

Data mining applications in civil aviation sector: State-of-art review

Musab Talha Akpinar
Ankara Yildirim Beyazit University
Management Information Systems Department
Aybu, 06400
Ankara, Turkey
e-mail: musabtakpinar@gmail.com

Muhammed Emin Karabacak
Ankara Yildirim Beyazit University
Management Information Systems Department
Aybu, 06400
Ankara, Turkey
e-mail: karabacakmemin@gmail.com

Abstract— Data mining is the process of analyzing data from different perspectives and summarizing it into useful information take place in the scope of business intelligence. Nowadays increasing of the globalization and integration of the world, transportation sector is one of the most required area that needs to be used for data analysis. Purely, planes equipped with flight recording data typically record up to 500 variables of data - described in these flight data recordings are time, altitude, vertical acceleration, and heading- per second for the duration the plane is being operated. Airline companies may use data mining in order to fuel cost optimization, planning take into consideration weather conditions, passenger analysis, cargo optimization, airport situation revenue per flight, profit per flight, cost per seat or more detailed one catering and handling expenses per seat. In this article, the current applications of data mining in civil aviation sector are reviewed based on certain critical factors including airlines, airports, cargo, passenger, efficiency and safety. The critical analysis of 63 empirical studies reveals that the usage of data mining in airlines sector is still in its early stages and the ability of these studies to generate knowledge may not be sufficient. It's tried to subhead papers and classified them subject by subject and also yearly. The intention of this study is that this is going to guide that sheds light on this field studies and staff. And also we want to put forth the tendency of civil aviation sector and evaluation progress of airlines in business intelligence concept. Given these findings, it can be suggested that there is more to dig for in order to obtain more managerially interpretable and acceptable results in further studies. Also, recommendations are made for other potentials of other business intelligence tools or data mining techniques research.

Keywords—Airlines; Business Intelligence; Civil Aviation; Data Mining

I. INTRODUCTION

Data, or pieces of information, have been collected and used right through history. However, in contemporary world, advances in digital technology have considerably boosted our ability to collect, store, and analyze data. All of the data, however, are merely that data until they are analyzed and used to inform decision-making [5]. Data mining can be viewed as a result of the natural evolution of information technology.

Business intelligence represents tools and systems

that allow a company to gather, store, access and analyze corporate data to aid in decision-making and strategic planning process. BI is an architecture and a collection of integrated operational as well as decision-support Applications and databases that provide the business community easy access to business data [56]. One of the most significant business intelligence tools and best practices is data mining. Generally, data mining is the process of analyzing data from different perspectives and summarizing it into useful information take place in the scope of business intelligence. Data mining software is one of a number of analytical tools for analyzing data. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases [33].

As a result of these rising trends, there is now a widespread realization that an enormous volume of data can be captured, stored, and processed. After that the knowledge gleaned from such data has advantages for everyone: business sectors, governmental issues, academic disciplines, engineering working, communities', and individuals' studies [5].

Nowadays increasing of the globalization and integration of the world, one of the most notable part of the transportation sector is civil aviation sector, has grown at an average annual 5% since 1980 [35]. Within this increase, automatically and manually accumulating data in airlines companies' systems is impracticable due to the mass amount of data produced every day. Because of the huge amount of the data, airlines corporations are obligated to utilize of business intelligence tools especially diverse data mining technics and applications.

The objectives of this paper are to provide a reviewed to Data Mining in Civil Aviation and subhead papers. The search strings were used on 3 different digital libraries relevance to Data Mining: "ScienceDirect" (www.sciencedirect.com), "Google Scholar" (scholar.google.com) and Inspec (www.iee.org/Publish/INSPEC/) in January 2017. Articles

search terms were; Data Mining, Civil Aviation, Airlines, Business Intelligence, Classification, Clustering, Decision Tree, Aviation Data Analytics and Aviation Association. It was also studied in binary and trio combinations of these terms.

In this article, the current applications of data mining in civil aviation sector are reviewed based on certain critical factors including airlines, airports, cargo, passenger, efficiency and safety. The critical analysis of 63 empirical studies reveals that the usage of data mining in airlines sector is still in its early stages and the ability of these studies to generate knowledge may not be sufficient. It's tried to subhead papers and classified them subject by subject and also yearly. The intention of this study is that this is going to guide that sheds light on this field studies and staff. And also we want to put forth the tendency of civil aviation sector and evaluation progress of airlines in business intelligence concept. Given these findings, it can be suggested that there is more to dig for in order to obtain more managerially interpretable and acceptable results in further studies. Also, recommendations are made for other potentials of other business intelligence tools or data mining techniques research.

The structure of the paper is as follows. In the first section, critical definition and description associated with business intelligence, data mining and a brief introduction general aviation circumstances are presented. And also in that section some data mining techniques illustrated in civil aviation sector, with the intent of providing readers unfamiliar with techniques some basic concepts in airline to understand the rest of the paper. Section 3 includes the method followed in accomplishing the critical analysis of the empirical studies is presented. Section 4 concludes the article by providing some discussions on the future of soft computing in segmentation research.

II. LITERATURE, DEFINITIONS AND DESCRIPTIONS

According to the Oxford dictionary definition of data is that; the quantities, characters, or symbols on which operations are performed by a computer, which may be stored and transmitted in the form of electrical signals and recorded on magnetic, optical, or mechanical recording media [86]. When we look the short but imposing story too past too present, World technological capacity to store information has roughly doubled every 3 years since the 1980s [32]. Another fascinating thing, we create 2.5 quintillion bytes of data daily so much that 90% of the data in the world today has been created in the last two years alone [63]. Namely 1018 bytes or 1 billion gigabytes data is equal to more than half billion songs or 90 years of the HD videos [38]. To illustrate it in another way, in 1992 generated data was about 100 GB per day,

otherwise it is approximately 50 thousand per second at the present time and the data store grows 4 times faster than the world economies [80].

About the business intelligence definition, a term that encompasses all the capabilities required to turn data into intelligence, has emboldened companies to strive for the ultimate goal: getting the right information to the right people at the right time through the right channel [11]. Grujar et al. diverse description for business intelligence is an overarching term which are includes the applications, infrastructure and tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance [27]. The term "Business Intelligence" was originally devised by Richard Millar Devens in the 'Cyclopedia of Commercial and Business Anecdotes' in 1865, about 150 years ago. The BI terms if we are to be expressed in shortly is; gathering and analyzing process include whole business information [85].

Data mining describes from oxford dictionary very basically practice of examining large pre-existing databases in order to generate new information. It helps in extracting and refining useful knowledge from different types and size datasets. According to the description of Barai; obtain and aggregated information can be used to form a prediction or classification model, identify trends and associations, refine an existing model, or provide a summary of the datasets being mined [74]. Definition of the Han and Kamber from their books is Data mining is the task of discovering interesting patterns from large amounts of data, where the data can be stored in databases, data warehouses, or other information repositories [30]. From the second half of the 19th century data collection and database creation has thieved stepwise. The last quarter of before the millennium database management systems flourish with hierarchical, network and relational database systems theories. After 80's advanced studies have been put forward about those fields. And then after the widespread use of the Internet to begin in commercial area, XML and web based databased systems improved and data-information integration has spread [79]. Data mining is a step in the knowledge discovery and data mining process consisting of particular data mining algorithms that, under some acceptable computational efficiency limitations, produces a particular enumeration of patterns [24]. The civil aviation industry is one of the earliest adopters of data science with amounts of data in their systems. Airlines begin to use analytics and intelligence of technologies techniques end of the 60's. Over the years, this industry has been making large investments to mine data and explore opportunities to improve operational efficiency and boost customer loyalty.

On the other side, one of the crucial sectors in 20th

century was aviation sector, not only military area but also civil and commercial field, total number of passengers carried on scheduled services reached 3.7 billion in 2016, a 6.0 per cent increase over last year [61]. Obviously, the nature of data is critical to the success of data mining application. Exactly, the nature of the data is related to its source, utility, behavior and description.

The other face of the coin, aircraft always record data down in their black box. Planes equipped with flight recording data typically record up to 500 variables of data -described in these flight data recordings are time, altitude, vertical acceleration, and heading- per second for the duration the plane is being operated [66]. Finding patterns in aviation data manually is impracticable due to the mass amount of data produced every day [60]. Airlines companies' domain to better make sense of current and historical data, and make predictions using descriptive behavior, a scalable analytics service because of their real needed [8]

In the field of not only army aviation but also civil aviation sectors' large amounts of data are generated during studies. To demonstrate, on flights information, passengers' data and analysis, cost expenditure analysis, mainly fuel prices and employee cost, airspace of countries and other political relations and circumstances, measurements of wastes and damage to nature, technical affairs and mechanical matters like traffic signals and radar issues, aircraft interior design, advertising and other managerial issues and inventory etc. Based on all these data, decision-makers who are managers arrive at decision to solve a respective problem and optimize their resource. Managers should look out for ways to ease the pain in obtaining access to and applying disparate datasets.

III. METHOD

In this paper, the method which was followed in this article can be found in most of the literature studies that include the data mining and its tools and applications of the current body of knowledge studies of any literature. Considering the complexities around the logistics of running an airline and the tough competition, the ones that exploit sophisticated analytical tools are the ones who would ultimately enjoy a competitive edge over the rest. One of the important business intelligence tool is data mining which include data analytics, dynamics dashboard, efficient handling of complex and relational data, classification and clustering techniques are predicted to be one of the strongest transformation factors for airlines because of several reasons which are going to be explained from next paragraph. Data can improve ground operations, supporting faster turnaround times and better airspace management solutions which are

driving efficiencies and also give airlines a critical look at passengers to enable better and more personalized experiences for each passenger, which in turn drives brand loyalty, increases customer satisfaction, enables stronger auxiliary revenue stream and finally supports scheduling/rebooking of passengers when delays occur.

Due to both economic and social reasons several journal articles were examined about aviation sector. But in this paper, it focused data mining tools and applications related airlines and civil aviation. Moreover, the lots of studies regarding aircraft production and some engineering issues about aviation in particular till 2000 years. The article selection procedure was based on some significant criteria. The first criterion is that the studies should be in empirical nature and should be real-world data. The second one is related to the main purpose of using data mining techniques; the tools or the analysis should be included all echelon of data mining and analytics such as, integration, cleaning, transformation, pattern evaluation and knowledge. The third criterion is that only articles considering airlines market or aviation sector issues except human automation interaction of pilots or stewards, employees' survey works, accident and aircraft production. Exclusively data mining and analytics about airlines, airports, cargo, passenger, efficiency and safety. Also, the excluded studies are human resources and psychology. The last one is only journal articles were examined, publications in other forms, such as conference paper, book chapters and research reports were not included in the study. Hence, all empirical studies were collected through the publication databases depending on the availability of the access to these databases. However, it can be said that the majority of the well-known science and social science journals were searched. At the end of the searching process, a total of 63 studies were selected include data mining for evaluation. The earliest date of these publications is in 1998, whilst the latest one is in 2016. The majority of the studies were published in science related journals.

Appraised article categorize 8 different title by the subject are airlines, airlines market, airports, air cargo, efficiency, low-cost airlines, passenger issues, and safety. It is shown that the number of articles in Figure 1.

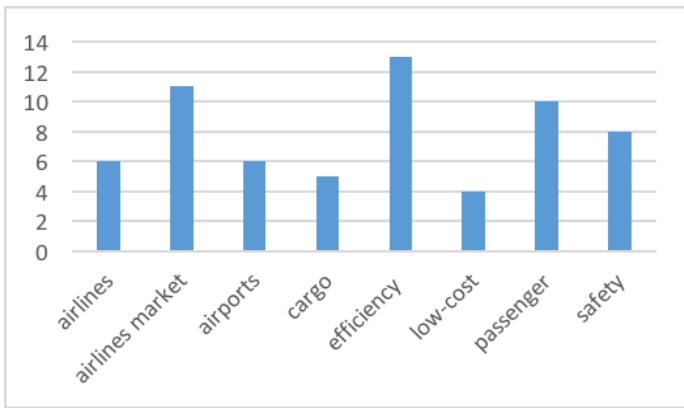


Fig. 1. Number of Article by Subject

A. Airlines

The airlines subject has general operations and process of civil aviation companies and airlines transaction and procedure [3, 17, 18, 65, 68, 83]. With a big data architecture, an airline can enrich historical data with real-time information to help forecast, while in flight, part failures. Alerts can be generated to an operations manager notifying them of a 99.5% chance for some part failing within the next two flights [78]. With this real-time alert, an airline can proactively have parts ready upon an aircraft's arrival, fix the part, and thereby reduce downtime, customer delays, overall maintenance checks, etc. One of the other segments where Predictive Analytics has been of big significance is around pricing of the flight tickets. Categorizing air tickets into different segments can lead to a better understanding of the demand-supply matrix, and the airlines can dynamically change their pricing to ensure all tickets are sold across different classes than having an empty flight due to conventional pricing.

One of the important studies in this title was written by Ryerson & Kim and they mentioned airlines hub optimization [65]. They used clustering and hierarchical algorithms which are popular tools in DM. Wu and et al. also refer to IATA delay codes for they simulated delay data and were benchmarked against the IATA system [83].

B. Airlines Marketing

The airlines market headline contains general marketing theme with financial and economics side [1, 6, 14, 19, 25, 26, 27, 41, 57, 59, 81]. Thanks to some of data mining techniques, oil expenses can be reduced as much as possible. To illustrate the last example, aircrafts ordinarily do not fly with full tank, because it causes really high damage in terms of cost. However, in the event of using data mining techniques, fuel cost optimization precisely, taking into

consideration weather conditions, passenger numbers, cargo weights, next flight, airport situation and whole lot more.

In this field, Arjomandi and Seufert evaluate the major airlines performances based on technical and environmental used with data envelopment analysis (DEA) techniques. They also classify the different regions and types of airlines according to IATA and efficiencies of the airlines. Their results are really empirical and they advocate that low cost airlines more environmentally oriented [6]. The other remarkable article was written by Cosmas et al. and aforementioned article is about clustering airlines performance in US market.

C. Airports and Cargo

In airports enclose airport management, operations and their analysis [2, 9, 23, 40, 50, 67]. Data mining has helped in analyzing passenger-movements and behaviors right from the time passenger arrives at the airport until they depart, to avoid possible bottlenecks at the security check, immigration counter, or even at the baggage counter. This will help to predict passenger flows to the identified critical bottlenecks. Based on this forecast, airline can efficiently do capacity management to help minimize long queues, which also results to provide a more pleasant travel experience for the customers. One of the most comprehensive article about the airports is Adler and Berechmans'. They used DEA for the measuring airports qualities.

Under the cargo title, studies regarding air cargo optimization and assessment [34, 52, 69, 82, 84]. In this headline all of the articles relevant Asian Airlines which are China, Hong Kong, Taiwan and Korea.

D. Efficiency

The heading efficiency is one of the most studied one, because every data mining works implicated optimization and efficiency tools [13, 15, 21, 31, 37, 39, 42, 44, 48, 55, 62, 75, 77]. The writers also use DEA, Network and Comparative Analysis and Classification.

Utilizing data mining during normal operations, airlines can combine historical data with real-time weather information, sensor data, and other pertinent data to optimize the flight path such as; altitude, speed or time thereby minimizing fuel and other cost. Namely, track weather and wind patterns in-flight in real-time to optimize flight path if needed and also reduce fuel-consumption. Proactively reduce flight delays by analyzing data from different flight paths tracked over time to expand their search for shorter routes around the globe – cutting costs on fuel consumption.

E. Low-Cost-Company

Farther there are some studies about low-cost airlines or low-cost carrier is an airline that generally has lower fares but also low comfort and quality [10, 16, 42, 64]. Not talked about inward but the studies were shown under a different title, because they were in that context exclusively. Generally low-cost studies include comparative analysis not only with amongst themselves and also with normal airlines [10, 42].

F. Passenger

One of the other title is passenger which include loyalty, rate, management, decision, optimization and also clustering [4, 20, 29, 36, 46, 47, 49, 51, 53, 58]. Utilizing historical information on customers' purchase history, in-flight entertainment -which is called IFE- browsing history, social media feeds, partner info (what hotel chain they prefer) an airline has the ability to know the customer information which are complete view enterprises gain into their customers. In real-time, an airline knows where a plane is headed, where a passenger is sitting, and what a passenger is viewing on the IFE or connectivity system watching a movie, browsing for hotels, looking for skis. Through a combination of the historically generated customer information, plus the current real-time information being gathered, specific and personalized advertising or programming can be offered from airline partners, thereby increasing ancillary revenues and improving customer satisfaction. And also, with historical customer knowledge and data regarding individual passengers' day-of-travel plans, an airline can offer customized price and once again improve customer satisfaction. At the present time, digital data enabled through connectivity is transforming the aviation industry including airlines and all other stakeholders in the aviation ecosystem, and the World digitalize rapidly. Additionally, an airline has specific passenger information for each flight. For example, airline companies have number of executive or platinum passengers, revenue per flight, profit per flight, cost per seat or more detailed one catering and handling expenses per seat.

G. Safety & Security

The last but not the least, due to safety and security is piece de resistance for aviation industry, during disruptions or emergency, airlines not only can optimize flight paths, but they also can choose to divert plans in each case in particularly assigned to each particular plane [7, 12, 45, 70, 71, 72, 76]. Real-time information on the performance of engines with the help of sensors installed to detect the likelihood of a failure to ensure best practices are followed to continuously improve

maintenance procedures and ensure maximum safety to passengers.

The salient paper in this title written by Li et al. which is specific modelling for flight operation and safety monitoring[45]. Predictive analytics which is one of the business intelligence tools may also help identify whether aircraft parts or the engine require maintenance or replacement. By predicting the likelihood of equipment failure, analytics can help airlines be more proactive and set best practices to continuously improve their maintenance procedures to ensure maximum safety to its passengers.

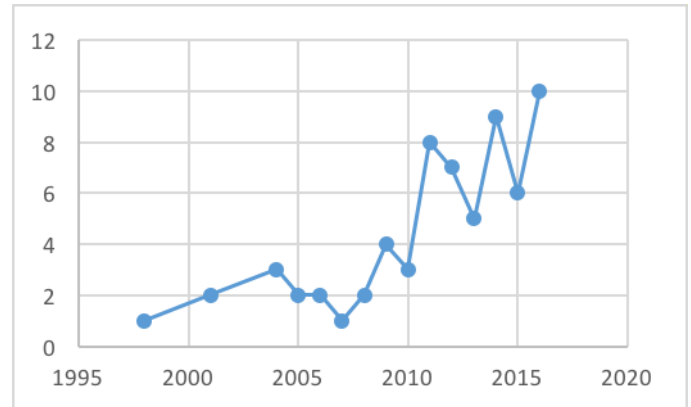


Fig. 2 Published Years by Article

When we check the timeline of the published paper, obviously increase emerged, especially after 2010. The following section of this study will provide conclude and some discussion in this field and future works comment.

IV. CONCLUSION, DISCUSSION, FUTURE WORK

The data mining applications include handling spatial data, engineering design data, hypertext and multimedia data, time-related data, stream data, and the World Wide Web [30]. Nowadays, one of the most significant and fast developing sector is civil aviation has lots of variables data -described in these flight data recordings are time, altitude, vertical acceleration, and heading- per second for the duration the plane is being operated. Flight cancellations, delays, and the subsequent expenditure on re-arrangements can result in millions of dollars in losses and concerned customers. In such a scenario, a re-accommodation system that would alert a passenger in case of flight delay or cancellation will be a win-win solution for both the airline company and the passengers. The other services like 'In-Flight recommender systems' that recommend frequent-flyers their preferred refreshments—based on historical data—will give the passengers a top-notch experience and also help airline companies boost ancillary sales.

For the future, extensive use of Internet of Things (IoT) technologies and analytics in the future will help the aviation industry evolve as everything will be completely monitored and orchestrated by sensors, thus improving their operational efficiency and giving passengers a legendary experience.

Actually there are a lot of different phrases and terms derived in after the 2000 and especially last decade. Some of them are evolved previous version and some of them used generally commercial issued. For instance, Fremont Rider has mentioned high volume data in 1944 and B. A. Marron and P. A. D. de Maine published about data comparison in however Michael Cox and David Ellsworth publish “Application-controlled demand paging for out-of-core visualization” was the first article to use the term “big data. That is really different and interesting subject and it’s going to be good proposal for future works. On the other hand, IoT, Machine Learning, Artificial Intelligence, Social Media Mining, Advertising Mining and Industry 4.

On the other hand, cybersecurity is a must as aircraft become nodes in the IoT environment. Aircraft are increasingly becoming internet-enabled machines – which can be prone to hacking and attacks. As regulators look closer at cybersecurity we must do everything we can to protect passenger information, payment details and IP of both on-board and streaming content. That is why we take a holistic view of security working with stakeholders to exceed applicable security requirements for connected systems.

REFERENCES

[1] Abdelghany, A., Guzhva, V. S., “Analyzing airlines market service using panel data,” *Journal of Air Transport Management*, vol. 16(1), pp. 20-25, 2010.

[2] Adler, N., Berechman, J., “Measuring airport quality from the airlines’ viewpoint: an application of data envelopment analysis,” *Transport Policy*, vol. 8(3), pp. 171-181, 2001.

[3] Adler, N., Golany, B., “Evaluation of deregulated airline networks using data envelopment analysis combined with principal component analysis with an application to Western Europe,” *European Journal of Operational Research*, vol. 132(2), pp. 260-273, 2001.

[4] Akartunalı, K., Boland, N., Evans, I., Wallace, M., Waterer, H., “Airline planning benchmark problems—Part I: Characterising networks and demand using limited data,” *Computers and Operations Research*, 40(3), pp. 775-792, 2013.

[5] Akerkar, R., “Analytics on big aviation data: turning data into insights,” *International Journal of Computer Science and Applications*, vol. 11, pp. 116-127, 2014.

[6] Arjomandi, A., Seufert, J. H., “An evaluation of the world’s major airlines’ technical and environmental performance,” *Economic Modelling*, vol. 41, pp. 133-144, 2014.

[7] Asgary, A., Ansari, S., Duncan, R., Pradhan, S., “Mapping potential airplane hazards and risks using airline traffic data,” *International journal of disaster risk reduction*, vol. 13, pp. 276-280, 2015.

[8] Ayhan, S., Pesce, J., Comitz, P., Sweet, D., Bliesner, S., Gerberick, G., “Predictive analytics with aviation big data,” *Integrated Communications, Navigation and Surveillance Conference (ICNS)*, 2013. IEEE, 2013.

[9] Bilotkach, V., Lakew, P. A., “On sources of market power in the airline industry: Panel data evidence from the US airports,” *Transportation Research Part A: Policy and Practice*, vol. 59, pp. 288-305, 2014.

[10] Bitzan, J., Peoples, J., “A comparative analysis of cost change for low-cost, full-service, and other carriers in the US airline industry,” *Research in Transportation Economics*, vol. 56, pp. 25-41, 2016.

[11] Buytendijk, F., “Dealing with dilemmas: where business analytics fall short.” Chichester: Wiley, 2010.

[12] Chen, W., Li, J., “Safety performance monitoring and measurement of civil aviation unit,” *Journal of Air Transport Management*, vol. 57, pp. 228-233, 2016.

[13] Chiou, Y. C., Chen, Y. H., “Route-based performance evaluation of Taiwanese domestic airlines using data envelopment analysis,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 42(2), pp. 116-127, 2006.

[14] Cosmas, A., Love, R., Rajiwade, S., Linz, M., “Market clustering and performance of US OD markets,” *Journal of Air Transport Management*, vol. 28, pp. 20-25, 2013.

[15] Dang, Y., Peng, L., “Hierarchy of air freight transportation network based on centrality measure of complex networks,” *Journal of Transportation Systems Engineering and Information Technology*, vol. 12(3), pp. 109-114, 2012.

[16] Daraban, B., Fournier, G. M., “Incumbent responses to low-cost airline entry and exit: A spatial autoregressive panel data analysis,” *Research in Transportation Economics*, vol. 24(1), pp. 15-24, 2008.

[17] Demirtas, O., “Evaluating the core capabilities for strategic outsourcing decisions at aviation maintenance industry,” *Procedia-Social and Behavioral Sciences*, vol. 99, pp. 1134-1143, 2013.

[18] Derudder, B., Witlox, F., “On the use of inadequate airline data in mappings of a global urban system,” *Journal of Air Transport Management*, vol. 11(4), pp. 231-237, 2005.

[19] Devriendt, L., Burghouwt, G., Derudder, B., De Wit, J., Witlox, F., “Calculating load factors for the transatlantic airline market using supply and demand data—a note on the identification of gaps in the available airline statistics,” *Journal of Air Transport Management*, vol. 15(6), pp. 337-343, 2009.

[20] Dupuis, C., Gamache, M., Pagé, J. F., “Logical analysis of data for estimating passenger show rates at Air Canada,” *Journal of Air Transport Management*, vol. 18(1), pp. 78-81, 2012.

[21] Duygun, M., Prior, D., Shaban, M., Tortosa-Ausina, E., “Disentangling the European airlines efficiency puzzle: A network data envelopment analysis approach,” *Omega*, vol. 60, pp. 2-14, 2016.

[22] ECFR — Code of Federal Regulations. (n.d.). Retrieved March 09, 2017, from http://www.ecfr.gov/cgi-bin/textidx?tpl=%2Fecfrbrowse%2FTitle14%2F14cfr121_main_02.tpl

[23] Fan, H., Tarun, P. K., Shih, D. T., Kim, S. B., Chen, V. C., Rosenberger, J. M., Bergman, D., “Data mining modeling on the environmental impact of airport deicing activities,” *Expert Systems with Applications*, vol. 38(12), pp. 14899-14906, 2011.

[24] Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., Uthursamy, R., “Advances in knowledge discovery and data mining.” AAAI Press/The MIT Press, Cambridge, MA, 1996.

[25] Gudmundsson, S. V., “New-entrant airlines’ life-cycle analysis: growth, decline and collapse,” *Journal of Air Transport Management*, vol. 4(4), pp. 217-228, 1998.

[26] Gürbüz, F., Özbakir, L., Yapici, H., “Classification rule discovery for the aviation incidents resulted in fatality,” *Knowledge-Based Systems*, vol. 22(8), pp. 622-632, 2009.

[27] Gürbüz, F., Özbakir, L., Yapici, H., “Data mining and preprocessing application on component reports of an airline company in Turkey,” *Expert Systems with Applications*, vol. 38(6), pp. 6618-6626, 2011.

[28] Gurjar, Y. S., Rathore, V. S., “Cloud business intelligence—is what business need today,” *International Journal of Recent Technology and Engineering*, vol. 1(6), pp. 81-86, 2013.

[29] Haensel, A., Koole, G., Erdman, J., “Estimating unconstrained customer choice set demand: A case study on airline reservation data,” *Journal of Choice Modelling*, vol. 4(3), pp. 75-87, 2011.

- [30] Han, J., Kamber, M., "Data mining: concepts and techniques (the Morgan Kaufmann Series in data management systems)", 2000.
- [31] Hess, S., Adler, T., Polak, J. W., "Modelling airport and airline choice behaviour with the use of stated preference survey data," *Transportation Research Part E: Logistics and Transportation Review*, vol. 43(3), pp. 221-233, 2007.
- [32] Hilbert, M., López, P., "The World's Technological Capacity to Store, Communicate, and Compute Information," *Science*, vol. 332(6025), pp. 60-65, 2011.
- [33] Huang, J. Y., Kao, Y. C., Lu, W. C., Shieh, J. C. P., "An inductive literature review of REITs by data mining methods," *Applied Engineering Sciences: Proceedings of the 2014 AASRI International Conference on Applied Engineering Sciences*, Hollywood, LA, USA, vol. 1. CRC Press, 2014.
- [34] Hui, G. W., Van Hui, Y., Zhang, A., "Analyzing China's air cargo flows and data," *Journal of Air Transport Management*, vol. 10(2), pp. 125-135, 2004.
- [35] IATA. (2016, June). Annual Review 2016, 72nd Annual General Meeting. <http://www.iata.org/about/Documents/iata-annual-review-2015.pdf>, (06.06.2016).
- [36] Jambulingam, T., Kathuria, R., Doucette, W. R., "Entrepreneurial orientation as a basis for classification within a service industry: the case of retail pharmacy industry," *Journal of Operations Management*, vol. 23(1), pp. 23-42, 2005.
- [37] Kasturi, E., Devi, S. P., Kiran, S. V., Manivannan, S., "Airline Route Profitability Analysis and Optimization Using BIG DATA Analytics on Aviation Data Sets under Heuristic Techniques," *Procedia Computer Science*, 87, pp. 86-92, 2016.
- [38] Khoso, M. (2016, July 14). How Much Data is Produced Every Day? Retrieved March 09, 2017, from <http://www.northeastern.edu/levelblog/2016/05/13/how-much-data-produced-every-day/>
- [39] Kilic, H. S., Zaim, S., Delen, D., "Development of a hybrid methodology for ERP system selection: The case of Turkish Airlines," *Decision Support Systems*, vol. 66, pp. 82-92, 2014.
- [40] Kilkis, S., Kilkis, S. "Benchmarking airports based on a sustainability ranking index," *Journal of Cleaner Production*, 130, pp. 248-259, 2016.
- [41] Lantseva, A., Mukhina, K., Nikishova, A., Ivanov, S., Knyazkov, K., "Data-driven Modeling of Airlines Pricing," *Procedia Computer Science*, 66, pp. 267-276, 2015.
- [42] Lee, B. L., Worthington, A. C., "Technical efficiency of mainstream airlines and low-cost carriers: New evidence using bootstrap data envelopment analysis truncated regression," *Journal of Air Transport Management*, vol. 38, pp. 15-20, 2014.
- [43] Lee, B. L., Worthington, A. C., "Technical efficiency of mainstream airlines and low-cost carriers: New evidence using bootstrap data envelopment analysis truncated regression," *Journal of Air Transport Management*, vol. 38, pp. 15-20, 2014.
- [44] Lesot, M. J., Rehm, F., Klawonn, F., Kruse, R., "Prediction of aircraft flight duration," *IFAC Proceedings Volumes*, 39(12), pp. 107-112, 2006.
- [45] Li, L., Hansman, R. J., Palacios, R., Welsch, R., "Anomaly detection via a Gaussian Mixture Model for flight operation and safety monitoring," *Transportation Research Part C: Emerging Technologies*, vol. 64, pp. 45-57, 2016.
- [46] Liou, J. J., "A novel decision rules approach for customer relationship management of the airline market," *Expert systems with Applications*, vol. 36(3), pp. 4374-4381, 2009.
- [47] Liou, J. J., "Variable Consistency Dominance-based Rough Set Approach to formulate airline service strategies," *Applied Soft Computing*, vol. 11(5), pp. 4011-4020, 2011.
- [48] Liou, J. J., "Developing an integrated model for the selection of strategic alliance partners in the airline industry," *Knowledge-Based Systems*, vol. 28, pp. 59-67, 2012.
- [49] Liou, J. J., Tzeng, G. H., "A dominance-based rough set approach to customer behavior in the airline market," *Information Sciences*, vol. 180(11), pp. 2230-2238, 2010.
- [50] Liou, J. J., Tang, C. H., Yeh, W. C., Tsai, C. Y., "A decision rules approach for improvement of airport service quality," *Expert Systems with Applications*, vol. 38(11), pp. 13723-13730, 2011.
- [51] Liou, J. J., Yen, L., Tzeng, G. H., "Using decision rules to achieve mass customization of airline services," *European Journal of Operational Research*, vol. 205(3), pp. 680-686, 2010.
- [52] Lo, W. W. L., Wan, Y., Zhang, A., "Empirical estimation of price and income elasticities of air cargo demand: The case of Hong Kong," *Transportation Research Part A: Policy and Practice*, vol. 78, pp. 309-324, 2015.
- [53] Majeske, K. D., Lauer, T. W., "Optimizing airline passenger prescreening systems with Bayesian decision models," *Computers and Operations Research*, vol. 39(8), pp. 1827-1836, 2012.
- [54] Miller Devens, R., "Cyclopedia of Commercial and Business Anecdotes; Comprising Interesting Reminiscences and Facts, Remarkable Traits and Humors of Merchants, Traders, Bankers Etc. in All Ages and Countries." New York and London, D. Appleton and company, 1864.
- [55] Min, H., Joo, S. J., "A comparative performance analysis of airline strategic alliances using data envelopment analysis," *Journal of Air Transport Management*, vol. 52, pp. 99-110, 2016.
- [56] Moss, Larissa T., Atre, S. "Business intelligence roadmap: the complete project lifecycle for decision-support applications." Addison-Wesley Professional, 2003.
- [57] Mumbower, S., Garrow, L. A., Higgins, M. J., "Estimating flight-level price elasticities using online airline data: A first step toward integrating pricing, demand, and revenue optimization," *Transportation Research Part A: Policy and Practice*, vol. 66, pp. 196-212, 2014.
- [58] Notomista, G., Selvaggio, M., Sbrizzi, F., Di Maio, G., Grazioso, S., Botsch, M., "A fast airplane boarding strategy using online seat assignment based on passenger classification," *Journal of Air Transport Management*, vol. 53, pp. 140-149, 2016.
- [59] O'Connor, K., Fuellhart, K., "Cities and air services: the influence of the airline industry," *Journal of Transport Geography*, vol. 22, pp. 46-52, 2012.
- [60] Pagels, D. A., "Aviation Data Mining," *Scholarly Horizons: University of Minnesota, Morris Undergraduate Journal*, vol. 2(1), Article 3, 2015.
- [61] Philbin, A., "Traffic growth and airline profitability were highlights of air transport in 2016," Retrieved March 09, 2017, from <http://www.icao.int/Newsroom/Pages/traffic-growth-and-airline-profitability-were-highlights-of-air-transport-in-2016.aspx>
- [62] Ravizza, S., Chen, J., Atkin, J. A., Stewart, P., Burke, E. K., "Aircraft taxi time prediction: comparisons and insights," *Applied Soft Computing*, vol. 14, pp. 397-406, 2014.
- [63] Researcher, I. (2016, October 31). Bringing big data to the enterprise. Retrieved March 09, 2017, from <https://www-01.ibm.com/software/data/bigdata/what-is-big-data.html>
- [64] Rey, B., Myro, R. L., Galera, A., "Effect of low-cost airlines on tourism in Spain: A dynamic panel data model," *Journal of Air Transport Management*, vol. 17(3), pp. 163-167, 2011.
- [65] Ryerson, M. S., Kim, H., "Integrating airline operational practices into passenger airline hub definition," *Journal of Transport Geography*, vol. 31, pp. 84-93, 2013.
- [66] S. Das, B. L. Matthews, A. N. Srivastava, N. C. Oza, "Multiple kernel learning for heterogeneous anomaly detection: algorithm and aviation safety case study," In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 47-56. ACM, 2010.
- [67] Sarkis, J., Talluri, S., "Performance based clustering for benchmarking of US airports," *Transportation Research Part A: Policy and Practice*, vol. 38(5), pp. 329-346, 2004.
- [68] Scheraga, C. A., "Operational efficiency versus financial mobility in the global airline industry: a data envelopment and Tobit analysis," *Transportation Research Part A: Policy and Practice*, vol. 38(5), pp. 383-404, 2004.
- [69] Shiao, G. C., Hwang, C. C., "Analyzing competition of international air cargo carriers in the Asian general air cargo markets," *Transport Policy*, vol. 27, pp. 164-170, 2013.

- [70] Shyr, H. J., "A quantitative model for aviation safety risk assessment," *Computers and Industrial Engineering*, vol. 54(1), pp. 34-44, 2008.
- [71] Shyr, H. J., Keng, H., Ia-Ka, I., Huang, C. L., "Using extended hazard regression model to assess the probability of aviation event," *Applied Mathematics and Computation*, vol. 218(21), pp. 10647-10655, 2012.
- [72] Stolzer, A., "Tutorial B: Data Mining for Aviation Safety: Using Data Mining Recipe," *Automatized Data Mining from STATISTICA*, 2009.
- [73] Stolzer, A., "Tutorial B: Data Mining for Aviation Safety: Using Data Mining Recipe," *Automatized Data Mining from STATISTICA*, 2009.
- [74] S. K. Barai, "Data mining applications in transportation engineering," *Transport*, vol. 18(5), pp. 216-223, 2003.
- [75] Sun, X., Zimmermann, C. M., Jackson, G. P., Bunker, C. E., Harrington, P. B., "Classification of jet fuels by fuzzy rule-building expert systems applied to three-way data by fast gas chromatography—fast scanning quadrupole ion trap mass spectrometry," *Talanta*, vol. 83(4), pp. 1260-1268, 2011.
- [76] Tanguy, L., Tulechki, N., Urieli, A., Hermann, E., Raynal, C., "Natural language processing for aviation safety reports: from classification to interactive analysis," *Computers in Industry*, vol. 78, pp. 80-95, 2016.
- [77] Tesfay, Y. Y., "Modified panel data regression model and its applications to the airline industry: Modeling the load factor of Europe North and Europe Mid Atlantic flights," *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 3(4), pp. 283-295, 2016.
- [78] Thales, G. (2016, May 22). Big Data for aviation: a new revolution is coming. Retrieved March 11, 2017, from <http://onboard.thalesgroup.com/2016/05/23/big-data-aviation-new-revolution-coming>
- [79] Vercellis, C., "Business intelligence: data mining and optimization for decision making." Chichester: John Wiley & Sons, 2010.
- [80] Walker, B. (n.d.). Every Day Big Data Statistics – 2.5 Quintillion Bytes of Data Created Daily. Retrieved March 09, 2017, from <http://www.vcloudnews.com/every-day-big-data-statistics-2-5-quintillion-bytes-of-data-created-daily/>
- [81] Wen, C. H., Chen, T. N., Fu, C., "A factor-analytic generalized nested logit model for determining market position of airlines," *Transportation Research Part A: Policy and Practice*, vol. 62, pp. 71-80, 2014.
- [82] Wen, C. H., Tsai, M. C., Lin, C. H., "Classification and competition analysis of air cargo logistics providers: The case of Taiwan's high-technology industry," *Journal of Air Transport Management*, vol. 17(2), pp. 106-109, 2011.
- [83] Wu, C. L., Truong, T., "Improving the IATA delay data coding system for enhanced data analytics," *Journal of air transport management*, vol. 40, pp. 78-85, 2014.
- [84] Yoon, S. H., Park, J. W., "A study of the competitiveness of airline cargo services departing from Korea: Focusing on the main export routes," *Journal of Air Transport Management*, vol. 42, pp. 232-238, 2015.
- [85] Okkonen, J., Pirttimäki, V., Hannula, M., Lönnqvist, A., "Triangle of Business Intelligence, Performance Measurement and Knowledge Management," In *IInd Annual Conference on Innovative Research in Management*, May 9-11, Stockholm, Sweden. ISO 690, 2002.
- [86] Stevenson, A., "Oxford dictionary of English." New York, NY: Oxford U Press, 2010. Print.