Analysis and Interpretation of Competitive Strategies and Consumer Behaviour Using ComScore Online Panel Data

Sahar Sameer Khogeer

A dissertation submitted to The University of Manchester for the degree of Master of Science in the Faculty of Engineering and Physical Science

School of Computer Science

2016
Declaration

No portion of the work referred to in the dissertation has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.
Intellectual Property Statement

i. The author of this dissertation (including any appendices and/or schedules to this dissertation) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this dissertation, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has entered into. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the dissertation, for example graphs and tables (“Reproductions”), which may be described in this dissertation, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this dissertation, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/display.aspx?DocID=24420), in any relevant Dissertation restriction declarations deposited in the University Library, and The University Library’s regulations (see http://www.library.manchester.ac.uk/about/regulations/_files/Library-regulations.pdf).
Acknowledgment

First and above all, I praise Allah the almighty for making my dream come true and giving me the strength to pursue it.

I would like to thank my supervisor Prof. C. Holland for his guidance throughout this project.

I am also so grateful to my mom, dad, and all my family and friends for their continuous love and support. And a special thanks to my sweet children for their patience and unconditional love.

And last but not least, I would like to thank my dear husband for always believing in me and helping me to shine.
Abstract

The high penetration of the Internet and the increasing confidence in its use have contributed to the growth of e-commerce, offering consumers multiple routes and influencing their behaviours in the buying process. These changes pose new challenges for businesses, which need to fully understand online consumer behaviour and search patterns and accordingly adjust their strategies to remain competitive in their markets.

The aim of this project was to analyse ComScore online panel data to map the search trajectory of consumers in the social network of the airline market. Studies which have used clickstream data in online marketing to understand consumer behaviour have focused on the individual level and not on the interactions between competitors. In the present research, the innovative application of Social Network theory in combination with market analysis to examine competitors’ relationships and the online customer journey provides a systemic approach for building digital marketing strategies.

The research findings show that, in the United States online airline market, a high degree of shared visitors switch among competitors’ websites, and online travel agents have more traffic than airlines in this competitive network. The findings indicate that there is a significant positive relationship between the number of visitors to airlines’ websites and the number of passengers. As well, attracting more visitors to online channels is especially important for airlines’ profit. Social Network Analysis found different search patterns among low-cost carrier and legacy carrier customers as low-cost carrier customers tend to search online travel agents’ websites more. Finally, the interaction levels in the findings indicate that customers’ crossover search behaviour is greater among competitors of the same type than competitors of different types.
These findings have implications for decision-makers and practitioners in the United States and similar markets worldwide. These findings provide a greater understanding of the competitors’ network, enabling decision-makers to better comprehend their customers’ search behaviour and to identify influential relationships and potential strategic partnerships and collaborations.
# Table of Contents

Declaration .......................................................................................................................... 2
Intellectual Property Statement .......................................................................................... 3
Acknowledgment .................................................................................................................. 4
Abstract .................................................................................................................................. 5
Table of Contents ..................................................................................................................... 7
List of Figures ......................................................................................................................... 9
List of Tables .......................................................................................................................... 10
1 Introduction .......................................................................................................................... 11
   1.1 Introduction and Motivation ......................................................................................... 11
   1.2 Aim and Objective ....................................................................................................... 13
   1.3 Description of the Project .......................................................................................... 14
2 Background ........................................................................................................................... 16
   2.1 Customer Journey ....................................................................................................... 16
   2.2 Online Search Behaviour ............................................................................................ 19
   2.3 Panel Data ................................................................................................................... 20
   2.4 ComScore ..................................................................................................................... 21
   2.5 Social Network Analysis ............................................................................................. 22
   2.6 Online Travel Agents ................................................................................................. 24
3 Research Methods .............................................................................................................. 25
   3.1 Research Design ......................................................................................................... 26
   3.2 Methodology .............................................................................................................. 27
      3.2.1 Market Share Analysis ........................................................................................ 27
      3.2.2 Online Market Share Analysis ........................................................................... 28
      3.2.3 Social Network Analysis .................................................................................. 28
   3.3 Data ............................................................................................................................ 30
   3.4 Evaluation ................................................................................................................... 33
4 Data Analysis ...................................................................................................................... 34
   4.1 The Online Travel Market in the USA ....................................................................... 34
      4.1.1 US Airlines Overview ....................................................................................... 38
      4.1.2 Online Travel Market Overview ....................................................................... 43
      4.1.3 Online Market Shares and Competitive Strategies .......................................... 46
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2</td>
<td>ComScore Data Analysis</td>
<td>51</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Descriptive Analysis</td>
<td>51</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Customer Journey Analysis</td>
<td>53</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Airlines Performance Analysis</td>
<td>59</td>
</tr>
<tr>
<td>4.3</td>
<td>Social Network Analysis</td>
<td>63</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Adjacency Matrix</td>
<td>64</td>
</tr>
<tr>
<td>4.3.2</td>
<td>SNA Analysis</td>
<td>66</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Visualisation</td>
<td>75</td>
</tr>
<tr>
<td>5</td>
<td>Discussion and Evaluation</td>
<td>80</td>
</tr>
<tr>
<td>6</td>
<td>Conclusions</td>
<td>85</td>
</tr>
<tr>
<td>6.1</td>
<td>Managerial Contributions</td>
<td>87</td>
</tr>
<tr>
<td>6.2</td>
<td>Limitations and Future Work</td>
<td>88</td>
</tr>
</tbody>
</table>

References

Word Count (excluding references): 17,104
List of Figures

Figure 1 Model of online customer journey ................................................................. 19
Figure 2 Hierarchical research model ........................................................................... 25
Figure 3 Conceptual framework of the research project ............................................... 26
Figure 4 US Airlines annual capacity and RPM 2014 ................................................... 37
Figure 5 Systemwide (Domestic & International) airlines market share Q1 2014 .... 47
Figure 6 Airlines market shares for Q1 2014 ............................................................... 48
Figure 7 Distribution of incoming traffic by source type .............................................. 53
Figure 8 Share of unique visitors for February 2014 .................................................... 54
Figure 9 Correlation analysis between UV and average number of passengers ...... 56
Figure 10 Percent of online traffic via Logon/Logoff for the month of February 2014 from UV to Airline’s website .................................................................................. 57
Figure 11 Incoming traffic from Search Engines for the month of February, 2014 .. 58
Figure 12 Incoming traffic from search engines for the month of February, 2014 (normalised) .................................................................................................................. 59
Figure 13 Airlines conversion rate 2014 ....................................................................... 61
Figure 14 Source/Loss adjacency matrix (pilot) ............................................................. 64
Figure 15 The final source/loss 10 x 10 adjacency matrix .............................................. 66
Figure 16 In-degree and Out-degree Measures ............................................................... 69
Figure 17 Correlation analysis between In-degree Centrality and passengers ....... 70
Figure 18 Online traffic ‘Density by group’ matrix ....................................................... 71
Figure 19 Difference matrix ........................................................................................ 73
Figure 20 Initial visualisation random layout ............................................................... 76
Figure 21 Visualisation Scaling/ordination layout ......................................................... 77
Figure 22 Visualisation difference matrix ‘By group’ layout ....................................... 78
List of Tables

Table 1 Airlines ranked by Scheduled Domestic and International Enplanements... 38
Table 2 Forecast of the airlines’ online sales ................................................................. 43
Table 3 Airlines business models’ elements ................................................................. 49
Table 4 Revenue from fares (online) per UV February 2014 ....................................... 62
Table 5 Density by Group (Sum of tie strengths) ......................................................... 71
Table 6 Net Gain online traffic network ................................................................. 74
Table 7 Airlines’ visitors interaction with OTAs as adapted from difference matrix 83
1 Introduction

This dissertation project is centred around an exploration and analysis of ComScore data, using social network analysis and visualisation. It is divided into four chapters. First, the introduction will highlight the aim and objective of the project. Second, the background chapter will introduce the base information regarding the concepts involved in the research. Third, the research methods chapter will set out the research design, data and methodologies for conducting the analysis and the evaluation. Then, data analysis chapter will present the analyses perfumed guided by the research focus. After that the discussion chapter will explore and evaluate the results and their managerial implications. Finally, the conclusion chapter will highlight the project achievements and shed light on limitation and future work.

1.1 Introduction and Motivation

The internet has contributed to the continuous growth of e-commerce over the past decade; the increased adoption of the medium and confidence levels in its use among a wide range of users have created an environment in which online trading can flourish. The internet is considered an essential source of product and price information, gives consumers more options and allows vendors to interact with their customers on many levels by utilizing this additional distribution channel. This flexible shopping environment, which is available 24 hours a day, seven days a week, has made it easier for consumers to purchase products or services in the comfort of their own spaces with just a few clicks.

The characteristics of this distribution channel have facilitated the offering of multiple routes to consumers and influenced their behaviours in the buying process. However,
these changes also pose new challenges to businesses, which need to fully understand online consumer behaviour and search patterns and adjust their strategies accordingly, in order to remain competitive in the market.

The travel industry has been significantly influenced by this rapid growth, especially with the low-cost and high-penetration nature of the internet. The airline sector, in particular, is enjoying the benefits; as announced by the International Air Transport Association (IATA), revenues by the end of 2015 were $727 billion, with a net profit of around $29.3 billion [1]. In 2015, online travel sales alone in the US generated approximately $208 billion [2]. Behind these high profits lies a dynamic and complex network of players, from consumers to travel intermediaries such as online travel agencies, social media sites, search engines, meta-search sites and airline websites.

The customer journey is “a description of modern multichannel buyer behaviour as consumers use different media to select suppliers, make purchases and gain customer support” [3]. More specifically, the online customer journey model shows that consumers’ searches involve many steps, iterations and interactions with various media [4].

ComScore is a leading commercial provider of business intelligence and consumer analytics, based on its worldwide panel of over two million online users in many markets. Its data has also been used in research into online retailing and the impact of comparison websites [5]. For this project, click stream data, obtained from ComScore, will be analysed using the innovative approach offered by Social Network Analysis (SNA) theory. SNA has been applied in many disciplines, including sociology, psychology, anthropology, economics, information sciences and communication.
Many studies on clickstream data to understand consumer behaviour focus on the individual level and not on the interactions between competitors [6]. It would be interesting, therefore, to map the customer’s journey through this complex network. There are also benefits to exploring the structure of the social network formed by ComScore data and studying the current and future impact on consumers’ behaviours and businesses’ performances. Such mapping would become a powerful enabler of successful strategies by the various airline sector businesses.

1.2 Aim and Objective

The aim of this project is to analyse ComScore online panel data to map the search trajectory of consumers in the airline market’s social network. It is hoped that the findings will contribute to managerial implications in the context of online marketing.

The objective is to study and analyse the network’s structure and the served market in order to address the following questions:

1. *What is the structure of the social network, as defined by the consumers’ search patterns?*

2. *What can the structure tell us about the following:*
   a) *The performance of the online competitors?*
   b) *The consumers’ search behaviour?*

3. *What is the interaction between the online travel agents and competitor airlines?*

Through an analysis of the network structure and by relating it to market shares and sales data, questions one and two will provide a deeper understanding of the online
customer journey (see Figure 1). This may help the airline companies and online intermediaries to improve their services, form strategic partnerships, introduce specific loyalty programs suited for the broader consumers and develop effective marketing campaigns, all of which could contribute to greater revenues.

Question three will identify interesting players that should be explored further by analysing their online marketing strategies. This could help weaker players to improve their strategies and stronger players to recognise the effective ones.

1.3 Description of the Project

The outcome of this project is a management report and a network visualisation. The report will consist of the served market overview, data collection and explanation, comprehensive analysis, results and interpretations, and managerial implications. The visualisation will help concerned organisations to understand their position in the network. This may be very useful in identifying opportunities for strategic alliances and achieving a competitive advantage.

SNA theory will be used to conduct data analysis. Specifically, online traffic source and loss data will be acquired to construct the social network in order to study online consumers’ search behaviour and its relation to other actors’ performance. Primary market data will be collected and analysed to measure the performance of those involved and to identify their competitive strategies.

The focus of the project will be on the US airlines market, which is the largest worldwide and among the most well-established online markets. This makes it a good indicator of online consumers’ behaviours and interactions with other actors in the network.
UCINET 6 and NetDraw are the tools that have been selected to analyse and visualise the social network, as built from the abovementioned dataset.
2 Background

This chapter provides an overview of specific topics that are relevant to the project. First, the marketing concepts of ‘customer journey’ and ‘online search behaviour’ will be explained. Next, in order to better understand the data source and collection method that is used in this project, the panel data will be discussed, followed by an overview of the company ComScore. Then, as an important player in the travel industry’s online market, it is also important to provide an explanation of the business model and services provided by the online travel agent. Finally, an overview of the SNA, the main methodology used for this project, and its background and terminology will be introduced.

2.1 Customer Journey

“The customer journey represents different touchpoints that characterise a person’s interaction with a brand, product or service of interest” [7]. This statement could be said to describe the entire structure of engagement levels from one brand or product to the next. During this journey, a mixture of feelings, impressions and preferences is accumulated in the customer’s mind, which will affect their conscious and subconscious purchasing decisions. From start to end, this journey can be described through different, often sequential, stages; the number of these stages varies in the marketing literature, but all reports are similar in context. According to David Clark [7], the customer journey and experience stages are as follows:

- Brand/product awareness. This is the first stage of the journey, wherein the potential customer becomes aware of a brand or product and starts formulating impressions and thoughts about it. It could be a trend that triggered their interest or the
realisation of a specific need or desire. Inputs affecting this stage could either be formal, such as advertisements, or informal, such as friends and social media.

- **Brand/product connection.** In this stage, the customer connects the need or desire to a specific brand or product range that is available. They also start to explore methods of procurement— in shop, online, pre-order, and so on—and endeavour to learn about other people’s experiences.

- **Brand/product evaluation.** Here, the customer thoroughly evaluates their options and starts to eliminate less appealing or irrelevant routes. In this process, they might read other customers’ reviews or use free trials or samples.

- **Shopping experience.** This is the stage at which the customer initiates the purchase transaction. This experience has an impact on the overall impression of the brand or product.

- **Out-of-box experience/deployment.** At this stage, the customer has completed the purchase already and is unpacking, launching or configuring the product. The level of difficulty of this process, and the extent to which it meets their expectations, will affect their views on the product and brand, which will probably be shared with others.

- **In-life product experience.** The customer, at this stage, is assessing the performance of the product during long-term use and possibly sharing his satisfaction levels with others. This is an important stage in terms of deciding whether to develop an affinity for the brand or to discontinue its use in the future.

- **Service and support.** How are these elements provided when needed? The customer decides whether this experience was satisfying and whether the issue was resolved smoothly and communicates the outcomes to others.
• Long-term commitment. Based on the accumulated impressions and opinions from the previous stages, the customer decides whether to engage in other interactions with the same brand and whether to encourage others to do so.

Having said this, in online retailing, Holland and Mandry believed that this linear framework does not fit with the ever-changing customer behaviours in the online channel. They proposed a revised model (see Figure 1) that includes iterative movements between different stages in the online customer journey. It also incorporates both online and offline factors that can influence the customer’s behaviour during the journey. As can be seen in the first step on the left-hand side, a combination of online advertising, pre-existing knowledge and needs and desires would lead the customer to the next step, that of building a set of potential choices of brands, products or service providers, termed the online consideration set. Once developed, this set can be constantly re-examined and revised, based on many factors, such as further search activities or interactions with other customers. This is illustrated by the iterative movement arrows between the steps of evaluation, development of the online consideration set and the first step. It should be noted that the model incorporates the customer’s online interactions in the purchase and service stages of the buying process. For example, after completing their purchase, customers share product reviews on social media or discuss post-purchase services with the supplier or other customers on various online platforms. The outcomes of such dialogues will feed back into their evaluation step and that of other customers and may influence their decision over whether to become a repeat customer or not [4].
2.2 Online Search Behaviour

The approach a consumer follows in their search for information is referred to as the search behaviour and may be either directed or undirected. As cited by Chaffey [3], Lewis and Lewis classified web users into five types, based on their reason for using the web and the distinct search behaviour they display.

- Directed information seekers: these users search for product, market or leisure information and tend to be familiar with the use of the web and expert in using search engines and directories.
- Undirected information seekers: also called ‘web surfers’, they browse websites and follow hyperlinks and tend to be more likely to click on banner advertisements.
- Directed buyers: these customers are looking to buy a specific product. They tend to visit intermediaries’ websites that compare product features and prices.
- Bargain hunters: these customers look for offers or take advantage of sales promotions.
- Entertainment seekers: these users look for enjoyment on the web, such as the playing of games.

These types of different behaviours are not mutually exclusive and could be exhibited by the same user in either one online session or a range of sessions.

In general, consumer online search behaviour can be classified into navigational, transactional, and informational purposes. Wendy W. Moe, as cited in [8], agreed with Lewis and Lewis in that the search depends largely on the online user’s purchase intent. For example, users with direct buying intentions exhibit a narrower search behaviour while users with low purchase intentions perform much broader search patterns.

2.3 Panel Data

“Panel data are typically thought of as information obtained by interviewing a sample of respondents, a panel, at two or more points in time” [9]. This is sometimes also referred to as longitudinal data, as it involves responses or observations to the same or similar queries, by the same entities, for more than one time period. Panel data has been commonly used in marketing research for multiple purposes, such as the study of consumer behaviour or to test the effectiveness of a promotional campaign. It has many advantages over traditional marketing research methods; for instance, by comparing responses from one time period with those from another, the precision of the survey’s
measurement can be increased. In addition, it allows for the observation of changes in an individual’s behaviour over a period of time. Finally, it is generally more accurate than cross-sectional data and enables more information to be collected [10].

Similarly, online panel data collection gathers data from online participants who agree to be part of a survey in exchange for some kind of compensation or reward. The survey owner usually retains key demographic information on all participants in order to ensure the selection of a representative sample [11] and the performance of further research-specific analyses, based on attributes such as cluster analysis, as needed. The literature shows that online panel data has been used in multiple sales-related research areas, such as social media and market orientation [11].

2.4 ComScore

ComScore Inc. is an American internet analytics company that describes itself as “a leading cross-platform measurement company that precisely measures audiences, brands and consumer behaviour everywhere” [12]. By capturing 1.9 trillion global interactions monthly, it is able to offer various analytical products to more than 3200 clients from 75 countries.

ComScore has introduced the Unified Digital Measurement™ methodology as a hybrid approach between census-based site analytics data and panel-based audience measurement data in order to overcome the usual vulnerability in digital measurements caused by the frequent inconsistencies when each approach is used separately. In order to cover 100% of a property’s audience, this approach combines weighted consumption with person-level metrics from the two million users of the ComScore global panel. In addition, to avoid inconsistencies caused by user configuration, such as cookie deletion, blocking or rejection, ComScore has developed an exclusive
methodology that combines panel and server-side metrics to calculate measures such as audience reach [13].

2.5 Social Network Analysis

“Networks are a way of thinking about social systems that focus our attention on the relationships among the entities that make up the system.” Networks can be seen in different cultural and natural structures, such as the neural networks in the brain, authoritarian (hierarchical) networks in organisations and food networks in ecologies. It is important to identify the position of an entity (known as an actor or node) in a network, as a generic hypothesis on network theory states that this position partly determines the constraints and opportunities this actor will encounter. Identifying the position helps to predict the actor’s performance, behaviour or beliefs [14].

Social network analysis (SNA) is a research stream that is concerned with the study and analysis of networks. Its foundations can be linked to the disciplines of sociometry and structural sociology [15]. Some traditional applications of SNA have been used to study family and community structures, group dynamics, social stratification and health issues. SNA helps researchers to discover the relationships (ties) between entities. It can also detect clusters and identify the positions and roles of entities. Therefore, the application of SNA in the marketing discipline has been found to be very beneficial, as it helps marketers to develop a more effective marketing strategy after understanding the different relationship patterns between consumers and/or the diffusion of information in a network [16]. In the past few decades, it has been applied in many areas in the field of marketing. Some examples include analyses of customer referral behaviour, investigations of the impact of word-of-mouth communication and
information sharing in new product development processes [15]. Social network analysis methodologies helped in understanding the purchase transition behaviour among consumers by considering brand switching patterns as networks [17]. In addition, networks were used in mapping multiple branding effects and answering branding questions either by looking within a single consumer associative network or across several networks [18].

There are three important elements in a social network: actors, ties and relationships. Actors are the entities being studied, and can be people, organisations or events. A relationship among actors is represented by a tie, which is a path between them. The tie could be described as strong or weak, depending on the strength of the relationship. Varying network characteristics can be caused by different relationships (Hanneman in [16]). The network structure graph is a visual representation of the social network. Where the actors are represented by nodes, the ties are lines connecting the associated nodes.

In recent years, with the global increase of internet penetration and rise in online communities and social media users, greater interest has been paid to SNA. The data for conducting the analyses can now be easily collected from electronic sources, unlike the traditional methods that relied on humans and were challenging for researchers. In addition, the advancement in computer science discipline and technologies have helped to open new doors for SNA applications. Business applications in the marketing industry have benefitted from this shift in SNA research as the outcome of an analysis can facilitate decisions regarding marketing activities and, more specifically, digital marketing, such as direct advertisements and social marketing [19].
2.6 Online Travel Agents

Online travel agents (OTAs) play a similar role to real-life travel agents, but on the web. That is to say, they are an intermediary between the customer and the travel service provider (airlines, accommodation, car rental and so on). OTAs provide the customer with an online platform to search for flights, compare fares and make a selection that is booked and ticketed by the OTA [20]. The OTA then receives a commission from the airline. OTAs can also generate revenues from banner advertisements and promotions [21].

An OTA provides multiple services for its online customers, typically in the following order: reservation information and recommendation services, reservation services and ticketing services. In the first step, the OTA collects the customer’s requirements and retrieves reservation information that meets those requirements. It may also combine this with recommendations based on the customer’s search history or profile or promotional offers. After the customer selects the most convenient flight, the OTA performs the reservation on their behalf. The customer confirms the selection by authorising the OTA to proceed with the ticketing service, which is linked to the commission, as mentioned previously [20]. Some of the major OTAs in the USA include Booking.com and Expedia.com.

This chapter has attempted to provide a brief summary of the literature relating to marketing theories of customer journey, online search behaviour, and social network analysis. In addition, it has described online panel data, followed by an overview of the company ComScore and online travel agents. The next chapter describes the research framework and discusses the procedures and methods used in this project.
3 Research Methods

This chapter highlights the research framework and the analytical methods and tools that are used in the project.

The hierarchical research model in Figure 2 was followed for this research. The uppermost level depicts the underlying theories, which are the consumer search, customer journey and social network analysis theories. The second layer is the operational construct for the research. In the main, the consumer’s search trajectory (search flow) will be measured using raw source/loss data from the ComScore report, which lies on the lowest level of the hierarchical model. The arrows indicate that each level relates to the level above.

![Hierarchical research model](image)

*Figure 2. Hierarchical research model*
3.1 **Research Design**

A mixed-method approach is employed in this research. It is broadly quantitative, followed by a qualitative interpretation of some strategies’ data on the airline market. In other words, quantitative analytical methods will be applied to the various data, as explained in Section 3.3, in order to understand the consumer behaviour in an online market context. Then, data from additional resources, mainly published materials such as company reports and websites, will be collected and qualitative methods will be applied to understand the marketing strategies and differences between competitors, followed by interpretations of the observations made. Finally, the correlation between both outcomes will be examined and the combination will be studied in terms of support for the customer journey theory.

![Conceptual framework of the research project](image)

*Figure 3. Conceptual framework of the research project*

---

1 The charts are for illustrative purposes only.
The conceptual research framework (see Figure 3) illustrates the analytical components involved in the study. The first component is the analysis of the market served. As the focus of the project is the US airlines market, the top six airline companies will be studied. A market share analysis will be conducted, based on the revenues and sales data. In addition, their digital marketing strategies will be compared. Next, an online market share analysis will be performed for these airlines and finally, a social network visualisation of these bodies, and other key players in the online travel market, will be constructed. SNA theory will be adopted in order to analyse the structure and relationships in this network. A comparative analysis will tie the three components together and demonstrate the significance of the findings. The ultimate goal is to reach an answer for the three research questions suggested previously.

3.2 Methodology

This section is concerned with the methodologies employed for this study.

3.2.1 Market Share Analysis

Given the simplest description of a company’s market share in the formula:

\[
\text{Company sales volume} = \text{Whole industry sales volume} \times \text{The company market share}
\]

For the purpose of this study, the number of passengers for each airline will be interpreted through the number of tickets sold, which will also be considered for the sales volume. The ‘whole industry sales volume’ will be calculated from the total sales of the six airlines analysed in the study, merely in order to create a comparable ground with the SNA.
Regarding calculation of the revenues, it is assumed to be linked to passenger miles. However, this is only one stream for generating revenue for an airline. Other streams may include service charges, cargo and overweight charges, advertisements, in-plane sales and so on; however, for the purpose of this analysis, the focus will be purely on sales, which is assumed to represent the number of passengers.

Other market data will be collected and analysed to identify the competitive strategies of those involved in the study. The sources of this data will be company published reports, company websites and published journal articles.

3.2.2 Online Market Share Analysis

The online market share (OMS) of a company can be defined as the ratio of the number of online visitors to that company’s website to the total number of online visitors among all competitors. The number of unique visitors (UV) from ComScore’s Media Metrix report will be used to calculate the OMS, as follows:

\[
OMS \times = \frac{UV \times}{\sum UV}
\]

3.2.3 Social Network Analysis

UCINET 6 has emerged as a powerful tool in studying social networks. It will be used for this project to analyse the social network. The data sets will be merged into a one data set. The outcome dataset will be used in creating the network data. The Network data can be represented in different ways depending on the application at hand. In other words, networks can be conceptualized mathematically in multiple ways, such as
graphs or matrices. An adjacency one-mode matrix\(^2\) will be constructed from the source/loss data. The number of unique visitors in source will represent the valued tie between the source (row) and destination (column). It will be assumed that this value represent the strength of the tie between the two actors. A value ‘0’ will be inserted in the matrix where no tie is present. The matrix is asymmetric as it represents a directed relation, from source to destination. This relationship can be reciprocated. It could be, for example, that there are online visitors reaching American Airlines website from Expedia.com, and Expedia’s website receiving online visitors from AmericanAirlines.com. It is worth mentioning that, by convention, the direction goes from the rows to the columns.

A whole network consists of all nodes (entities) in the dataset. In the case of this study it can be represented by 120 X 120 matrix. Where 120 is the number of all entities in the combined dataset. However, the network boundaries must be set according to the focus of this research. Hence, a new network will be constructed from a set of focal nodes comprising the six airlines and the major OTAs.

The outcome matrix may include some missing data which must be handled. Further examination of the best imputation approach must be conducted as it is recommended to try different approaches and compare the results of the analyses.

NetDraw is the software that will be used for visualising the network from the constructed matrix. The visualisation is called a network diagram, which consists of a set of points representing the nodes, and a set of lines representing the ties. The layout of the diagram will not be random. An ordination layout approach is more suitable due

\(^2\) Adjacency matrix has the same number of rows and columns. One-mode matrix means that the rows and columns both refer to the same single set of entities.
to the nature of the ties being valued. As such, the output diagram must clearly represent the strength of the ties by placing the nodes with the stronger ties closer to each other. Moreover, nodes attributes will be embedded in the colour, shape, size of the points in the diagram.

Network structure analysis is then conducted by calculating important centrality measures such as Degree centrality, which measures the number of ties that a node has [14].

### 3.3 Data

One major challenge in social network analysis tends to be the collection of information on the relationships between entities. Frequently, this is conducted via surveys or interviews, which are time-consuming and strenuous and usually result in a limited number of respondents compared to other data collection methods, such as online questionnaires. Another key challenge is the quality of the data collected. Many studies have indicated that information about the presence and strength of relationships, as obtained by questioning respondents, can be subject to substantial errors. As cited in the work of Kim and Srivastava [15], Marin pinpointed a number of reasons for the possible occurrence of such errors. Respondents tend to forget to mention certain relationships that could be important. Additionally, they may deliberately provide incorrect information or opt to omit details as a result of personal perceptions and interpretations. These types of errors may add bias to the data and the validity of the results of the social network analysis could be questionable.
The data used for the social network analysis in this project has certain advantages that enable it to overcome such challenges. First, it is obtained from reliable electronic sources provided by ComScore (for more details, see ComScore methodology). Second, it is collected from a panel of over two million online users.

The research data in this project is drawn from two main sources, one for each component explained in the conceptual research framework.

1. Statistical data from the Bureau of Transportation Statistics (BTS)\(^3\) is gathered. Specifically, raw data from the report “Passengers All Carriers - All Airports” \([22]\) is to be obtained. The number of passengers for each of the six airlines for the first quarter of the year 2014 will be extracted and used to calculate the market share.

2. Data from ComScore’s ‘Media Metrix Source/Loss Report’ is acquired\(^4\). The Media Metrix reports contain key metrics for measuring an industry’s online audience compositions. Six reports are obtained, one for each airline studied (US airlines: AA, Delta, JetBlue, SWA, United, US Airways) for the month of February 2014. For every airline, the ‘source’ measure represents the number of unique visitors for incoming traffic from multiple online media. Similarly, the ‘loss’ measure represents the number of unique visitors for outgoing traffic to multiple online media.

   a. Both source and loss data will be used to construct the social network for studying online consumers’ search behaviour and its relation to other actors’ performance.

\(^3\) BTS is a statistical agency of the U.S. Department of Transportation which provides extensive transportation statistics including air travel data. It is part of the Research and Innovative Technology Administration (RITA) \([58]\).

\(^4\) The acquired data is authorised for use within this project only and is not public.
b. This data will be used to calculate the OMS. The number of unique visitors for each of the six airlines will be used to perform the calculations.

c. The multiple online media sources and destinations are categorised into four types, as defined below:

- Airlines website: any airline website.
- OTAs: websites of OTAs providing travel information, reservations and ticketing services.
- Other travel: websites providing travel-related services, such as hotels and price comparison or meta-search engines.
- Non-travel: all other websites, such as social media sites, search engines, banks and so on.

In addition to these main sources, secondary data from other multiple sources will be collected and used in this project. These include but not limited to: journal articles, company annual reports, financial reports, government and private reports, and online news articles.

Data management and pre-processing will be performed in order to reach a compatible form with the analyses requirements. For example the formation of the adjacency matrix that will be used in the main analysis as mentioned previously under social network analysis methodology.
3.4 Evaluation

The results of this project should shed light on the theoretical and managerial implications of online consumer behaviour in the airline industry. The evaluation will be based mainly on two criteria:

1. How these results compare to the concepts and theories in the literature, and
2. The practicability of these findings and insights for managerial decisions.

It is essential to assess the usefulness of the outcomes in developing digital business plans and online marketing strategies for similar fields.

This chapter has explained the hierarchical research model and the mixed-method analytical approach that is employed in this research. Furthermore, it has described the data and the methods used in this project. The evaluation criteria has been presented as well. The next chapter describes, in detail, these analytical procedures and the results obtained from them.
4 Data Analysis

This chapter describes the qualitative and quantitative analyses that have been performed on the data. It is divided into three sections. First, in regard to the online travel market in the USA, a market overview is presented along with an overview of each of the study airlines. Online market share and competitive strategies analyses for those six airlines will also be presented. Second, there will be a discussion regarding online data analysis, which describes the analyses performed on ComScore data such as traffic source analysis, conversion rate, and revenue per online visitor and how they link to the customer journey and the airlines’ performance. Third, in regard to Social Network Analysis, an explanation and analysis of the constructed adjacency matrix will be presented in addition to network structure analysis and visualisation of this matrix.

4.1 The Online Travel Market in the USA

This section presents the initial analytical component in the study. It provides a market analysis of the online travel market with a focus on the top six airlines in the USA. Data from company annual reports, press releases, and business intelligence reports concerning the year 2014 were gathered and analysed.

First, some industry specific terminology will be explained. In the US airlines industry, the airlines are classified, based on their revenues, by the US federal government into three categories:
1. **Major airlines**: airlines that generate more than one billion USD annual revenues. This category includes, but is not limited to, American Airlines, Delta, Southwest, United, and US Airways.

2. **National airlines**: airlines that generate annual revenues between $100 million and $1 billion. They operate small- to medium-size aircraft and cover shorter destinations within the country or some international. Examples of this type include: Atlas Air, Airtran, Hawaiian, and Midwest Express.

3. **Regional airlines**: airlines that generate annual revenues of less than $100 million and serve specific regions of the country. [23]

Another common way of classifying airlines is based on their business model, which may be described as:

1. **Network carriers**: also referred to as legacy or full service airlines, are the most evolved type of airlines operating multiple types of aircraft and serving a global network with large fleets. American, Delta, and United fall under this category.

2. **Low cost carriers**: operate shorter flight distances with higher frequencies, offering lower fares with limited passenger services, such as Southwest and JetBlue airlines.

3. **Ultra-low cost carriers**: operate with the objective to offer the lowest ticket fare possible. Generally providing very basic service with additional charges for services such as on-board food or checked in baggage. Spirit and Frontier are two examples of this type [24].

4. **Other types** such as cargo or private hire models.
Throughout this analysis, the second classification will be used. As it is sensible to distinguish those airlines in a way that makes it easier to pinpoint and differentiate their competitive strategies.

**US Airlines Market Key Highlights 2014**

In the past decade, the US airline industry experienced considerable changes. From the consolidation of major players to the growth of low cost and regional players. Significant mergers took place between Delta and Northwest, American Airlines and US Airways, America West, and Trans World. United Airlines combined with Continental [25]. These changes have posed more challenges on the airlines to remain competitive with the competitors’ increased capacities and geographic network coverage. In addition, non-US carriers compete with the US airlines on international routes, as well, particularly Asian and Middle Eastern carriers.

Due to the seasonality characteristic of the industry, with peak travel in the summer from June to August; the industry must calibrate capacity to match supply with demand. Figure 4 illustrates the placement of the US airlines with regard to their annual capacity and revenue per seat mile. It can be seen that American Airlines, Delta, and United are placed in the top right corner of the chart, indicating longer routes, and they also make more revenue due to the higher pricing per fare. On the other hand, Southwest is placed somewhat in the middle with shorter routes and only a few international destinations, while regional and other low-cost carriers fly shorter distances and generate less revenue per mile [25].
Overall, in 2014 North American airlines scored slower growth of around 3% in international RPK (revenue passenger kilometres). The strong performance of the US economy among developed economies and trade gains, stimulated business-related travel [26].

According to RITA, in 2014 and, for the fifth year in a row, Delta airlines carried more total (domestic and international) passengers than any other US airline. On the other hand, United Airlines carried more international passengers than any other airline for the third year [27].

As for low cost carriers, according to Euromonitor, in 2014 Southwest dominated the low cost carrier category with a 57% value share followed with a large gap by JetBlue with an 18% value share. On the other hand, the network category was dominated by
the three giant legacy airlines, American, Delta, and United with a combined value share of 75% for the same year [28].

4.1.1 US Airlines Overview

This sub-section presents, in detail, the airlines business models and an overview of the six studied US airlines. Table 1 compares the six airlines’ ranked enplaned passengers, both domestic and international. It can be seen that, in 2014, the top five airlines remained at the same position from the previous year, while JetBlue escalated from 7th to 6th place.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Delta</td>
<td>129.210</td>
<td>1</td>
<td>120.389</td>
<td>7.3</td>
</tr>
<tr>
<td>2</td>
<td>Southwest</td>
<td>127.195</td>
<td>2</td>
<td>115.323</td>
<td>10.3</td>
</tr>
<tr>
<td>3</td>
<td>United</td>
<td>90.373</td>
<td>3</td>
<td>90.113</td>
<td>0.3</td>
</tr>
<tr>
<td>4</td>
<td>American</td>
<td>87.830</td>
<td>4</td>
<td>86.823</td>
<td>1.2</td>
</tr>
<tr>
<td>5</td>
<td>US Airways</td>
<td>57.514</td>
<td>5</td>
<td>57.007</td>
<td>0.9</td>
</tr>
<tr>
<td>6</td>
<td>JetBlue</td>
<td>32.056</td>
<td>7</td>
<td>30.428</td>
<td>5.4</td>
</tr>
</tbody>
</table>

Table 1 Airlines ranked by Scheduled Domestic and International Enplanements [27]
Airline Business Models

After the US airline deregulation in 1978, low cost carriers (LCC) emerged in the airline industry, making air travel more affordable. A new business model had been introduced by the pioneer in LCC, Southwest, and adopted by JetBlue and other LCCs. Unlike the traditional hub and spoke model adopted by legacy carriers such as Delta, American Airlines, and United, the point-to-point model focuses on short distance, regional, nonstop routes. Even though the hub-and-spoke model offers a wider network of coverage via a large and diversified fleet, it also has higher operational complexity and cost.

With the highly competitive nature of the airline industry, and customers expecting broader network coverage at lower fares, several airline carriers have adopted a mixed business model that can be referred to as the hybrid model. As the name implies, it combines the wider options of routes and services of the network model, with the cost efficiency of the low cost model. Nowadays, many legacy airlines own a brand that operates point-to-point such as United Express® under United Airlines. Likewise, low cost carriers offer a mix of long and short routes, and sometimes even international routes. By adopting the hybrid model, airlines aim to increase the number of passengers and ultimately improve their profitability [29].

American Airlines

American Airlines is an international airline based in Dallas, Texas in the United States. In 2014, it served 184 routes in Asia, Europe, North America, and South America with a large fleet of 627 aircraft. It formed a regional partnership with

---

3 In 1978 the government of the United Stated regulated the prices of commercial airlines which led the competitive carriers to shift from differentiation strategies to cost leadership strategies [59].
American Eagle\textsuperscript{6}; both offering an average of approximately 6700 flights daily to 350 destinations in 50 countries [30]. It is also a founding member of the OneWorld alliance, of which the current member airlines and members elect serve more than a thousand destinations in more than 150 countries [31].

Moreover, AA has a Trans-Pacific Joint Business Agreement with Japan Air and a Trans-Atlantic Joint Business Agreement with British Airways and Iberia through which it can coordinate their fares, services, and expand network routes.

In December 9, 2013, AMR Corporation (the company owning American Airlines) and US Airways Group officially formed the American Airlines Group and announced the merger of the two airlines. However, it wasn’t until October 17, 2015 when they began operating as one airline with a single passenger service system [32]. As of December 31, 2014, American’s and US Airways’ combined mainline fleet had a total of 983 operating aircraft with an average age of 12 years [30]. With this merger, the American Airlines Group led the industry with a 22% value share in 2014 [28].

**Delta Airlines**

Delta Airlines is a US-based legacy airline with its headquarters in Atlanta. It serves nearly 180 million customers annually. Delta operates a mainline fleet of over 800 aircraft covering 335 destinations in 61 countries on six continents. It is a founding member of the SkyTeam global airline alliance and has alliances with other foreign airlines, including Aero México and GOL. Moreover, it participates in international joint ventures, mainly the leading transatlantic joint venture with Air France,

\textsuperscript{6} American Eagle is a network of 10 regional carriers that operate under a codeshare and service agreement with American. Together they operate 3,400 daily flights to 240 destinations in the U.S., Canada, the Caribbean and Mexico.
KLM and Alitalia as well as a joint venture with Virgin Atlantic, in addition to agreements with many domestic regional carriers that operate as Delta Connection® [33].

**United Airlines**

United Airlines is based in Chicago, IL in the USA. In 2014, it operated an average of 5,055 flights a day to 373 airports across six continents. That year, it operated almost two million flights carrying 138 million passengers. United claims that it has the world’s most comprehensive route network including world-class international gateways to Asia and Australia, Europe, Latin America, Africa, and the Middle East. It operates a fleet of nearly 700 mainline aircraft. The airline is also a founding member of Star Alliance⁷ [34]. The current Star Alliance membership contains 28 Airlines offering 18,500 daily flights to 1,330 airports in 192 countries [35]. United is considered a traditional legacy carrier with a strong focus on service.

**US Airways**

US Airways was a major full service airline with its headquarters in Tempe, Arizona until it merged with American Airlines and ceased operations in October 2015 as mentioned previously. Before that, it operated a broad domestic and international network of 193 destinations in 24 countries across North America, South America, Europe, and the Middle East. US Airways operated around 3031 flights a day with a fleet of 343 mainline aircraft and 278 regional aircraft operated by subsidiary airlines under code sharing agreements.

---

⁷ “Star Alliance is a global airline network which was established in 1997 by five airlines, Air Canada, Lufthansa, Scandinavian Airlines, THAI and United.” [35]
Southwest Airlines

Southwest Airlines is a Dallas-based major passenger carrier that offers scheduled air transportation in the United States. It operates a network of 97 destinations across the US and seven other countries with a fleet of 680 aircraft. In 2014, Southwest acquired AirTran Airways and the integration was complete by the end of the year. In July 2014, Southwest became an international airline with added destinations to Nassau, Bahamas; Montego Bay, Jamaica; and Aruba. Since its beginnings in the 70’s, Southwest is considered a pioneer in facilitating affordability in air travel by lowering fares and increasing passenger traffic, and has become a leading low cost carrier [36].

JetBlue Airlines

JetBlue Airlines was founded in 2000 with headquarters in Long Island City, New York, and focus cities in Boston, Fort Lauderdale, Los Angeles, Long Beach, Orlando, and San Juan. It serves almost 35 million customers annually with an average of 925 daily flights to over 96 destinations in the US, Caribbean, and Latin America. It operates a fleet of 219 aircraft [37]. It is considered to be one of the major low cost carriers in the US with its unique value proposition of providing high class services to its customers at affordable prices. JetBlue offers point-to-point service, similar to other low cost carriers, but also includes differentiated products and perks that are generally provided by network carriers. Therefore, it can be said that JetBlue operates according to a hybrid business model. As such, JetBlue can be considered to be targeting the niche market of so-called underserved customers, who are seeking better services, more than other LCCs can offer, at a more reasonable price compared to network carriers [38].
4.1.2 Online Travel Market Overview

Online travel sales to residents in the US scored US$255.6 billion in 2014 with a rise of 6%. With the high penetration of internet in the US, consumers are spending an increasing amount of time comparing products and prices on multiple platforms. US travellers use the internet for all phases of the travel cycle from gaining inspiration, selecting a product and making a booking, to finalising the transaction and completing the purchase [39].

According to the Euromonitor ‘Airlines in the US’ 2014 report, almost 74% of total airline sales were through online platforms. Out of which, 55%, which equals approximately 67 billion USD, occurred through air intermediaries. Moreover, the online sales from air intermediaries are projected to steadily increase until 2019 (see Table 2) [40]. It can be observed that intermediaries account for more than half of online airline sales, which highlights the importance of intermediaries and their contribution to airline revenues in addition to the important role they play in the customer journey.

![Table 2](image1)

*Table 2 Forecast of the airlines’ online sales [40]*
The Importance of OTAs

In 2014, intermediaries’ sales reached almost US$188 billion with a growth rate of 2%. Moreover, as a result of the increasing number of consolidations in the travel industry, intermediaries gained more negotiating power. They provide a broader customer reach, in addition to other services such as marketing support and PR and social media management. Expedia and Priceline, for instance, established and managed Facebook pages. Such pages are used for information updates, customers’ discussions and complaints, and even allow companies to participate and communicate with their customers interactively [41].

Expedia

Expedia is an online full-service travel brand, amongst many leading brands, which was founded in 1996 and is available in 33 countries around the world. The brand’s purpose is to help travellers plan and book easily by providing them with a wide selection of vacation packages, flights, hotels, rental cars, cruises and in-destination activities, attractions, and services [42]. In 2013, Expedia got involved in a mutually beneficial marketing agreement with Travelocity by which Travelocity gained technology platforms which they used to promote their brand and market their travelling services which were provided by Expedia and in return, Expedia gained access to Travelocity’s supply and customer services. Travelocity-branded websites were used to launch Expedia’s hotel and air products for the United States towards the end of 2013 and for the Canadian website in 2014. In January 23, 2015, Expedia obtained the Travelocity brand, amongst other assets from Sabre for US$280 million [43].
Expedia’s marketing channels includes different types of online advertising providing a direct e-mail communication with travellers. Based on Expedia’s supplier relationships, they could also include promotional offers such as coupons and seasonal or periodic special offers. Expedia is also making use of affiliate marketing where bookings are received through different affiliate partner websites by providing affiliates with technology and access to a wide range of products and services such as travel products. The brand competes with many competitors from different regions due to more favourable offerings for travellers such as pricing and supply breadth.

In September 2015, Expedia gained all the assets and brands of Orbitz Worldwide Inc. this was beneficial for Expedia because Orbitz Worldwide Inc. brought passionate teams that built well-recognized brands. The acquisition enhanced Expedia’s customer service and assisted into enhancing the marketing and the distribution abilities. In addition, due to Expedia’s acquisition of Orbitz, most air ticket volumes increased 35% in 2015 in addition to the expansion of the brand Expedia. Furthermore, in 2014, the marketing agreement between Expedia and Travelocity improved the air volumes 30%, along with continuous improvements for the Brand Expedia sites themselves [44].

**Priceline**

Priceline is a major online travel agent, which specialises in providing access to online travel products, travel related reservation, and search services through its multi-brand websites. The brand’s aim is to lead in the online travel related services by providing the consumers with the best services that meet their needs and expectations. It intends to accomplish this by growing the business through continuous investments and innovations.
During 2014, the total online advertising cost was US$2.4 billion which was spent on internet search engines, meta-search and travel research services, and affiliate marketing [45].

### 4.1.3 Online Market Shares and Competitive Strategies

This section presents the online market share analysis as defined in the methodology for the studied airlines and identifies their relevant competitive strategies with regards to digital marketing.

**Market Share Analysis**

Adapted from RITA report [22], Figure 5 presents system-wide market share for the six airlines as defined in Section 3.2.1 in the methodology. System-wide refers to the combined number of passengers of both domestic and international flights. It can be seen that, for the first quarter of 2014, Delta and Southwest share the leading positions with 24% and 23%, respectively. JetBlue holds a moderate 6% market share. These results are compared with RITA’s overall ranking of airlines based on scheduled domestic and international enplanements (Table 1), and have been found to be generally consistent. The only difference appears in the ranking of American Airlines and United, which is a marginal 1% that could be due to the seasonality of the quarter compared to the whole year.
Further down, Figure 6 illustrates the breakdown of market shares among the six studied airlines for the first quarter of 2014 [22]. In the domestic travel market, Southwest Airlines is in the lead with 28% of enplaned passengers, followed by Delta Airlines with 23%. On the other hand, the international market is dominated by United Airlines, American Airlines, and Delta Airlines with a total of 85% of market shares. It is obvious that US Airways is the weakest player in both markets and that could be linked to the merger with American Airlines in October 2015, the time which it ceased its operations under the brand ‘US Airways’ [20]. Southwest and JetBlue operate mainly in domestic markets, while American Airlines, Delta and United generate 30% to 40% of revenue from international markets [25]. The distinction between the domestic and international market is important for the analysis. Findings about the customer search trajectory from the SNA could be related to characteristics of this distinction.
Competitive Strategies Analysis

Low cost carriers (LCC) have a low cost structure compared to legacy carriers. This imposes the adoption of cost-cutting strategies in different cost categories such as maintenance, fuel, inflight services, and marketing. For the purpose of this analysis, the focus will be on competitive strategies related to marketing. LCCs have to reduce their marketing efforts to sell tickets in order to be able to offer lower prices. One essential approach of doing so is selling the tickets directly through company-owned websites and avoiding the commission costs on sales through intermediaries and the additional labour cost of sales staff. Legacy carriers have borrowed particular strategies from LCCs to improve their efficiency and remain competitive in the market as well. Those strategies include reducing on-board services and substituting older aircraft with newer ones. Nonetheless, few cost efficiencies remain unique to LCCs that cannot be adopted by legacy carriers, resultant from a point-to-point business model, shorter routes, and single type aircraft [29]. Table 3 shows a list of elements
differentiating the network or legacy carriers from LCC. Airlines adopting the hybrid model apparently choose overlapping elements.

<table>
<thead>
<tr>
<th><strong>Full-Service (Network) Carriers</strong></th>
<th><strong>Low-Cost Carriers</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Core business: Passenger, Cargo, Maintenance</td>
<td>Core business: passenger air-service, ancillary offers</td>
</tr>
<tr>
<td>Hub-and-spoke network</td>
<td>Point-to-point network</td>
</tr>
<tr>
<td>Global player</td>
<td>Secondary airports</td>
</tr>
<tr>
<td>Alliances development</td>
<td>Single aircraft fleet</td>
</tr>
<tr>
<td>Vertical product differentiation</td>
<td>Aircraft utilisation</td>
</tr>
<tr>
<td>Customer relationship management</td>
<td>Minimised sales/reservation costs</td>
</tr>
<tr>
<td>Yield management and pricing</td>
<td>Ancillary services</td>
</tr>
<tr>
<td>Multi-channel sales</td>
<td></td>
</tr>
<tr>
<td>Distribution system</td>
<td></td>
</tr>
</tbody>
</table>

*Table 3 Airlines business models’ elements*

*Adapted from “The Airline Industry: Challenges in the 21st Century” [46]*

The air travel market can be distinguished under two types: business travel and leisure travel markets. Business travel market is more concentrated with fewer number of traveller who travel frequently and are generally not price conscious. Whereas the leisure travel market may consist of wider range of traveller who fly less frequently and are usually price sensitive [47]. Each market’s target customers have different requirements that airlines have to fulfil and market for in order to maintain a competitive market share. Legacy carriers usually target both types of travellers sometimes with a more focus on one, while LCC generally target leisure traveller.
**Strategies at Southwest The Unique Case**

Southwest is very keen on its online strategy as it states that it is an essential differentiation point from its competitors. The fact is that the company website is the only channel through which Southwest customers can purchase and manage travel online. In addition, the website is designed with additional features to help personalise the customer’s experience. Such features include providing deals based on a customer’s recognised location, and shopping cart functionality offering other travel products such as hotels and car rentals. Southwest also initiated its mobile app and website and mobile boarding passes, offering its customers the ability to access and manage services through their mobile devices. It was reported that, for the year 2014, almost 80% of the company’s passenger revenues came through its website [36]. Southwest’s online strategy is unique as it is the only airline within this study that doesn’t include its offerings on travel intermediaries’ websites.

**Successful Stories**

American Airlines took their ecommerce hub into the next level by implementing the “Abandoned Cart Program”. The program targets online shoppers who add flights into their shopping carts but do not proceed to checkout points and complete their purchases within 24 hours. Consequently, the program is triggered to send customized emails to the travellers by offering them options that are more flexible and cheaper yet related to their abandoned carts. Compared with standard fare sale campaigns, The “Abandoned Cart Program” resulted in a boost in open rates, click-through rates, and conversion rates by 300%, 200% and 400% respectively. In addition to that, American made sure to support every email sent to the site visitors with appealing visuals of their dream destinations, which played an important role in the conversion formula.
Southwest Airlines adopts an approach that is highly individual based, which profiles customers and designs custom-made offerings through sending weekly emails with a maximum of 15 different placements. This approach depends on several elements such as departing dates, current location and past behaviours. Impressively, 66% of those individually targeted visitors end up purchasing one of those placements according to Kevin Krone, the airline’s chief marketing officer [48].

4.2 ComScore Data Analysis

This section presents the second analytical component in the research design. It is concerned with studying ComScore panel data. The main objective of this analysis is to understand the customer journey and relate it to the airlines’ performance.

4.2.1 Descriptive Analysis

This section illustrates descriptive analysis of the ComScore data. The data has been obtained from ComScore’s Media Metrix Source/Loss report for February 2014. There are a total of six reports on each airline. The following data preparation process was performed:

- Data management and analysis were performed using Microsoft Excel. Source and loss data from six separate reports were merged and compiled into one report.
- Online media sources and destinations were assigned values of type as categorised under Airlines, OTA, Travel, or Non-Travel as explained in Section 3.3
- Descriptive analysis was performed and initial results were generated.
An overview of the distribution of incoming online traffic to the airlines’ websites is displayed in Figure 7. The multiple online media sources are categorised into four types, as defined in Section 3.3: airlines website, OTAs, travel, and non-travel. A further breakdown of both types of airlines and non-travel resulted in logon and search engine, respectively. This is to demonstrate, in particular, the customer search behaviour. What stands out in the chart is that, for all six airlines, most of the visitors come from search engine sites. As for the visitors logging in to the airlines websites, the percent ranges from only 9% for JetBlue to almost 13% for Delta. Likewise, OTA source type and Airlines have similar ranges of contribution to each airline’s website, apart from US Airways which has a high of 20% unique visitors arriving from other airlines’ websites. It is notable that, overall, the percentage of incoming visitors from travel websites is considerably low compared to other source types. This may be viewed as an opportunity for airlines to increase their presence on travel websites to gain a competitive advantage. It is interesting that one quarter of JetBlue’s visitors come from non-travel websites. Perhaps this is related to an online ad campaign. It is also interesting to note that the second major source of visitors for all six airlines, except US Airways, is non-travel websites, while other airline websites are the second major contributor to US Airways’ visitors. This may indicate that those airlines are trying to trigger the interest of the non-travel websites visitors by placing banner advertisements with, for example, exciting offers, therefore spreading their reachability to new customers.

The logon source means visitors arrived at the airline website directly by logging in. Airline in the chart represents visitors that arrive from other airlines’ websites. The logon and search engine data are further analysed to better understand visitor behaviour.
4.2.2 Customer Journey Analysis

Online visitors are regarded as customers in this online market place. Taking a look at the studied market share, Figure 8 shows the general distribution of online unique visitors among the six airlines for February 2014 as adapted from ComScore report. Southwest has around 30\% of the online unique visitors, followed by United at 20\%, then Delta which holds 18\%. JetBlue and US Airways come in last with less than 10\% shares of the online unique visitors. US Airways’ small share of UV may be linked to the announcement of the merger with AA that took place in December 2013. It is worth mentioning that Southwest and JetBlue are the only LCCs among the studied airlines, while the remaining are considered legacy carriers. The shares of unique visitors could
indicate the online visibility and presence of the airlines’ websites. However, this doesn’t necessarily always mean that a website with a larger number of unique visitors has a stronger online presence. A number of other factors may contribute and must be considered, such as market size and number of products offered online.

After looking at the airlines’ charts (Figure 5 and 8) and comparing between system-wide market shares and the shares of unique visitors, it seemed necessary to investigate the relationship between the two variables. Hence, a linear regression model was used to measure the strength of the said correlation. Figure 9 shows a definite positive correlation between the number of unique visitors to the airlines’ websites and the average number of passengers. With a determination coefficient of 0.682, it can be said that 68% of the variability in the number of passengers can be explained by the
variability in the number of unique visitors. In other words, the greater the number of unique visitors to an airline’s website, the greater the contribution is to the number of seats sold. However, this is not a causal relationship, as having more unique visitors doesn’t necessarily result in having more passengers, as other factors play significant role in the conversion. As a case in point, it can be seen from the chart that Delta and Southwest have a very similar number of passengers despite the variance in their number of unique visitors. This could be explained by the different strategies for each, as Southwest is a LCC that sells its tickets via its website only, so travellers are probably booking and buying themselves, whereas Delta is a legacy airline with presence on OTAs and a strong focus on business travel, so travellers could be booking their flights via other websites or corporate travel services. On the other hand, even though American and United are both legacy airlines with similar shares of passengers (Figure 5) they exhibit quite a variance in the number of unique visitors. That could be due to United’s website usability flaws that make it difficult for visitors to efficiently search and book flights [49].

It is worth noting that, even though unique visitors represent the customers reaching the airlines’ shops through a digital channel; the combination of traditional offline media and other channels must not be neglected as they contribute to the customer journey’s different stages, such as product awareness and connection. This strong correlation must be compared with insights from the social network structure analysis.
Next, ComScore source/loss data is analysed to enhance the understanding of the online customer journey. If it is assumed that the normal path for online customers starts from their login on the airline website and ends by logging off, Figure 10 illustrates a quantitative analysis of this path for all six airlines as a percentage of the total unique visitors for each. It can be seen that Delta Airlines has the highest percentage of logons among all its incoming traffic. This may be viewed as a positive result of having strong loyalty programs, for instance, or a negative outcome of not having enough presence on other websites that will contribute to incoming traffic.

As for the logoff rate, visitors can be regarded as having proceeded from the landing home screen to perform some kind of activity on the website, ranging from account management to the purchasing of a flight ticket. Hence, a high logoff rate may imply that the airline has been successful in retaining those online visitors on their platforms and pushing them forward in the online customer journey. Delta has investments in redesigning its website, with the objective of creating a great online experience for
their customers [49], and it appears that these investments have paid off as they enjoy the highest percentage of log-off rates.

![Online Traffic via Logon /Logoff](image)

*Figure 10 Percent of online traffic via Logon/Logoff for the month of February 2014 from UV to Airline’s website*

To further explore the customer search behaviour, Figure 11 compares the incoming traffic from the top four search engines: AOL, Google, Microsoft, and Yahoo. As expected, Google is responsible for the highest incoming traffic to all airlines’ websites. It is also observed that Microsoft and Yahoo’s engines have similar contributions to the incoming traffic for those airlines while AOL is always responsible for less than 5% for each airline.
However, to better understand the impact of the search engines on the competitors, a normalised chart has been created, taking into account the airlines’ market share that was presented earlier. The number of incoming traffic from each source has been adjusted by multiplying it by the ratio of total UV to total passengers for the same period for each airline accordingly.

As such, it can be seen in Figure 12 that United is receiving significantly more UV than American through Google even though they have nearly equal market share. This may indicate that United’s strategy with Google has been successful and helped it gain a competitive advantage over its rivals. A simple search for both airlines on Google showed that United results appeared in three different prime locations on the first page while American on two only. Such results can help the airlines target a better digital marketing strategy and invest in the appropriate media to maximise profit.

*Figure 11 Incoming traffic from Search Engines for the month of February, 2014*
4.2.3 Airlines Performance Analysis

This section analyses the performance of the six airlines from an online prospective, that is, with regards to their online visitors’ data. This data is then related to different measures: conversion rate and revenue.

Conversion rate

The ultimate goal for every airline is to increase its profits. One way of reaching this goal could be by maximising the number of passengers and minimising costs. ComScore data regarding the number of unique visitors to the airlines’ websites has been analysed in combination with passenger data from IATA [22] to compare the conversion rate for each airline (the rate of converting visitors to passengers). It is assumed that all other factors are similar for the six airlines throughout the year; thus, an average number of passengers for the year 2014 has been calculated and the results have been used to calculate the conversion rate. Then, taking into account that 74% of
total airlines sales are through online channels, as reported by Euromonitor [40], the number of passengers has been calculated accordingly. Furthermore, according to the same source, 45% of the airlines’ online sales are derived from the airlines’ own online channels which they refer to as air direct. Hence, this ratio has been applied to the number of passengers, except for SW, as it is known that it doesn’t utilise intermediaries in its sales. Finally, the resultant number of passengers is then divided by the number of total UV for each airline and the conversion rate in Figure 13 is computed.

It is obvious that Southwest is doing extremely well, with a high conversion rate of approximately 70% which could be linked, again, to its strategy of restricting online sales to its own websites. American, Delta, and US Airways show high conversion rates ranging from around 45 to 55%. This means that almost half of their visitors are moving from the search and evaluation stage in the customer journey to the purchase stage, and are most likely purchasing flight tickets. On the other hand, JetBlue has the lowest rates with less than 30%. Upcoming SNA analysis will examine the interaction among those competitors and its impact, if any, on the conversion rate.
Revenue per Online Visitor

The author is aware that *Revenue Passenger Mile* (RPM)\(^8\) is the unit used to measure airlines passenger revenues in the US. However, to obtain a ratio of revenue per online visitor, a different approach was used. Airlines’ annual revenue figures were gathered from the Bureau of Transportation Statistics report [50]. Applying the same factor of 74\% for online channels [40] (Euromonitor), annual online revenues have been calculated. Then, with the assumption that each of the twelve months equally contributes to the annual revenue, the figure was divided by ‘12’ to compute a monthly online revenue for each airline. It must be highlighted that revenue streams for airlines do not solely rely on passengers, as they may include cargo, advertisements, and so on. Therefore, to achieve a more accurate analysis, revenues from fares only were

---

\(^8\) “One revenue passenger transported one mile in revenue service. Revenue passenger miles are computed by summation of the products of the revenue aircraft miles on each inter-airport segment multiplied by the number of revenue passengers carried on that segment.” [58]
considered. Table 4 shows the results of the online revenues from fares by unique visitor.

It has been observed that the results are plausible, showing both Southwest and JetBlue with lower fares. If the numbers can be regarded as the airfare price the unique visitor has paid, totals of less than 100$ for LCCs and a range between approximately 200$ and 250$ for legacy carriers, appear to be reasonable. It is important to include this analysis in the study as it provides what can be regarded as the incremental value of online visitors for each airline. This will complement the findings of the correlation analysis.

<table>
<thead>
<tr>
<th>AIRLINE</th>
<th>ONLINE REVENUE PER UV (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMERICAN</td>
<td>210</td>
</tr>
<tr>
<td>DELTA</td>
<td>254</td>
</tr>
<tr>
<td>JETBLUE</td>
<td>95</td>
</tr>
<tr>
<td>SOUTHWEST</td>
<td>89</td>
</tr>
<tr>
<td>UNITED</td>
<td>200</td>
</tr>
<tr>
<td>US AIRWAYS</td>
<td>188</td>
</tr>
</tbody>
</table>

*Table 4 Revenue from fares (online) per UV February 2014*

*Adapted from: ‘2014 Airline Financial Data’ Bureau of Transportation Statistics [50]*
4.3 Social Network Analysis

In this section, the third analytical component is presented, which is concerned with Social network analysis. SNA concepts and techniques have been applied to the ComScore online panel data to perform network structure analysis and visualisation. First, the formation of the adjacency matrix will be discussed, followed by the network structure analysis. Finally, the visualisation representation will be presented and analysed.

The matrix representation of the data was selected to perform on the social network analysis. As a pilot, a 10X10 adjacency matrix (Figure 14) was created from the combination of source/loss data. The matrix involved the six airlines and the top four OTAs which, according to a Forbes report, are Expedia, Priceline, Orbitz Worldwide, and Travelocity who control around 95% of the US OTA market [51]. The empty cells in the matrix are missing data or there is no tie. For the latter case, the cells are imputed with a ‘0’. The diagonal values that indicate the tie of the actor to itself (reflexive ties) are all replaced with ‘0’ as well. This was done primarily because the selected SNA routines ignore the diagonal values. Hence, the actors’ ties to themselves must be analysed separately. Therefore, the logon/logoff traffic was analysed previously in Section 4.2.2, customer journey analysis. As for the missing values, further investigation was done to determine the most appropriate approach for imputation and this is explained in the following sub-section, adjacency matrix. After that, key relevant social network measures were analysed and a visualisation of the network will be discussed.
The pilot matrix was evaluated and the best ways for handling the missing values were investigated. It was noticed that most of the missing values represent the ties between actors of the type OTA. Given the purpose of this study, it appears that it is significant to have this data in order to build a more representative social network and gauge the interactions among key actors more precisely. Additional data was acquired. Ideally, source/loss reports for the four OTAs for the month of February 2014 must be obtained to maintain consistency with the airline data. However, the available data was only for one OTA (Expedia) and for the month of February 2015. Therefore, further data pre-

### Figure 14 Source/Loss adjacency matrix (pilot)

#### 4.3.1 Adjacency Matrix

![Adjacency Matrix Table]

The pilot matrix was evaluated and the best ways for handling the missing values were investigated. It was noticed that most of the missing values represent the ties between actors of the type OTA. Given the purpose of this study, it appears that it is significant to have this data in order to build a more representative social network and gauge the interactions among key actors more precisely. Additional data was acquired. Ideally, source/loss reports for the four OTAs for the month of February 2014 must be obtained to maintain consistency with the airline data. However, the available data was only for one OTA (Expedia) and for the month of February 2015. Therefore, further data pre-

---

**Figure 14 Source/Loss adjacency matrix (pilot)**

- **American Airlines**: 49 82 133 46 149 168 632 70 145 231
- **Delta Airlines**: 117 26 140 81 177 108 33 76 177 55
- **Expedia**: 236 185 87 201 227 77
- **JetBlue Airways**: 20 59 51 61 16 201 25 30
- **Orbitz Worldwide**: 94 163 48 177 190 52
- **Priceline**: 149 97 48 56 144 17
- **Southwest Airlines**: 72 329 323 126 174 162 75 120 189 81
- **Travelocity**: 31 118 24 108 96 27
- **United Airlines**: 106 136 167 17 169 153 23 103 109 120
- **US Airways**: 317 65 41 23 29 24 18 38 114 94
processing was performed with the objective of imputing the missing values in the matrix. Building on that, the goal was to assume the source/loss data for Expedia for February 2014, then estimate the other three OTA’s source/ loss data by modelling it on a pro-rata basis relative to their market share after researching the possible methods or tools that can be used to reach this goal. The following methodology was used using data gathered from published reports [44] [2]:

1. From the gross bookings of 2014 and 2015 for Expedia, calculate total market growth rate (gross booking).

2. Based on the previous correlation analysis between the number of unique visitors and passengers, it can be inferred that the number of UV is related to gross booking as well. Hence, apply that rate to calculate Expedia 2014 total UV from the available 2015 ComScore data.

3. Calculate the percentage of contributions for each source of incoming traffic to the total UV of Expedia 2015.

4. Assuming that the change in UV numbers is equal to the growth rate of gross bookings, calculate the total UV of 2014 by applying the growth rate in step 1 to total UV for Expedia 2015.

5. Calculate total market in terms of UV relevant to market share for 2014. Then, derive the total UV for other three OTA.

6. Assuming that all source channels had the same contribution the previous year, apply the calculated contribution percentage to the derived total UV to calculate Expedia, Orbitz, Priceline, and Travelocity 2014.

7. Finally, use those numbers to complete the adjacency matrix.
Figure 15 shows the final 10 x 10 adjacency matrix after handling the missing data. As the values represent the strength of ties between two actors, a maximum value of 1677 is observed from Priceline to Expedia and this can be described as the strongest tie in the network. On the other hand, the weakest tie is present from JetBlue to Priceline with a value of 16. This matrix is used to conduct the social network analysis and visualisation as presented in the following sections.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>82</td>
<td>133</td>
<td>46</td>
<td>149</td>
<td>168</td>
<td>632</td>
<td>70</td>
<td>145</td>
<td>231</td>
</tr>
<tr>
<td>2</td>
<td>117</td>
<td>0</td>
<td>140</td>
<td>81</td>
<td>177</td>
<td>108</td>
<td>33</td>
<td>76</td>
<td>177</td>
<td>55</td>
</tr>
<tr>
<td>3</td>
<td>236</td>
<td>185</td>
<td>0</td>
<td>87</td>
<td>682</td>
<td>1299</td>
<td>201</td>
<td>463</td>
<td>227</td>
<td>77</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>59</td>
<td>51</td>
<td>0</td>
<td>61</td>
<td>16</td>
<td>201</td>
<td>100</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>94</td>
<td>163</td>
<td>710</td>
<td>48</td>
<td>0</td>
<td>271</td>
<td>177</td>
<td>271</td>
<td>190</td>
<td>52</td>
</tr>
<tr>
<td>6</td>
<td>149</td>
<td>97</td>
<td>1677</td>
<td>48</td>
<td>838</td>
<td>0</td>
<td>56</td>
<td>639</td>
<td>144</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>72</td>
<td>329</td>
<td>323</td>
<td>126</td>
<td>174</td>
<td>162</td>
<td>0</td>
<td>120</td>
<td>189</td>
<td>81</td>
</tr>
<tr>
<td>8</td>
<td>31</td>
<td>118</td>
<td>536</td>
<td>24</td>
<td>268</td>
<td>204</td>
<td>108</td>
<td>0</td>
<td>96</td>
<td>27</td>
</tr>
<tr>
<td>9</td>
<td>106</td>
<td>136</td>
<td>167</td>
<td>17</td>
<td>169</td>
<td>153</td>
<td>23</td>
<td>103</td>
<td>0</td>
<td>120</td>
</tr>
<tr>
<td>10</td>
<td>317</td>
<td>65</td>
<td>41</td>
<td>23</td>
<td>29</td>
<td>24</td>
<td>18</td>
<td>38</td>
<td>114</td>
<td>0</td>
</tr>
</tbody>
</table>

*Figure 15 The final source/loss 10 x 10 adjacency matrix*

### 4.3.2 SNA Analysis

The 10 x 10 adjacency matrix (Figure 15) was formed from the combination of source/loss data, estimations, and calculations. The matrix is valued, asymmetric, and involves the six major US airlines and the top four OTAs. The software UCINET 6
was used to conduct the SNA and the software NetDraw was used as the visualisation tool. With regard to the social network analysis for this study, it mainly involves three exercises: network and individual measures, a difference matrix, and source/loss rate analysis. Each is explained in detail as follows.

**Network and Individual Measures**

Here, some important network measures are examined and analysed to understand the network structure.

**Centralisation and Density**

Whole network measures help to reveal some features of the network that characterise the entire network. Density is a network cohesion measure that can describe the extent to which a network is integrated within itself. In a valued network, density can be defined as the ratio of all present ties to the number of possible ties [52]. Density can be considered for the entire network or can be calculated for subgroups as well. Centralisation is another network measure that describes the extent to which a single actor controls the network [14]. Analysing those measures is important for understanding the network’s properties.

Linton Freeman has introduced basic measures of actors’ centrality based on their degree, and the overall Centralisation of graphs. According to the Freeman approach, the Centralisation measure indicates the degree of inequality or variance in the network from that of a perfect star-shaped network of the same number of ties, expressed in percentage [52]. In this case, since the network consists of directed relations, two Centralisation scores are calculated: out-degree and in-degree Centralisation. The Centralisation statistic is divided by the maximum value in the input dataset. However, both out-degree and in-degree network Centralisation here are almost equal with
around 15% graph Centralisation each. Having said that, this network can be described as relatively decentralised, meaning that the power of individual actors in the network does not vary greatly with a similar distribution of the overall positional advantages in this network. This finding is further examined to identify the actors contributing to the Centralisation score.

**Degree Centrality**

Degree centrality is an individual measure that refers to the level of involvement or activity of an actor in the network. The position of a node (actor) in a network is very important for its achievement. To enhance the understanding of this property, centrality measures have been analysed. Since the relationship in this network is directed with valued ties, both in-degree and out-degree centrality measures have been analysed. In-degree centrality refers to the number of ties received by an actor from other actors, while out-degree centrality denotes the number of ties sent or given by that actor to other actors in the network. They are used as a measure of prominence or expansiveness, respectively [53]. It is usually perceived that actors which attract many ties (high in-degree centrality) as prominent or important in the network. Also, actors sending out many ties (high out-degree centrality) can be characterised as influential in the network. For this dataset, highly central actors (both in-degree and out-degree) would be those who demonstrate high online traffic volume to and from other actors, meaning more interaction with other players in the network. In contrast, actors with low degree centrality are more peripheral to the network.
The chart in Figure 16 presents the in-degree and out-degree measures for the ten actors in the network using the Freeman approach in UCINET 6. Overall, it is obvious that the OTAs have greater in-degree and out-degree scores than the airlines, indicating that they have more traffic through their websites in general. This demonstrates the importance of OTA’s role in the network and can help in understanding the consumer behaviour in this network. This could be interpreted as, within the studied network, regardless of the size of the website, the majority of the online consumers prefer to use OTAs. However, the actual activity cannot be determined as it may be searching for travel information, comparing prices, booking a flight, or completing the booking transaction. This explains the importance of price comparison websites in the customer journey. For LCCs, it was noticed that SW has substantially greater in-degree and out-degree centrality than JetBlue. Whereas for legacy airlines, United has the most in-degree centrality while American Airlines has the highest out-degree centrality, indicating greater propensities for American for losing visitors.
Figure 17: Correlation analysis between In-degree Centrality and passengers

To examine the impact of the volume of incoming traffic within this network into the airlines’ websites on their performance, a linear regression model has been applied on the in-degree scores and the number of passengers for the same time period. Figure 17 illustrates a strong positive correlation between the in-degree centrality scores of the airlines’ websites and the average number of passengers for that quarter. These results are consistent with the previous correlation analysis between total unique visitors and the number of passengers that was presented in Section 4.2.2, which indicates that, with the set network boundaries, the actors’ roles are quite similar to that of the whole network before applying the boundaries. This may suggest that airlines that attract more visitors to their online platforms are more likely to have more passengers. It is worth mentioning that this is not a causal relationship, meaning that having more online visitors will not necessarily yield more passengers and more evidence would be required to analyse such a relationship.
Density by Group

It is now important to better understand the interaction between actors in the network. Therefore, a density by group analysis has been conducted. The two groups represent the actor type as categorised earlier (airline, and OTA). The resultant matrix is shown in Figure 18. Furthermore, Table 5 illustrates that the interaction between actors within the same group is greater than cross group interaction. With the ties expressing online visitors, it can be interpreted that the flow of unique visitors between websites from the same type is more than that of different types of websites. This could be a result of consumers switching between competitors’ websites to compare prices for instance. This flow was further analysed by looking at the visualisation of the network.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Amer</td>
<td>82</td>
<td>55</td>
<td>117</td>
<td>317</td>
<td>72</td>
<td>106</td>
<td>149</td>
<td>31</td>
</tr>
<tr>
<td>2</td>
<td>Delta</td>
<td>231</td>
<td>81</td>
<td>55</td>
<td>65</td>
<td>329</td>
<td>136</td>
<td>97</td>
<td>118</td>
</tr>
<tr>
<td>10</td>
<td>US Airways</td>
<td>46</td>
<td>33</td>
<td>23</td>
<td>20</td>
<td>81</td>
<td>120</td>
<td>48</td>
<td>27</td>
</tr>
<tr>
<td>4</td>
<td>JetBlue Airways</td>
<td>145</td>
<td>177</td>
<td>114</td>
<td>281</td>
<td>126</td>
<td>17</td>
<td>56</td>
<td>24</td>
</tr>
<tr>
<td>7</td>
<td>Southwest Airlines</td>
<td>168</td>
<td>108</td>
<td>24</td>
<td>16</td>
<td>162</td>
<td>153</td>
<td>639</td>
<td>284</td>
</tr>
<tr>
<td>9</td>
<td>United Airlines</td>
<td>133</td>
<td>76</td>
<td>4</td>
<td>100</td>
<td>120</td>
<td>103</td>
<td>167</td>
<td>536</td>
</tr>
<tr>
<td>6</td>
<td>Priceline</td>
<td>149</td>
<td>31</td>
<td>236</td>
<td>149</td>
<td>72</td>
<td>106</td>
<td>149</td>
<td>72</td>
</tr>
<tr>
<td>8</td>
<td>Travelocity</td>
<td>118</td>
<td>27</td>
<td>185</td>
<td>27</td>
<td>329</td>
<td>136</td>
<td>97</td>
<td>118</td>
</tr>
<tr>
<td>3</td>
<td>Expedia</td>
<td>227</td>
<td>24</td>
<td>87</td>
<td>87</td>
<td>126</td>
<td>120</td>
<td>48</td>
<td>27</td>
</tr>
<tr>
<td>5</td>
<td>Orbitz Worldwide</td>
<td>190</td>
<td>96</td>
<td>281</td>
<td>281</td>
<td>188</td>
<td>177</td>
<td>190</td>
<td>24</td>
</tr>
</tbody>
</table>

Figure 18 Online traffic ‘Density by group’ matrix

<table>
<thead>
<tr>
<th></th>
<th>Airlines</th>
<th>OTAs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Airlines</strong></td>
<td>3670</td>
<td>2752</td>
</tr>
<tr>
<td><strong>OTAs</strong></td>
<td>2652</td>
<td>7858</td>
</tr>
</tbody>
</table>

Table 5 Density by Group (Sum of tie strengths)
**Difference Matrix**

Taking a closer look on the dyadic level, in an attempt to understand the relationship between each pair in the network, a difference matrix has been formed. Figure 19 depicts this matrix that is derived from the adjacency matrix. In UCINET 6, the ‘**TRANSFORM > SYMMETRIZE**’ routine was performed on the original matrix and the ‘**difference**’ criterion was selected. The resultant matrix is a symmetric one with the values being the variance between the pair’s correspondent values (tie strengths). For ease of representation, only the positive values are shown in the matrix in the figure, indicating that the variance is in the favour of the actor in the column (the destination) and the cell is coloured in green. For example, in the tie between Delta and American, the difference is equal to 35 for American. This means that American Airlines receives more visitors from Delta websites than Delta does from American airlines by a value of 35. This analysis is essential for finding the beneficiary party in the relationship between each pair of actors in the network from a website source-loss perspective, which indicates the shared visitors among those competitors. Moreover, this matrix will be used to visualise those relationships.
Taking a step further to the individual level, the following analysis examines the net gain for each actor in the network as defined by the variance between the total number of incoming visitors and the number of outgoing visitors within the network boundaries. One could argue that, in reality, it is impossible to have fewer incoming visitors than outgoing. However, given the network boundaries set as per the study focus, it is recognised that there exist other sources of visitors outside those boundaries. Hence, the negative net gain is noticed in some actors of the network. Nevertheless, a negative figure doesn’t necessarily imply loss. To illustrate this, the proportion of the net gain to the overall flow (total incoming visitors to each website) is presented in Table 6. It is noticed that Priceline is losing a significant 28% of its

Source/Loss Rate

![Figure 19 Difference matrix](image_url)
total visitors to other websites in the network followed by American Airlines, which has a nearly 13% loss. As for JetBlue and Southwest, it can be seen that the relatively low marginal loss may have little impact on their performance with losses of around 3% and 2%, respectively. On the other hand, Travelocity and Orbitz are successfully attracting and keeping online visitors from the remaining nine websites with almost 11% and 10%, respectively. This could be related to the loss of Priceline, with consumers visiting its website first then switching to the other OTA’s to compare deals, for example.

<table>
<thead>
<tr>
<th>Total Source</th>
<th>Total Loss</th>
<th>Net Gain</th>
<th>% Net Gain to Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Airlines</td>
<td>1141</td>
<td>-516</td>
<td>12.51%</td>
</tr>
<tr>
<td>Delta Airlines</td>
<td>1233</td>
<td>269</td>
<td>5.76%</td>
</tr>
<tr>
<td>JetBlue Airways</td>
<td>500</td>
<td>-62</td>
<td>2.59%</td>
</tr>
<tr>
<td>Southwest Airlines</td>
<td>1449</td>
<td>-127</td>
<td>1.69%</td>
</tr>
<tr>
<td>United Airlines</td>
<td>1306</td>
<td>314</td>
<td>6.76%</td>
</tr>
<tr>
<td>US Airways</td>
<td>689</td>
<td>21</td>
<td>0.95%</td>
</tr>
<tr>
<td>Expedia</td>
<td>3777</td>
<td>320</td>
<td>2.74%</td>
</tr>
<tr>
<td>Orbitz Worldwide</td>
<td>2548</td>
<td>572</td>
<td>9.79%</td>
</tr>
<tr>
<td>Priceline</td>
<td>2405</td>
<td>-1259</td>
<td>28.27%</td>
</tr>
<tr>
<td>Travelocity</td>
<td>1879</td>
<td>468</td>
<td>10.52%</td>
</tr>
</tbody>
</table>

Table 6 Net Gain online traffic network
4.3.3 Visualisation

The objective of conducting the visualisation analysis is to enhance the understanding of the overall structure of the network and to help in the interpretation of the findings. This is considered to compliment the quantitative social network analysis performed in the previous section, to help understand the network in a qualitative manner that may, otherwise, be difficult to achieve.

The final 10 x 10 adjacency matrix (Figure 15) is used as input data in the software, NetDraw 2.158 [54]. Due to the properties of the ties being directed and valued; a series of iterations has been performed to find the most appropriate representation of the network that would facilitate answering the research questions.

Figure 20 shows the outcome of the initial exploratory iteration with a random layout\(^9\), meaning no specific layout algorithm has been selected. This can be described as displaying the data in its raw format without any manipulations or inclusions, except for the actor’s type and online market share that have been embedded in the network graph as a node attribute where the shape and colour correspond with the actor’s type, and the size corresponds to the online market share. This visual representation of the attributes enriches the understanding of, and aids in interpretation of, the network. The random layout is evaluated, and it is apparent that the network is generally connected, with all actors being connected to one another through some kind of path. It seems to be quite difficult to obtain much information from this visualisation; however, the overall decentralisation characteristic of the network that has been described earlier, is clearly visible in this graph.

\(^9\) “The layout of the network diagram refers to the position of the points in the diagram.” [14]
Following the initial iteration, several iterations and evaluations have been performed. Figure 21 displays the results of running the algorithm *Scaling/ordination layout* with the criterion *similarities (strength of ties)* without specifying any threshold value. The actor’s node here is placed with relevance to the strength of its ties, meaning the position in the network reflects the actor’s power. As can be seen, Expedia and Priceline are placed in the middle of the diagraph which means they play a central role in the network, while US Airways and JetBlue are relatively peripheral. These results mirror those of the previous analysis that examined the degree centrality measures.

Although this layout has helped in identifying some network characteristics, it could still be improved. Depending on the research at hand, a better visualisation may be displayed after including additional data into the graph. This data could be either actor attribute data or data derived from previous analyses performed on the network. Another way to improve the visualisation would involve selecting different layout algorithms as appropriate and manipulating parameters to observe changes in the
network visualisation. Introducing a threshold for the values of ties to be included in the visualisation could be a useful method as well.

Figure 21 Visualisation Scaling/ordination layout

To build on the difference matrix analysis that was presented previously (Section 4.3.2.2) and, in an attempt to reduce the complexity and simulate a clearer visualisation of the relationships between actors, the graph in Figure 21 was formed. This was created by using the difference matrix (see Figure 19) as input and selecting a ‘by group’ layout. The reason for choosing the difference matrix as input for the visualisation is to enhance the understanding of the beneficiary actors in the relationship. In addition, by embedding the actors’ type and online market share into
the nodes’ colours, shapes, and sizes as in the previous iterations, the strength of ties is embedded here as well. Lines of stronger ties, those consisting of a higher number of unique visitors, appear thicker with larger arrow heads. At the same time, weaker ties, representing a lower number of unique visitors, have thinner lines with smaller arrow heads.

Taken together, this graph shows a number of interesting issues:

- Thicker lines have a tendency to be within the groups and not between them (i.e. actors from the same type seem to share stronger ties with others in the group), which can be interpreted as the flow of online visitors between websites belonging to actors from the same type representing more than cross over flows.
- American Airlines has a strong directed tie with Southwest, which means it is losing a considerable number of its unique visitors to Southwest.
• Priceline is sending strong ties to its competitors, which indicates that a significant number or Priceline’s visitors are leaving its website to go to Expedia, Orbitz, and Travelocity.

• Interestingly, Expedia primarily sends ties to airlines’ websites, except for Southwest and the other three OTAs, as it receives ties from them. This means Expedia is losing online visitors to airlines’ websites and not to other OTAs, but is receiving visitors from Southwest. With this pattern, Expedia may gain a competitive advantage over the other OTAs by attracting their online visitors. This may be explained in terms of consumers visiting Expedia first and then moving on to a specific airline website to compare details (prices, offers, and so on); however, the case is different for Southwest customers as they have to visit Southwest’s website first (being the only online channel for booking) to check fares or offer details before moving on to other websites to search for other similar flights and compare prices.

With this final part of the analyses, the data analysis chapter is concluded. This chapter has described the three analytical components used in this study. The online travel market in the US has been reviewed, shedding light on the top airlines in terms of their market shares and competitive strategies. A ComScore data analysis followed, providing a comprehensive analysis of the online panel data with regard to the customer journey and airline performance. Finally, the social network analysis was conducted, examining the adjacency matrix, studying the network structure, and discussing the network’s visualisation. The next chapter will describe synthesis and evaluation of the analyses results and their implications.
5 Discussion and Evaluation

The previous chapter presented an extensive analysis of the datasets at hand. The ComScore data analysis complements the SNA, providing interesting insights and a clearer picture of the online market. This chapter discusses the results of these analyses and presents a synthesis of the insights and interpretations within the online marketing context. Important managerial implications from these findings are also highlighted.

The clearest finding to emerge from the analysis is that the online network of the studied market is well connected, with paths connecting all the actors to one another. These connections and the suggested flow of online customers’ cross-visiting websites confirm the fierce competition in the field noted in numerous studies and experts’ reports.

Another important result from the analysis is the significant contributions of search engines as sources of airlines’ incoming traffic. All the airlines received nearly half of their total incoming visitors from search engines, even exceeding visitors who came directly by logging onto the site. Therefore, airlines are encouraged to invest in search engine marketing, especially with Google, as it accounts for the greatest contribution (see Figure 7). Airlines should regularly evaluate their brand search results against universal searches which include images, videos, and maps in addition to websites. Airlines should implement search-engine marketing techniques through either paid placements (pay-per-click) or placement in organic listings (search engine optimisation) [3]. Moreover, airlines should use web analytics to direct search traffic to their web pages that display customised content to create personalised shopping experiences for visitors. In these ways, airlines can push visitors further along the customer journey from the search-and-evaluation phase to the purchase stage.
The present study also found that larger websites generally attract more visitors than smaller ones, but by implementing a focused, dynamic online marketing strategy, even smaller websites can reach more potential customers and improve their conversion rate. Based on the distribution of UV sources, airlines should review their strategies to take advantage of existing high-traffic sources and to evaluate and improve lower-traffic sources. As a case in point, the analysis identified that the percentage of incoming visitors from travel websites is low compared to other source types. This can be viewed as an opportunity for airlines to increase their presence on travel websites and trigger the interest of customers in the pre-purchase stage to gain a competitive advantage. For instance, travel giant Tripadvisor.com recently introduced a new feature allowing visitors to rate and review airlines as they do hotels and restaurants [55]. Airlines should evaluate such initiatives by travel websites and their impact on business.

To test the common view of the importance of online traffic to sales volume, the correlation between the number of passengers and the number of unique visitors to airlines’ websites was examined. The results show a statistically significant positive correlation between the two. This finding is regarded as compelling evidence that attracting visitors to digital platforms is extremely important for airlines’ profits. In particular, researchers have reported that online sales contribute 74% of airlines’ revenue [40].

Although the quantitative analysis shows that OTAs are responsible for the majority of online visitors within the Travel source type, the SNA analysis identifies Expedia as the main source among OTAs for incoming visitors to airline websites (with the exception of SW’s website) (see Figure 22 Visualisation difference matrix ‘By group’ layout). An explanation of this result is that these airlines benefit from an online
presence on Expedia’s platform. These visitors’ search behaviours can be described as those of directed information seekers, directed buyers and bargain hunters, as defined earlier. This behaviour indicates the iterative movement between the different stages of the online customer journey model (Figure 1). However, it is important to note that the numbers analysed here represent unique visitors who might have left the website before reaching the purchase stage, gone to another website and then returned to complete the transaction, in which case they are difficult to capture. It is then reasonable to assume that visitors leaving OTAs’ websites for airlines’ websites most likely are considering directly purchasing from airlines’ websites, but more information is needed to establish an accurate understanding of this matter. Nevertheless, this observation seems to align with Expedia’s report that airline ticket sales generated only 8% of its 2015 revenue [44].

In contrast, the SNA showed that Priceline plays a key role in the shared visitors’ network of OTAs. The other three competitor OTAs receive rather substantial number of visitors from Priceline. A possible explanation is that the Priceline’s high online visibility stimulates users to visit it as a starting point in their search journey.

Unsurprisingly, Southwest has the highest rate of converting online visitors to passengers, possibly as its online platforms are the only medium where online purchases take place. Interestingly, this strategy has helped Southwest gain positional advantage as a hub in the network with both high in-degree and out-degree centrality scores, indicating a significant number of shared visitors with other players in the market.

The SNA analysis revealed a different pattern for Southwest and JetBlue visitors, who exhibit stronger ties with OTAs. Table 7 shows the percentage of visitors leaving airlines’ websites for OTAs. It is clear that LCC customers behave differently than
those of network carriers and tend to switch more to OTAs’ websites. Price-sensitive travellers, who are generally the target of LCCs, tend to search and compare to find the lowest price before making purchase decisions. This result is consistent with Holland and Georghiades’s [56] finding that Southwest customers have quite different search behaviours than legacy airline customers and engage in a more active search process.

<table>
<thead>
<tr>
<th>Total source</th>
<th>Loss to OTA</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Airlines</td>
<td>1,141</td>
<td>113</td>
</tr>
<tr>
<td>Delta Airlines</td>
<td>1,233</td>
<td>25</td>
</tr>
<tr>
<td>JetBlue Airways</td>
<td>500</td>
<td>89</td>
</tr>
<tr>
<td>Southwest Airlines</td>
<td>1,449</td>
<td>240</td>
</tr>
<tr>
<td>United Airlines</td>
<td>1,306</td>
<td>16</td>
</tr>
<tr>
<td>US Airways</td>
<td>689</td>
<td>18</td>
</tr>
</tbody>
</table>

*Table 7 Airlines’ visitors interaction with OTAs as adapted from difference matrix*

Although the SNA analysis found that the studied network is decentralised, it also confirmed that OTAs generally have more online traffic than airlines. This again highlights the importance of travel intermediaries for the airline industry to reach more customers. In a relevant note, further mapping this flow of online visitors, the SNA’s density by group analysis and the visualisation both revealed that the interaction levels are greater within groups. It can be inferred that these customers are in the search-and-
evaluation stage of the customer journey and that, by switching between websites, they are trying to develop a consideration set. This pattern can characterise the behaviour of customers who use OTAs and tend to make comparisons with other OTAs’ websites offerings more than airlines’. Airlines should try to prevent their visitors from switching to other airlines’ websites not only by offering competitive prices but also by enhancing the online customer experience and improving interactions with the airlines’ online channels.

Finally, the SNA established that an actor’s role in the network has a strong correlation with the number of passengers. More central actors exhibit a stronger correlation, indicating that competitors with higher online visibility tend to have more passengers.
6 Conclusions

The aim of this project was to analyse ComScore online panel data and map consumers’ search path in the airline market’s social network. Using online visitors’ source/loss data and applying concepts from the theories of consumer behaviour, customer journey and SNA enabled studying the network structure of online competitors in the selected market and the impact on consumers’ search trajectory. The findings contribute managerial implications in the context of online marketing.

This study showed that the social network of online competitors is well connected and generally decentralised and that a high degree of shared visitors switch among websites. Airlines must develop or leverage multi-brands to reach different customer segments among those floating online visitors. Moreover, airlines must gain knowledge from big data derived from their customer database and use it in innovative ways to retain existing customers by creating a customised online experience, predicting customers’ needs and providing deals that better match those needs. These findings and conclusions answer the first research question about the structure of the social network.

With respect to the second research question, it was found that, within the studied network, OTAs generally have more traffic than airlines’ websites. This finding highlights the significant role that travel intermediaries can play for the airlines industry to reach more customers. Additionally, the significant positive results of the correlation analysis established that attracting more online visitors to airlines’ online channels is especially important for their profits. Regarding consumer behaviour, the
SNA analysis found that LCC customers have different search patterns than those of legacy carriers and engage in more cross-visiting with OTAs.

The final research question about the interaction between OTAs and competitor airlines has a two-fold answer. First, the analysis found that the interaction levels indicating customers’ crossover search behaviour are greater among competitors of the same type than competitors of different types. Second, the analysis showed that Expedia in particular had great impacts on the incoming volume of visitors for all competitor airlines, except for Southwest.

These findings have significant implications for the understanding of how key players in the competitive landscape of the online travel market interact in terms of the search behaviour of their consumers. Overall, this study strengthens support for the idea that OTAs are extremely important in the online traveller’s customer journey.

The data used in this study have several advantages from two perspectives. The data were collected from the clickstream data of an online panel of more than 2 million users and consequently are highly reliable compared to self-reported or survey data. Another advantage comes from the use of retrospective data, which is being able to identify and evaluate the strategies adapted for the present.

Examining a competitor relationship network by applying SNA techniques to online source/loss data provides a sophisticated analysis of the customer journey underlying this complex network structure and enables understanding the flow of online visitors among the studied competitors. Although this methodological approach can be applied to networks of online competitors in other industries, the results should be interpreted with caution as the ties in the network could represent different relationships.
6.1 Managerial Contributions

This research has implications for decision-makers and practitioners in the United States and similar markets worldwide (particularly well-developed online market settings). The innovative application of SNA, in combination with market analysis, to examine competitors’ relationships and the online customer journey provides a systemic approach to build digital marketing strategies. SNA illustrates the structure of the complex network, so decision makers can review and quantify relationships and behaviours and detect specific threats and opportunities. Thus, they can better understand the network and customers’ search behaviour and identify influential relationships and potential strategic partnerships and collaborations.

A key challenge facing decision makers in a mature market is keeping up with increased demand from their experienced customers. Airlines should continuously gather information about existing and potential customers and use information they already have about travellers’ demographics, requirements and even attitudes. By conducting Big Data analysis, airlines can reach more customers and predict any changes in customer needs.

As the SNA showed that the number of shared online visitors is greater within websites of the same type, airlines are encouraged to expand their marketing alliances with other airlines with the objective of reaching a wider customer base. Moreover, airlines should dynamically review and adapt existing marketing-alliance activities, such as frequent flyer programs and travel packages, to meet customer needs by means falling under their corporate strategy.
6.2 Limitations and Future Work

The findings reported here are subject to three limitations. First, the generalisability of the results is subject to certain constraints. For instance, the specific nature of relationships between competitors in a network may differ, so interpretations might vary and must be made carefully. Therefore, even though the methodology can be applied to online markets in other fields, the findings cannot be generalised. Second, the study scope was limited as it considered only key players in the air-travel online marketplace. Meta-search engines, for example, were not included in the scope but might play an essential role in the network; therefore, a study of the wider network is a task for future work. Finally, due to availability constraints, the data used to form the matrix of the social network represents a snapshot of the competitors’ online visitors at a specific point in time, in this case, one month. Ideally, data on online visitors from over a year should be considered to more accurately capture the trends and developments in this dynamic online marketplace.

A possible area for future research is to investigate the methodologies used in building an analytical model that combines the theories used in this study with customer segmentation theory. Merging understanding of competitors’ relationships in an online market with understanding of individual visitors’ attributes and demographics could help organisations develop specific marketing strategies. For example, mobile travel gross bookings in the United States recorded growth of 390% from 2012 to 2016, according to Phocuswright [57]. In this context, reaching this level of granular analysis of the device used by customers could have great impacts on airlines’ profitability.
References


