Transportation Research Part C 69 (2016) 1–15

Contents lists available at ScienceDirect

Transportation Research Part C

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

A methodology for identifying similar days in air traffic flow management initiative planning

Kenneth D. Kuhn

RAND Corporation, United States

A R T I C L E I N F O

Article history:
Received 16 October 2015
Received in revised form 20 May 2016
Accepted 22 May 2016
Available online 27 May 2016

Keywords:
Air traffic control
Air traffic flow management
Ground delay program
Cluster analysis
Machine learning

A B S T R A C T

This article describes a methodology for selecting days that are comparable in terms of the conditions faced during air traffic flow management initiative planning. This methodology includes the use of specific data sources, specific features of calendar days defined using these data sources, and the application of a specific form of classification and then cluster analysis. The application of this methodology will produce results that enable historical analysis of the use of initiatives and evaluation of the relative success of different courses of action. Several challenges are overcome here including the need to identify the appropriate machine learning algorithms to apply, to quantify the differences between calendar days, to select features describing days, to obtain appropriate raw data, and to evaluate results in a meaningful way. These challenges are overcome via a review of relevant literature, the identification and trial of several useful models and data sets, and careful application of methods. For example, the cluster analysis that ultimately selects sets of similar days uses a distance metric based on variable importance measures from a separate classification model of observed initiatives. The methodology defined here is applied to the New York area, although it could be applied by other researchers to other areas.

1. Introduction

The Federal Aviation Administration (FAA) oversees commercial air carriers and other organizations involved in civil aviation in the United States. The FAA and carriers work together to ensure safe and efficient operations, in part through air traffic control. For example, dispatchers at airline operations centers (AOCs) will delay, reroute, and cancel flights based on strategic-level guidance from the FAA Air Traffic Control System Command Center (ATCSCC). This guidance includes air traffic flow management (TFM) initiatives. An example of such an initiative would be when flights scheduled to fly to an airport when and where there will be thunderstorm activity are delayed at their origin airport. This particular type of TFM action is known as a Ground Delay Program (GDP). Safety and efficiency gains are achieved by avoiding routing aircraft through potentially hazardous conditions and by replacing airborne delay with ground delay. Other forms of TFM initiatives include ground stops, reroute advisories, flow metering (via miles-in-trail or minutes-in-trail restrictions), airspace flow programs, fix balancing, and capping.

There were 1305 GDPs implemented in 2006 (Kotnyek and Richetta, 2006). The costs to the economy of the resulting flight delays were substantial. Delays at certain airports, particularly busy airports in the New York area, can propagate throughout the air transportation system (Churchill et al., 2010; Pyrgiotis et al., 2013). GDPs can have environmental impacts that translate to euro or dollar costs that are roughly the same order of magnitude as direct economic costs (Carlier et al.,

E-mail address: kkuhn@rand.org

http://dx.doi.org/10.1016/j.trc.2016.05.014
0968-090X/

2007). In summary, GDPs and other TFM actions have important economic and other implications. This is reflected in the large body of research proposing novel methodologies for making strategic air traffic control decisions. Researchers have, in particular, proposed many forms of decision support tools, with one study claiming that "computer-based decision support systems might improve performance significantly" (Lulli and Odoni, 2007).

Analysis of the past use of TFM initiatives can demonstrate the relative success of courses of action but must account for the conditions faced during planning and operations. An identification of days that are similar can help. Analysts could, for example, focus on the 10 days in the past two years when there was thunderstorm activity at the key airports in the New York area between 8am and 11am, local time, but clear weather the rest of the day. The analysts could then examine what actions were taken during these days, why these actions were taken, and what the results of these actions were. As one study reported, "clustering techniques appear to be promising methods for identifying the major causes of Ground Delay Programs" (Grabbe et al., 2013). This explains the recent proliferation of research articles describing methods for clustering or classifying hours or days based on TFM conditions (Grabbe et al., 2013, 2014; Hoffman et al., 2003; Liu et al., 2014; Mukherjee et al., 2013, 2014; Smith and Sherry, 2009).

This article describes a methodology for characterizing and clustering days based on forecast aviation weather and scheduled air traffic conditions in a region, the inputs for TFM decision making. The goal is different than the goals of prior studies, explaining the need for a new methodology.

Mukherjee et al. (2014) and Smith and Sherry (2009) describe models that forecast GDPs and associated delays. Grabbe et al. (2013) apply k-means cluster analysis to data describing observed GDPs and then, separately, apply Expectation Maximization cluster analysis to data describing GDP initiative start and end times, observed airport weather, and airport arrival rates. Hoffman et al. (2003) identify similar days by applying cluster analysis to data describing the performance of the air transportation system, such as counts of flights canceled by day. Mukherjee et al. (2013) cluster days according to the impact of inclement weather on scheduled air traffic. The authors use observed weather data. Similarity among days in the cited studies is defined in terms of observed GDP actions, observed weather, and observed air transportation system performance. The cited studies will be more useful for analysts modeling system performance than this paper, but less useful for those interested in TFM decision making.

Liu et al. (2014) define a semi-supervised methodology for modeling the similarity between different weather forecasts at an airport. The authors begin by labeling a pair of hours as 'similar' whenever the runway configuration and meteorological conditions (IMC vs.VMC) are the same at the airport and airport arrival rates are within some threshold quantity from one another. Grabbe et al. (2014) find clusters of hours where the probabilities of there being a GDP at a particular airport were similar according to the authors own model of GDP activity, which is based on "historical weather data, scheduled arrival rates and traffic flow management logs." The models in the cited studies affix labels to particular hours or pairs of hours based upon conditions at a particular airport. Researcher defined labels are not used here, aside from the end product assignment of calendar days to clusters of similar days. Cluster analysis is applied directly to weather and air traffic data. The focus is on days rather than hours and on what happens in the New York area, a region that includes multiple airports. The New York area, like a handful of regions around the country, is a *metroplex* where there are important interactions among busy airports in close proximity to one another. It is also particularly important as the known source of substantial delays that propagate through the nation's air transportation system (Churchill et al., 2010; Pyrgiotis et al., 2013).

The methodology defined in this article includes the use of specific data sources, specific features of calendar days defined using these data sources, and the application of a specific form of classification and then cluster analysis. Note that developing a model of current traffic management decision making is not the primary goal. The methodology is applied to the New York area. The New York area is small enough that one might reasonably expect to be able to summarize aviation weather using a relatively small discrete set of possible condition states. It is also a region that is relatively busy in terms of air traffic and that generates a large volume of air traffic flow management initiatives.

There are several challenges here including the need to identify the appropriate machine learning algorithms to apply, to define distances between calendar days, to select features describing days, to obtain appropriate raw data, and to evaluate results in a meaningful way. These challenges are overcome via a review of relevant literature, the identification and trial of several useful models and data sets, and careful application of methods. For example, the cluster analysis that ultimately selects sets of similar days uses a distance metric based on variable importance measures from a separate classification model of initiatives. The following sections of this article describe: data collection, feature selection, and cluster and classification analysis.

2. Data collection

An archive of historical data covering hundreds of calendar days from the recent past is required for the development of statistical models that identify sets of similar days. The work presented here is based on publicly available data and data from the set described in Eshow et al. (2014) that are available to many aviation systems researchers. This ensures that other researchers can apply the methodology defined in this article. This section describes the ideal data for the specific problem faced here, the identification of similar days based on forecast aviation weather and scheduled air traffic conditions, and then compares the ideal to publicly available data. The section finishes with a description of the actual data used in the work presented here.