Journal of Air Transport Management 36 (2014) 120–128

Contents lists available at ScienceDirect

Journal of Air Transport Management

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

An airline connection builder using maximum connection lag with greedy parameter selection

Adam Seredyński [a,b,*], Franz Rothlauf [a], Tobias Grosche [b]

[a] Department of Information Systems & Business Administration, Johannes-Gutenberg University Mainz, Jakob-Welder-Weg 9, 55128 Mainz, Germany
[b] airconomy aviation intelligence GmbH, Amadeus IT Group, Hugo-Eckener-Ring FAC 1, 60549 Frankfurt, Germany

Keywords:
Airline connection
Itinerary builder
Connection lag
Hub connectivity
Transfer passengers

A B S T R A C T

This paper introduces a methodology for designing an airline connection builder (CB) and adjusting its parameter settings. The objective of the proposed CB is to construct relevant connections that attract passenger demand while avoiding operationally infeasible and commercially irrelevant connections. Using worldwide MIDT booking data, we examined the sensitivity of CB results to the setting of the standard CB parameters maximum connection time and geographical detour. We demonstrated that CB performance can be increased by replacing these two parameters with connection lag, a measure that combines the impact of connection time with geographical detour on the total travel time of a given connection. We also found that the willingness of a passenger to book slower connections and the relationship between the number of passenger bookings and generated connections strongly depends on the O&D distance. Better results can thus be obtained by greedily determining the distance-specific CB parameter settings. The greedy adjustment of parameter settings reduces the number of unattractive connections generated, while keeping the number of covered passenger bookings high.

1. Introduction

A connection builder (CB) is an algorithm that constructs flight connections using a set of rules and parameters. Although most studies on airline networks and connectivity make use of the results of a CB, only few pay attention to its design and parameter settings. In principle, two flights can always be combined to form a connection if the second flight departs after the arrival of the first. However, a CB which applies only this concept would construct many irrelevant connections. Thus, the CB rules and parameters must be defined in such a way that only feasible connections likely to attract passenger demand are formed. In general, a well-designed CB build for schedule and network planning purposes compromises between maximizing the number of competitive connections covering most passenger flows and minimizing the number of unattractive connections.

Many CBs tend to construct a relative large number of connections since a very tight and constrained parameter setting would exclude relevant transfer traffic flows. However, the generation of a large number of irrelevant connections may bias the results of various analyses supporting schedule and network planning, that are based on the CBs output. This is particularly relevant for practical applications using simple and transparent performance indicators calculated directly from the CBs output like e.g. the number of connections or city-pairs served at a given hub. For example, a CB that allowed up to 24 h of connection time in its schedule would obtain a similar number of connections as in a randomly generated schedule, thus failing to underline the performance difference between the two. Of course, the quality of the constructed connections could be evaluated by, for example, applying some kind of scoring index (see e.g. Burghouwt and Redondi (2013) for an overview of several approaches), or, by combining a discrete choice model with traffic estimations (see e.g. Garrow (2010); Coldren and Koppelman (2005)). However, the formulation of such indices or models as well as the estimation of their parameters is influenced by the settings of the applied CB. Consequently, appropriately setting the CB parameters may improve the results of applications using the given CB and thus lead to an overall improvement of the schedule and network planning process.

This paper presents a CB with a greedy parameter selection, and maximum connection lag as the main parameter. Connection lag measures the increase of the total connection's travel time due to geographical detour and connection time at a given transfer hub. Using the MIDT[1] booking data for the worldwide traffic, we analyze

* Corresponding author. airconomy aviation intelligence GmbH, Amadeus IT Group, Hugo-Eckener-Ring FAC 1, 60549 Frankfurt, Germany. Tel.: +49 69 690 27070.

E-mail addresses: adam.seredynski@airconomy.com, adam.seredynski@gmail.com (A. Seredyński), rothlauf@uni-mainz.de (F. Rothlauf), tobias.grosche@airconomy.com (T. Grosche).

[1] Marketing Information Data Tapes: Booking data collected by the computer reservation systems.

0969-6997/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jairtraman.2014.01.005

the sensitivity of the CB results to the settings of maximum connection time, geographical detour, and connection lag. Our analysis reveals how the CB parameters should be set to cover a given percentage of passenger bookings per distance class, and how the parameter settings impact the number of connections generated and passenger bookings covered. We discuss how to fine-tune the parameter settings to improve the CB performance and propose a greedy approach to adjust parameter values for each distance class in such a way that the number of constructed connections is minimized while covering a pre-defined percentage of the passenger bookings. By using MIDT passenger bookings as the calibration data set, we provide the most effective CB parameter setting for worldwide air traffic. Our experimental results show that the proposed greedy maximum connection lag CB leads to a smaller number of irrelevant connections in comparison to a standard CB that uses maximum connection time and maximum geographical detour.

In the next section we provide a brief literature review, present the basic CB concepts typically used in existing studies and introduce the concept of maximum connection lag. Then, in Section 3, we describe the data and methodology used in the analysis. Next, in Section 4, we study how the CB parameters: maximum connection time, geographical detour, and connection lag depend on the origin-destination (O&D) distance. Further, in Section 5, we analyze the sensitivity of CB results and compare the performance of a CB using maximum connection lag with a CB using a combination of maximum connection time and maximum geographical detour. Finally, in Section 6, we present the greedy parameter selection algorithm, which adjusts the parameter settings depending on the O&D distance. The paper ends with conclusions and a brief discussion of this study's limitations.

2. Basic concepts and parameters of CB

The most advanced and comprehensive CB to date were developed for airlines and computer reservation systems. The detailed design and parameters of such CBs are not shared in the scientific literature due to proprietary reasons. Some insights into general concepts of professional CB can be found in e.g. (Goedeking, 2010; Burghouwt, 2007). Other studies (e.g. on the estimation of passenger volumes (Coldren et al., 2003; Garrow, 2010)) use professional CB, but publish no details concerning the CB parameterization.

Many authors developed various forms of CB that were specifically designed to answer a particular research question. Various studies (Doganis and Dennis, 1989; Veldhuis, 1997; Burghouwt and de Wit, 2005; Danesi, 2006; Park et al., 2010) analyzed the hub connectivity of airports using different measures to evaluate the quality of flight connections. Most of these studies focused on connections with at most one stop and used simple connection rules based on connection times and geographical detours. Grosche (2009) developed a methodology for integrated airline scheduling and designed a CB that supported passenger estimations. He viewed the design of a CB as an optimization problem and calibrated the CB parameters using MIDT data. Malighetti et al. (2008) analyzed the connectivity of various networks and explored the potential of self-hubbing. They focused on minimum travel times between all airport-pairs and applied less restrictive connecting rules than the other studies mentioned above.

Although the CB implementations found in the literature differ significantly, they share common concepts. The following paragraphs briefly review the main shared concepts and restrictions, and discuss their standard parameter settings.

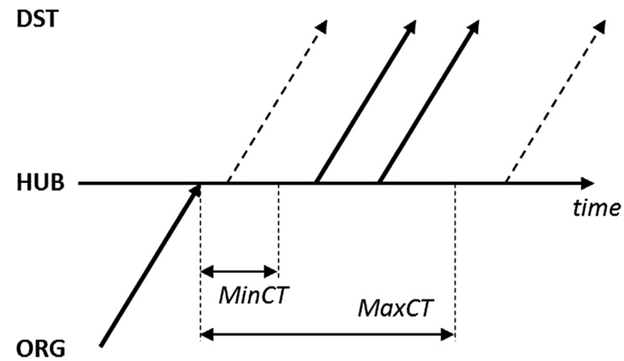

Fig. 1. Minimum and maximum connection time requirements: only the second and third outbound flights are feasible connections.

2.1. Number of stops

A connection may involve several stops. However, with increasing numbers of stop, the attractiveness of a connection decreases. Coldren and Koppelman (2005) analyzed passenger choice behaviour using booking data for the US market and found that 3-segment connections attract significantly less passengers than 2-segment connections. Although the multi-stop flight alternatives often provide the fastest (or the only) travel possibility on a significant number of airport pairs (Malighetti et al., 2008), the vast majority of transfer passengers prefer connections with only one stop. Our analysis of MIDT booking data for the worldwide traffic and DB1B[2] data for the US domestic market finds that multi-stop connections attract only about 2–3% of the total passenger demand. As a result, given the complexity of constructing and dealing with the resulting large number of 3-segment connections as well as their limited significance, most studies focus only on two-segment connection (Doganis and Dennis, 1989; Veldhuis, 1997; Burghouwt and de Wit, 2005; Danesi, 2006; Park et al., 2010).

2.2. Type of connection

Usually, airlines and computer reservation systems support only online and interline connections supported by some form of interline agreement, typically code-share. Passengers traveling on such connections are provided with airline support (ticket, baggage, etc.) in case one of the connecting flights is delayed or canceled. Usually, interline connections that are not supported by any airline agreement can only be booked as individual flights at the passengers own risk. Most of the literature focuses on online and/or code-share connectivity. Furthermore, some CB implementations simply assume that bilateral agreements (Grosche, 2009) or alliance-wide code-share agreements exist (Malighetti et al., 2008; Redondi et al., 2011; Suau-Sanchez and Burghouwt, 2012).

2.3. Minimum connection time

To build a meaningful connection, the departing flight must start within a connecting window, which is defined by the time period between the minimum and maximum connection time (see Fig. 1). The minimum connection time (*MinCT*) is the time necessary to transfer passengers and their baggage from the arriving to the departing flight. The value of *MinCT* is set by the airport and usually depends on the type of connection (domestic, domestic-to-international, or international) including exceptions for specific

[2] 10% sample of airline tickets collected by the US Bureau of Transportation Statistics.