Journal of Air Transport Management 67 (2018) 181–196

Contents lists available at ScienceDirect

Journal of Air Transport Management

journal homepage: www.elsevier.com/locate/jairtraman

Characterizing air traffic networks via large-scale aircraft tracking data: A comparison between China and the US networks

Pan Ren[a], Lishuai Li[b,*]

[a] Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong Special Administrative Region
[b] Department of Systems Engineering and Engineering Management, City University of Hong Kong, 83 Tat Chee Avenue, Hong Kong Special Administrative Region

ARTICLE INFO

Keywords:
Air traffic management
Air traffic network
Data mining
Data analytics

ABSTRACT

Air travel demand has continued to increase rapidly over the past decade, causing severe flight delays. To reduce such delays, Air Navigation Service Providers need to first understand the operational capacity and congestion risks associated with a network, and then develop strategies accordingly. However, limited studies have been conducted due to lack of data. New opportunities have arisen given the availability of large-scale aircraft tracking data and many other digitalized records of operations. In response, we develop a novel data-driven framework that characterizes the operational structure and dynamics of an air traffic network using actual tracking data. The framework includes several new statistical measures and data analytic techniques to summarize airspace availability, network structure, and utilization patterns. We then apply the framework to analyze the air traffic networks in China and the US. The results reveal distinctive characteristics of these two networks. Airspace availability for commercial flights is much more restricted in China than the US. The network in China has a clear structure with distinct utilization patterns, while the network in the US has a more flexible structure featuring complex dynamics. These operational differences indicate that China faces a greater chance of en-route congestion when compared with the US. The results also demonstrate that the data-driven approach is effective to identify the actual behavior and complexity of an air traffic network, which are not captured by existing methods.

1. Introduction

Air transport system capacity enhancements have failed to keep up with the increasing pace of demand growth all over the world, causing severe air traffic congestion and flight delays, which have had high economic costs and negative environmental effects (Ball et al., 2010). To reduce these delays, it is critical to understand the operational capacity, efficiency, and congestion risks associated with an air transport network, and to carry out strategic planning and tactical management interventions accordingly to mitigate these issues.

Despite extensive studies on air transport network, few research has been done to characterize the operational structure and dynamics of national-level air traffic networks, or compare how airspace is managed between different regions based on actual air traffic flows. A key limitation lies in the availability of operational data across different sources and regions. For example, aircraft tracks, air traffic control commands, fleet scheduling, these data were heavily regulated by national or regional agencies and airlines without proper information sharing among them. Air traffic service providers can rarely grasp a ‘big picture’ of the regional airspace.

New opportunities have arisen from the increasing availability of digitized air traffic data. For example, with the implementation of Automatic Dependent Surveillance – Broadcast (ADS-B), a satellite-based surveillance technology that tracks and broadcasts the location of each aircraft via satellite, it is now possible to track and analyze aircraft movement data on a global scale. With the appropriate analytical tools, post-event analysis can be carried out to examine the characteristics of the actual air traffic network.

Thus, the development of analytical tools to analyze large-scale multi-source operational data can significantly contribute to the improvement of air transport system. To support this effort, in this study, we aim to characterize the structure and dynamics of the air traffic network via a data-driven approach using large-scale aircraft tracking data. The analysis result will allow Air Navigation Service Providers (ANSP) to understand the current network better, identify deficiencies in Air Traffic Management (ATM) procedures, and provide recommendations for improving system capacity and efficiency.

Relevant literature exists in two groups. The first group of research focus on understanding the air transport system using network analysis, in which the air transport system is considered as a complex network

---

[*] Corresponding author.
E-mail addresses: panren2-c@my.cityu.edu.hk (P. Ren), Lishuai.li@cityu.edu.hk (L. Li).

https://doi.org/10.1016/j.jairtraman.2017.12.005
Received 22 September 2017; Received in revised form 8 December 2017; Accepted 16 December 2017

composed of many interlinked subsystems. Some of these studies use complex network theory to analyze the topological characteristics of the system (Guimera et al., 2005; Vespignani, 2012; Cook et al., 2015; He at al., 2004; Li and Cai, 2004). In these studies, the air transport system is abstracted to a directed/undirected, weighted/unweighted network, where nodes are represented by airports and edges are direct flights linking two airports. The network is evaluated and optimized using the robustness metrics in network theory, which include betweenness, degree, centrality, and connectivity (DeLaurentis et al., 2008). Wei et al, (2014) designed a robust air transportation network by maximizing algebraic connectivity to reduce air traffic congestion. However, these studies only considered a static graph, meaning that the dynamics of the network were not explicitly considered. To address this limitation, other studies have focused on studying how delays propagate in the network. Empirical studies using flight schedule and delay data have analyzed the causes for primary delays and assessed how they spread over the network, causing reactionary delays (Beatty et al., 1999; Fricke and Schultz, 2009; Hansen, 2002; Abdelghany et al., 2004; Jetzki, 2009; Rebollo and Balakrishnan, 2014). Furthermore, Fleurquin et al. (2013) defined metrics that quantify the level of network congestion and the macro-scale behavior of delay dynamics. Péter and Szabó (2012) proposed an exact mathematical model for large-scale dynamic networks from the perspective of control theory using a class of positive systems, which could be used to achieve minimized delays. Another stream of studies used queueing theory to model and simulate flight delays within a network (Pyrgiotis et al., 2013; Hansen et al., 2009; Long and Hasan, 2009; Peterson et al., 1995a, 1995b; Xu, 2007; Shah et al., 2005); however, the performance of these models relies on how well the underlying conceptual network can represent real air traffic operations.

Recent studies suggest that representing air transport systems as airport networks is an oversimplified approach, as it discards important operational information of ATM (i.e., flow management). Some scholars propose that the air transport network can be regarded as aggregated multi-layers of airline networks (Zanin and Lillo, 2013; Du et al., 2016; Belkoura et al., 2016), airport networks, air navigation route networks (Sun et al., 2015), and air traffic management networks (Wang et al., 2017). Although the findings of these studies helped to better conceptualize the complexity of air transport systems, there is much work to be done to obtain a more comprehensive picture. One drawback associated with the current work on multi-layer air transport networks is that this endeavor requires a significant amount of detailed information from ANSP and airlines that can construct the different layers of the system, such as air route information, sector maps, letter of agreements (LOA), airline route networks, and so on; this makes it cumbersome to use across the air transport networks of different regions.

The second group of relevant research focus on to how to analyze and use large-scale air traffic operational data. Several studies have been carried out that use clustering techniques on actual tracking data for air traffic flow identification. Eckstein (2009) developed a flight track taxonomy to decompose a set of radar tracks according to their lateral, vertical, and conformance segments, for the purpose of evaluating procedural conformance of individual flights in the terminal airspace. Gariel et al, (2011) developed an analytical framework that uses density-based clustering to learn the typical patterns of traffic flows, which are then used as benchmarks to monitor the behavior of an aircraft in a given airspace. Other studies use hierarchical or spectral clustering to identify air traffic flows to and from an airport (Rehm, 2010; Enriquez, 2013). Nevertheless, these studies focused on flow identification on a relatively small scale – i.e., at the terminal area or by examining flows to and from an airport, without further analysis on ATM operations. The exception to this is the work of Conde Rocha Murca et al. (2016), as this research team stepped forward along the direction of characterizing large-scale ATM operations. The authors developed a data mining framework to characterize the air traffic flows in the transition/terminal airspace, and demonstrated how to use the results to assess the performance of tactical operations in the New York region on a daily basis.

In summary, further research is needed to study the air transport system, understand its network's structure and dynamics at different scales, and ensure that the analysis results are useful for ATM improvement. The availability of large-scale operational data and the advancements in data mining techniques have created unprecedented opportunities for such research. The aim of this investigation is thus to develop a data-driven framework to analyze an air traffic network using aircraft-tracking data, including route availability, network structure, and utilization patterns. The framework includes an innovative statistical measure – a modified Ripley's K-function – to assess air route availability, clustering techniques to identify major air routes, network analysis to quantify the actual air traffic network structure, and spatial–temporal analysis to reveal airspace utilization patterns. In order to demonstrate the proposed framework, we apply it to analyze and compare the air traffic networks in China and the US using one month of historical flight-tracking data.

The remainder of this paper is organized as follows. Section 2 describes the aircraft-tracking data used in this study. Section 3 presents the proposed data-driven framework, including the algorithms and metrics used in each module of the framework. In Section 4, we conduct a comparative study of China and the US airspaces using the proposed framework. Finally, Section 5 summarizes our study and suggests some future research directions.

2. Dataset

Automatic Dependent Surveillance – Broadcast (ADS-B) is a key component of the US Next Generation Air Transportation System (NextGen) and it enables aircraft to track their positions and broadcast them via satellites (Gugliotta, 2009). ADS-B greatly enhances the safety of air travel by not only providing air traffic control with real-time, consistent, and visible position updates, but it also notes other aircraft equipped with ADS-B. Currently, 70% of all commercial passenger aircraft are equipped with an ADS-B transponder and the percentage is steadily increasing, as the ADS-B transponder will become mandatory for most aircraft around the world by 2020 (Flightradar24, 2015). Air traffic management decision makers can benefit significantly from such data when engaging in performance assessments, real-time monitoring, and strategic planning. Currently, several flight-tracking service providers operate a worldwide network of ADS-B receivers to collect and share live flight tracks, such as Flightrader24 and FlightAware.

The dataset used in this study features flight-tracking data collected from Flightradar24 every minute for 30 consecutive days, from November 1 to November 30, 2016, covering the airspace in China and the US. The geographic ranges of the airspace in China and the US are set as (17.37, 46.00, 92.86, 126.50) and (24.80, 49.20, −124.90, −60.40), respectively. The first two values are the boundary latitudes, while the last two values are the boundary longitudes. Fig. 1 shows all flight trajectories collected. For this research, we focus on air traffic networks consisting of the top 10 busiest airports in China and the US, as ranked by annual passenger traffic (CAAC, 2015a; ACI-NA, 2015). The information related to these airports is presented in Appendix A. In addition, we filter out airport pairs where the sample size is less than one flight per day. As a result, 40 airport pairs in China and 45 airport pairs in the US are left in our analysis. The flight trajectories of top 10 airports in China and the US are visualized in Fig. 2.

We choose to compare the air traffic network in China and the US because: 1) they are the busiest two regions in the world (ranked by passengers carried in 2015 (WBG, 2015)); 2) both are faced with significant flight delays (in 2015, 18.27% of flights were delayed and the average delay time relative to scheduled time was 11 min in the US (BTS, 2015), while 31.67% of flights were delayed and the average delay time was 21 min in China (CAAC, 2015b)); and 3) the airspaces