/3527602755.ch2. URL <http://doi.wiley.com/10.1002/3527602755>.
- Nowzari, C., Preciado, V. M., and Pappas, G. J. Analysis and Control of Epidemics: A Survey of Spreading Processes on Complex Networks. *IEEE Control Systems*, 36(1):26–46, 2016. ISSN 1066033X. doi: 10.1109/MCS.2015.2495000.
- Pastor-Satorras, R. and Vespignani, A. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200–3203, 2001. ISSN 00319007. doi: 10.1103/PhysRevLett.86.3200.
- Pastor-Satorras, R., Castellano, C., Van Mieghem, P., and Vespignani, A. Epidemic processes in complex networks. *Reviews of Modern Physics*, 87(3):925–979, 2015. ISSN 15390756. doi: 10.1103/RevModPhys.87.925.
- Perseguers, S., Lewenstein, M., Acín, A., and Cirac, J. I. Quantum random networks. *Nature Physics*, 6(7):539–543, 2010. ISSN 1745-2473. doi: 10.1038/nphys1665. URL <http://dx.doi.org/10.1038/nphys1665>.
- Pyrgiotis, N., Malone, K. M., and Odoni, A. Modelling delay propagation within an airport network. *Transportation Research Part C: Emerging Technologies*, 27:60–75, 2013. ISSN 0968090X. doi: 10.1016/j.trc.2011.05.017. URL <http://dx.doi.org/10.1016/j.trc.2011.05.017>.
- Rebollo, J. J. and Balakrishnan, H. Characterization and prediction of air traffic delays. *Transportation Research Part C: Emerging Technologies*, 44:231–241, 2014. ISSN 0968090X. doi: 10.1016/j.trc.2014.04.007. URL <http://dx.doi.org/10.1016/j.trc.2014.04.007>.
- Reggiani, A., Signoretti, S., Nijkamp, P., and Cento, A. Network measures in civil air transport: A case study of Lufthansa. *Lecture Notes in Economics and Mathematical Systems*, 613:257–282, 2009. ISSN 00758442. doi: 10.1007/978-3-540-68409-1-14.
- Sabidussi, G. The centrality index of a graph. *Psychometrika*, 31(4):581–603, 1966. ISSN 00333123. doi: 10.1007/BF02289527.

- Sachtjen, M. L., Carreras, B. A., and Lynch, V. E. Disturbances in a power transmission system. *Phys. Rev. E*, 61(5):4877–4882, 2000. ISSN 1063-651X. doi: 10.1103/PhysRevE.61.4877. URL <http://link.aps.org/abstract/PRE/v61/p4877>.
- Seidman, S. B. Network structure and minimum degree. *Social Networks*, 5(3):269–287, 1983. ISSN 03788733. doi: 10.1016/0378-8733(83)90028-X.
- Shields, R. Cultural Topology: The Seven Bridges of Konigsburg, 1736. *Theory, Culture & Society*, 29(4-5):43–57, 2012. ISSN 0263-2764. doi: 10.1177/0263276412451161. URL <http://tcs.sagepub.com/cgi/doi/10.1177/0263276412451161>.
- Stumpf, M. P. H., Wiuf, C., and May, R. M. Subnets of scale-free networks are not scale-free: sampling properties of networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102(12):4221–4, 2005. ISSN 0027-8424. doi: 10.1073/pnas.0501179102. URL <http://www.pnas.org/content/102/12/4221.full>.
- Travers, J. and Milgram, S. An Experimental Study of the Small World Problem. *Sociometry*, 32(4): 425–443, 1969.
- U.S. Department of Transportation. Air Travel Consumer Reports. Technical Report October, 2016. URL <https://www.transportation.gov/airconsumer/air-travel-consumer-reports>.
- Verma, T., Ara'ujo, N. a. M., and Herrmann, H. J. Revealing the structure of the world airline network. *Scientific Reports*, 4:5638, 2014. ISSN 2045-2322. doi: 10.1038/srep05638.
- Wang, J., Mo, H., Wang, F., and Jin, F. 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, 2011. ISSN 09666923. doi: 10.1016/j.jtrangeo.2010.08.012. URL <http://dx.doi.org/10.1016/j.jtrangeo.2010.08.012>.
- Watts, D. J. and Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature*, 393(6684): 440–442, 1998. ISSN 0028-0836. doi: Doi10.1038/30918. URL <http://www.nature.com/nature/journal/v393/n6684/abs/393440a0.html>.
- Welman, S., Williams, A., and Hechtman, D. Calculating Delay Propagation Multipliers for Cost-Benefit Analysis. (February), 2010.
- Wilson, R. J. *Introduction to Graph Theory*. Addison Wesley Longman Limited, 4th edition, 1996. ISBN 9788578110796. doi: 10.1017/CBO9781107415324.004.
- Wu, Z., Braunstein, L. A., Colizza, V., Cohen, R., Havlin, S., and Stanley, H. E. Optimal paths in complex networks with correlated weights: The worldwide airport network. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 74(5):1–9, 2006. ISSN 15393755. doi: 10.1103/PhysRevE.74.056104.
- Zanin, M. and Lillo, F. Modelling the air transport with complex networks: A short review. *European Physical Journal: Special Topics*, 215(1):5–21, 2013. ISSN 19516355. doi: 10.1140/epjst/e2013-01711-9.
- Zhang, J., Cao, X. B., Du, W. B., and Cai, K. Q. Evolution of Chinese airport network. *Physica A: Statistical Mechanics and its Applications*, 389(18):3922–3931, 2010. ISSN 03784371. doi: 10.1016/j.physa.2010.05.042. URL <http://dx.doi.org/10.1016/j.physa.2010.05.042>.