Journal of Air Transport Management xxx (2017) 1–12

Contents lists available at ScienceDirect

Journal of Air Transport Management

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

Identifying similar days for air traffic management

Sreeta Gorripaty [a, *], Yi Liu [b], Mark Hansen [c], Alexey Pozdnukhov [c]

[a] Uber Technologies Inc, San Francisco, CA, 94104, USA
[b] Amazon, Seattle WA, 98109, USA
[c] Institute of Transportation Studies, University of California, Berkeley, CA, 94720, USA

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Similar days
Clustering capacity and demand data
Decision support
Air traffic management

ABSTRACT

Air traffic managers face challenging decisions due to uncertainity in weather and air traffic. One way to support their decisions is to identify similar historical days, the traffic management actions taken on those days, and the resulting outcomes. We develop similarity measures based on quarter-hourly capacity and demand data at four case study airports—EWR, SFO, ORD and JFK. We find that dimensionality reduction is feasible for capacity data, and base similarity on principal components. Dimensionality reduction cannot be efficiently performed on demand data, consequently similarity is based on original data. We find that both capacity and demand data lack natural clusters and propose a continuous similarity measure. Finally, we estimate overall capacity and demand similarities, which are visualized using Metric Multidimensional Scaling plots. We observe that most days with air traffic management activity are similar to certain other days, validating the potential of this approach for decision support.

1. Introduction

Aviation contributes to the development and growth of the global economy, transporting people and enabling trade. Recent studies (Perkins, 2010) have shown that aviation contributes to 0.7–6% of a country's GDP (Gross Domestic Product), with approximately 4.9–5.2% being the contribution towards the US GDP. Increasing air travel demand has put immense pressure on existing aviation infrastructure (FAA, 2014) and necessitates the improvement of airport performance. Poor airport performance leads to immense losses to the airlines, passengers and the economy. A study by Zou and Hansen (2012) predicts the cost savings to US airlines alone of improved operational performance ranges from $7.1–13.5 billion for 2007.

Capacity is a critical input in operational performance and air traffic management decision-making. Depending on the context and the precise definition, airfield capacity may depend on factors such as weather conditions, fleet mix, traffic management actions, and human factors such as the experience of controllers (Newell, 1979; Venkatakrishnan et al., 1993). Owing to constraints on the airport infrastructure and regulations, airport capacity is often difficult to expand. For airports such as JFK (John F. Kennedy International Airport) and EWR (Newark Liberty International Airport) the capacity limitations result in an estimated loss of over $6 billion of lost travel spending in 2016 due to unmet demand (Eno, 2013).

One way to improve operational performance is to make better use of existing capacity. Unfortunately, capacity is difficult to predict on many days-of-operation, while penalties for releasing more flights for a given airport than its capacity can accommodate—such as airborne delay, diversions, and excess controller workload—are substantial. Thus, in many cases airports may provide conservative estimates of capacity that are an underestimate of the existing capacity. In the U.S., traffic specialists at the Air Traffic Control System Command Center, when faced with a possible demand-capacity imbalance at some airport, are responsible for planning traffic management initiatives (TMIs) that balance competing priorities of efficient capacity utilization and avoiding the over-delivery of flights.

The decisions made to manage air traffic play a significant role in keeping the operations in the national airspace system safe and efficient. Numerous studies in the past have developed algorithms and simulation studies to incorporate the stochastic nature of weather and capacity into decision-making (Ball et al., 2010; Cook and Wood, 2010; Dhal et al., 2013; Liu and Hansen, 2015; Mukherjee, 2004; Mukherjee and Hansen, 2007; Nilim and El Ghaoui, 2004; Provan et al., 2011; Smith and Sherry, 2008; Wang, 2011, 2012). Much of this literature concerns models and

* Corresponding author.
E-mail address: gorripaty@uber.com (S. Gorripaty).

http://dx.doi.org/10.1016/j.jairtraman.2017.06.005
0969-6997/

algorithms for making decisions that minimize the expected value of a loss function that takes into account ground delay and more expensive airborne delay. More recent contributions incorporate other performance goals, such as equity and predictability (Liu and Hansen, 2015). It has proven difficult, however, for such methods to gain traction with air traffic specialists, who generally place more stock in their own judgment and experience than recommendations from tools developed by researchers.

Recognizing this, the research community has in recent years turned more attention to developing tools that allow air traffic management specialists to tap their own experience. In essence, the idea is to identify days in the past that are similar to some reference day, and consider the TMI actions taken on these days and the operational outcomes that resulted. This information can be used in post-operational analysis to find if an experience in a recent day has echoes in the past, and for day-of-operations decision-making when the necessary forecasts are available.

With such applications in mind and in the light of recent advances in data mining, there has been considerable research on how to identify similar days in the National Airspace System (Bloem and Bambos, 2015; Grabbe et al., 2012, 2014; Kulkarni et al., 2013; Liu et al., 2014). The studies have used data mining algorithms to identify similar days using historical weather, traffic and capacity data. A variety of methods have been used, including clustering, Support Vector Machines (SVM), Ensemble Bagging Decision Trees (BDT) and Neural Networks. The research has been conducted on a range of geographic scales, from individual airports to the system as a whole. Features considered in assessing similarity also vary, including weather conditions, TMI actions, and operational outcomes. This paper contributes to similar days literature by developing methods for identifying similar days based on features of specific relevance to air traffic managers assessing—either prospectively or retrospectively—TMIs for balancing arrival demand and capacity at individual airports.

In this setting, the most salient features are profiles of arrival capacity and demand for a given day (in a day-of-operations application, these profiles would be derived from forecasts). Similarity between days depends on how closely both the demand and the capacity profiles match. We therefore propose and implement methods for assessing profile similarity and identifying similar profiles. There are three interesting research questions surrounding the investigation of similar days. First, is similarity between two days better viewed as a categorical or continuous variable? In other words, are there natural clusters of similar days, or simply a range of similarity without clear boundaries between similar and dissimilar? Second, how much guidance can similar days provide in situations in which TMI decisions may be required. Are TMI inclusive days more likely to be “odd balls” dissimilar from most historical days, or are TMI decisions often made in situations similar to those that have been encountered many times before? Third, how can the estimated similarity measure between pairs of days be used to aid efficient decision-making in air traffic management?

To address these questions we develop capacity and demand similarity metrics for historical days between 2007 and 2015 for four major U.S. airports: Newark (EWR), Kennedy (JFK), Chicago O'Hare (ORD), and San Francisco International (SFO). We then combine the demand and capacity similarity to determine a metric for overall (capacity-demand) similarity between days. We also characterize each day in terms of its level of TMI activity in order to assess the availability of similar days for the subset of days with significant TMIs. Finally, we offer case studies on how the similarity measures can be used to make better decisions in the context of air traffic management.

The remainder of this paper is organized as follows. Section 2 describes the data used and pre-processing done for the analysis. Section 3 presents the correlation analysis performed to explore the data. In section 4, we perform PCA on capacity data at the four airports and summarize the results. Section 5 describes the clustering analysis on capacity data and discusses the possibility of developing a discrete measure of similarity between days. PCA on the demand data is presented in section 6. Section 7 presents clustering analysis on demand data with a discussion on developing discrete measures of similarity based on demand data. In section 8, we identify similar days using continuous measures of similarity and visually represent the similarity between days with different levels of TMIs. Section 9 summarizes the findings of the study and identifies future research needed to make similar days a useful decision support tool.

2. Data

We use quarter-hour Aviation System Performance Metrics (ASPM) data for the analysis. We use two variables from the data: arrival demand and capacity. Our dataset covers the period from January 2007 through August 2015, and covers four airports: EWR (Newark Liberty International Airport), SFO (San Francisco International Airport), ORD (O'Hare International Airport) and JFK (John F. Kennedy International Airport).

The arrival demand variable in ASPM dataset reflects the total number of flights that would land in a given quarter-hour time period in the absence of capacity constraints. This includes flights that have been held at their departure airports or while en route in order to avoid queues from building in the terminal area of the arrival airport. We used these data to compute a “new” demand variable for each time period by subtracting flights that were also contributing to demand in the previous time period. This is done to avoid counting the same flights in multiple time periods and estimate the real demand experienced by the airport. The demand data at an airport for an observation, which is one day, is a vector of quarter-hourly demand data for that day.

The capacity variable is the Airport Acceptance Rate (AAR), which reflects FAA estimates of the number of arrival flights that could land in each quarter-hour time period. Discussions with Air Traffic Controllers (ATCs) reveal that the AAR may not be a reliable estimate of airport capacity, particularly when demand is low enough that estimation errors are inconsequential. However, during the hours of the day with a high demand, the controllers are likely to be more attentive to update the AAR regularly due to higher pressure to provide more reliable airport capacity estimates. To take advantage of this behavior, we only use hours that have a high demand-capacity ratio for our analysis. To determine which hours are appropriate, the median value of the demand-capacity ratio is calculated for each quarter-hour and airport, and the hours for which this value is over 0.75 were identified. (For this purpose, we used the original demand variable reported in ASPM rather than the “new” demand.) Thus, on the majority of days in these hours, demand is at least 75 percent of the announced capacity (AAR). Such periods can be viewed as busy and we can expect more accurate estimates of AAR in them. On this basis, we identify contiguous time periods in the day in which the airport is usually busy (These contiguous periods may contain a few short periods of lower demand). The hours selected as the busy period are: 7 to 22 for EWR; 7 to 22 for JFK, 8 to 22 for SFO; 7 to 20 for ORD. The capacity data at an airport for an observation, which is one day, is a vector of quarter-hourly AAR data for that day. It is important to note here that one day means different time windows for each of the airports.

Please cite this article in press as: Gorripaty, S., et al., Identifying similar days for air traffic management, Journal of Air Transport Management (2017), http://dx.doi.org/10.1016/j.jairtraman.2017.06.005