Enhancement of a model for Large-scale Airline Network Planning Problems

Kölker, K.; Lopes dos Santos, Bruno; Lütjens, K.

Publication date
2016

Document Version
Accepted author manuscript

Published in
20th ATRS World Conference

2016 ATRS World Conference, Rhodes, Greece, June 23 – 26, 2016

Enhancement of a model for Large-scale Airline Network Planning Problems

Katrin Kölker[1]
Institute of Air Transportation Systems, German Aerospace Center (DLR), Hamburg, Germany, Email: katrin.koelker@dlr.de

Bruno Santos
Air Transport and Operations, Faculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 1, 2629 HS Delft, The Netherlands

Klaus Lütjens
Institute of Air Transportation Systems, German Aerospace Center (DLR), Hamburg, Germany

The main focus of this study is to solve the network planning problem based on passenger decision criteria including the preferred departure time and travel time for a real-sized airline network. For this purpose, a model of the integrated network planning problem is formulated including scheduling and aircraft rotation. The input for the model is mainly the passenger demand, passenger preferences and the setup of the fleet of one airline. Hence, the final output includes the frequency for different flight legs and a passenger flow. The optimization is conducted using a genetic algorithm combined with a local search algorithm. A use case for a medium sized fleet is presented to demonstrate the functionality of the algorithm.

Keywords: Airline Network Planning, Genetic Algorithm

1 Introduction

Network planning and scheduling is a complex but well analysed field of research. Current models propose an integrated approach to solve multiple airline planning processes simultaneously and globally. Such an approach has been previously published by the authors, in which the general model setup and a solver based on genetic algorithms have been presented. In this paper we present an enhancement of the model including a more general approach and a new optimization method.

The main goal of the model is to generate the network of an airline with the information of where to fly, when to fly and with which aircraft to fly based on passenger preferences. However, the schedule should generate a maximum profit, which includes the operating costs

[1] Corresponding author

but also the revenue generated by sold tickets. In fact, the schedule is mainly driven by the passenger demand for travel between different airports. Passengers have various preference statements and like to fly at specific time, with a specific route and they have even more detailed preferences during the flights including entertainment and food supply. Not all of these preferences have an impact on the level of strategical planning, where network planning is based. Nevertheless, the model presented in this paper includes scheduling of aircraft and, therefore, time of departure is an important factor influencing the number of passengers that are willing to fly with a given flight. Passengers tend to be willing to fly in the morning and evening, which leads to the fact that airlines strive to operate flights during these hours. Additionally, the flexibility of passengers is of concern because some passengers might be willing to deviate from their desired departure time while some would want to change the airline rather than changing the departure time. A model of network planning needs to take this aspect into account. Finally, passengers are sensible to travel times which include potential transfers. This means that passengers will often prefer direct flights over connecting flights and shorter itineraries over long itineraries. This passenger behaviour has to be considered also when setting up a schedule.

Differentiating passenger preferences in a more detailed level becomes more important with the size of the network. In smaller airline networks there are not many optional itineraries for the passenger such that the effect of changed passenger flows is irrelevant. Hence, larger network have to be calculated to consider passenger flow effects appropriately. To do so, an algorithm has been developed to solve larger instances. The new approach is a hybrid algorithm consisting of a genetic algorithm and a fast local search algorithm.

The focus in this paper is not to optimize the network but rather to build a network for the airline which is close to the reality. Nevertheless, an airline always seeks for maximizing the profit, which means that the model must necessarily reproduce the behavior of optimizing the network. This study is intended to demonstrate the general setup of a model and to solve the network planning problem with various passenger preferences influencing the network.

Additionally, a central requirement for the model is that the framework should be able to evaluate aircraft characteristics. This means that the framework and especially the network planning model should be able to define requirements on aircraft and the impact of different aircraft characteristics on the decision of operated routes, the frequency on these routes and operated fleet types. These characteristics can be, for example, the speed, its capacity, its fuel efficiency and other aircraft performance attributes. The presented approach allows for calculating parameter studies considering the mentioned features.

The paper models and illustrates the two different side of the market. On the one hand, the generation of the schedule is generated representing the supply side. The airline generates a network based on the given parameters and passenger flow. On the other hand, the demand side is required to model the passenger behaviour when faced with a fixed schedule. These two sides allow for modelling a complete airline network. An example for an airline is provided demonstrating the different optimization techniques.

2 Mathematical Model for Network Planning

The airline planning process consists of several sub problems that define different aspects of the operation planning and become more detailed in the course of the planning period. First, a general demand forecast is conducted and used as a basis and input for the following optimization approaches. The first task is to fix the topological network being built up by the airport and the possible connections. The topological structures can be used to generate the flight network. To do so, the frequencies of the flights on all segments need to be fixed and the scheduled departures of these flights are retrieved afterwards. Hereafter, fleet types are allocated to the scheduled flights during fleet assignment resulting in several flight networks for different fleet types. Next, tail numbers are assigned to the scheduled flights which is called aircraft rotation planning and usually maintenance constraints are considered. Later steps consider crew planning and more detailed steps closer to the day of departure.

This process is driven by the demand forecast. Passenger demand is usually given as origin-destination demand, which means that the origin and final destination of every passenger willing to fly is known. However, this passenger can travel directly or connect at one or more airports. The detailed passenger route is called passenger itinerary and comprises all segments of the passenger.

2.1 Literature Review

A large number of models have been presented describing single sub problems of the planning process or integrated approaches considering two and more sub problems. However, earlier models do not consider the whole planning process but concentrate on solving single problem to optimality. Lederer and Nambimadom (1998) retrieve the optimal frequency for flights based on aircraft rotations in a network to compare direct and hub-and-spoke networks. Adler and Berechman (2001) focus on the topological spread of the network but develop a model to set aircraft size and frequency based on the origin-destination demand and to maximize profit. The passenger decision is modelled by a utility function as a function of the frequency, airfare and number of connections. Barnhart et al. (2000) consider the fleet assignment problem applying a passenger flow model to integrate demand effects into this decision step. They consider the impact of spill and recapturing effects when assigning passengers to a fixed schedule.

Other, more recent studies focus on the integration of scheduling with other sub problems. Scheduling models can usually be differentiated in two groups. On the one hand, models take a given schedule as an input and reschedule flights within a given time window. These models are used to perform slight changes in the airline operations, which may occur when the circumstances are subject to changes close to the operation of the schedule. For example, Mercier and Soumis (2007) use an aircraft routing model and allow for shifting flights slightly. Rexing et al. (2000) enhance a fleet assignment model by integrating additional flight legs being copies of given legs and, therefore, having the possibility to shift flights in a given time window.

On the other hand, models are generated to build up schedules from the scratch. These are used as first inputs for schedule planning, which is usually a long time before the actual operations. Lohatepanont and Barnhart (2004) develop a complex model considering

frequencies, scheduling from the scratch and fleet assignment combined with a passenger mixed model in a second step. Demand is spilled and recaptured by other passenger itineraries on same origin-destination market, where the recapture rate is based on a quality of service index. The schedules are on a daily basis. Braun et al. (2010) present an integrated algorithm including frequency planning, scheduling and fleet assignment. Passenger demand is given as accumulated origin-destination demand and passenger are allocated to itineraries without considering additional restrictions. Erdmann et al. (2001) present an integrated approach for frequency planning, scheduling and fleet assignment based on aircraft rotations for charter airlines. Passenger demand is given on a daily origin-destination basis and spilled if it exceeds capacities but not recaptured. Sherali et al. (2010) formulates a model based on passenger itineraries including network planning, scheduling and fleet assignment as a mixed-integer program and use Benders' decomposition for solving the problem. They differentiate between mandatory and optional legs which allows for setting flights in advance and adding profitable flight when solving the problem. Recapture was integrated in a follow-up study (Sherali et al., 2013). Other integrated models building integrated schedules from the scratch are for example for freight networks and presented by Tang et al. (2008) and Yan et al. (2006).

Passenger demand can be formulated and integrated in multiple ways. Passenger flow can be formulated as spill and recapture approach where passenger are spilled from one leg if capacities is exceeded and recaptured by another leg operated by the same airline. Jacobs et al. (2008) integrate this approach into fleet assignment by solving the leg-based fleet assignment problem and the passenger flow problem iteratively. Cadarso and Marín (2013) develop a model to solve scheduling and fleet assignment at the same time and consider passenger itineraries. Demand is given on an origin-destination basis and depends on the departure time. However, every passenger has a set of feasible departure times. Therefore, itineraries in the same market at different departure times compete for the same passengers. Spill and recapture does happen for itineraries with different but a priori fixed departure times at no extra cost for the airline. Pita et al. (2013) present a model for network planning where flights are restricted to slots. Spill and recapture is modelled and the departure time for passengers is considered while passengers can be recaptured by itineraries in other time windows. Another possibility is to evaluate the passenger flow after the schedule has been set. A model setting up the schedule a priori and then evaluating the passenger flow based on the fixed or slightly changed schedule is from Derigs et al. (2009) and applied to freight networks. In order to optimize passenger routes and airlines' frequencies, Hsu and Wen (2000) use grey theory and multi-objective programming. They apply grey theory to model and handle the uncertainty of demand. Hsu and Wen (2003) extended this by the interaction of frequency setting of the airline and passenger behaviour and optimizing both until equilibrium is found.

Algorithms solving airlines' planning problems using metaheuristics have also been discussed. Büdenbender et al. (2000) propose an approach to optimize the frequency of smaller freight networks. They apply a two-phase algorithm where feasibility is constructed in the first phase and the solution improved in second phase. The second phase utilizes a hybrid algorithm consisting of a tabu-search algorithm and a branch-and-bound algorithm. Grosche et al. (2001, 2008) propose a model to solve the scheduling and fleet assignment problem with

genetic algorithms, where a solution is represented by a string of flights and ground times. Passengers are spilled and recaptured based on a customized function that uses the preferred departure time of a passenger. Mashford and Marksjö (2001) use simulated annealing to solve the scheduling problem and define the final rotations of the aircraft. The demand is distributed on time windows and assigned to flights in the schedule evaluation step.

2.2 Model Formulation

In the following the mathematical model formulation consisting of the flight network and the passenger flow network is presented. However, as the algorithm is based on genetic algorithms, it solves the problem in two steps. A solution is given for the flight network meaning that a rotation for every aircraft is fixed. Secondly, the market model, which is described later on, distributes the passengers onto flights regarding the constraints presented in the formal description of the model.

The model is based on a set of airports and candidate flights in between. A set of aircraft is given to operate these flights and representing the entire fleet of the airline. The fleet must not necessarily consist of one fleet type. A candidate flight is a scheduled flight between two airports with a fixed aircraft assigned to it. This means, that the capacity, the flight time and the direct operating costs of the flight can be defined directly because the aircraft type is known. Additionally, the exact departure time is known. Therefore, not every candidate flights needs to be operated but the model choses the most profitable flights. However, the required turnaround time after the flight is integrated in the flight time. This means, that it does not have to be considered in the model directly and that the flight time is the amount of time the aircraft is not available for another flight independent from the current location of the aircraft.

For the market model, a set of origin-destination markets is defined, which connect two airports. For every market, a set of scheduled itineraries is defined. A scheduled itinerary is a set of scheduled flights having the same origin and destination as the market. For every market the overall demand of the whole considered time period is given. To model passenger preferences, it is broken down onto different time windows. For example, a weekly demand is given and broken down to slots of one hour such that the sum of the demand of all hours is equal to the demand of the whole week. Two additional factors describe the willingness of passengers to be allocated to other time windows than their desired time of departure. First, a factor is defined describing the share of passengers that is actually assigned to a specific time window but in general willing to depart at another. Secondly, we define a share of passengers that is also assigned to a specific time window and willing to be assigned to an itinerary in another time window. The latter might be smaller, because some passengers that want for example to depart between 9 a.m. and 10 a.m. might also be willing to depart between 10 a.m. and 11 a.m. However, there might be two itineraries with different travel times in that time slot. It is assumed that the share of passengers that is willing to take the longer itinerary is smaller than the share of those passengers willing to take the shorter itinerary.

Sets and Parameters

Flight network, Supply side

A: set of all airports
K: set of all aircraft
F: set of all scheduled candidate flights
Rot_k: set of all feasible rotations for aircraft $k \in K$, where $\forall r \in Rot_k$: r $\subset$ F
$cap(k)$: capacity of aircraft $k \in K$
$cost(f, k)$: direct operating costs of flight $f \in F$ operated by aircraft $k \in K$

Demand side

M: set of all markets
T: set of all time windows
$dem(m)$: total demand on market $m \in M$
I_m: set of all scheduled passenger itineraries on market $m \in M$
I: set of all scheduled passenger itineraries, $I = \bigcup_{m \in M} I_m$
$price(i)$: ticket price of itinerary $i \in I$
$p_t(m, t)$: proportion of demand from market $m \in M$ that is willing to travel at time $t \in T$, where $\sum_{t \in T} p_t(m, t) = 1$
$p_{ti}(m, t, i)$: share of passengers of the of demand from market $m \in M$ willing to travel at time $t \in T$ that is interested in travelling on itinerary i

Auxiliary functions

$$\delta_{f,r_k} = \begin{cases} 1, & \text{if flight } f \in F \text{ is in rotation } r_k \\ 0, & \text{else} \end{cases}$$

$$\zeta_{f,i} = \begin{cases} 1, & \text{if flight } f \in F \text{ is in itinerary } i \in I_m \\ 0, & \text{else} \end{cases}$$

$$\eta_{i,t} = \begin{cases} 1, & \text{if itinerary } i \in I_m \text{ starts at time } t \in T \\ 0, & \text{else} \end{cases}$$

Decision variables

$$x_{r_k} = \begin{cases} 1, & \text{if rotation } r_k \in Rot_k \text{ is operated by aircraft } k \in K \\ 0, & \text{else} \end{cases}$$

pax_i: number of passenger travelling on itinerary $i \in I$

$alloc(m, t, \tilde{t})$: number of passenger willing to travel at $t \in T$ and $m \in M$ and actually travelling at $\tilde{t} \in T$

$$max \sum_{i \in I} rev(i) pax_i - \sum_{k \in K} \sum_{r_k \in R_k} \sum_{f \in F} \delta_{f,r_k} cost(f,k) x_{r_k}$$

$$\text{s.t.} \quad \sum_{r_k \in R_k} x_{r_k} \leq 1 \qquad \forall\, k \in K \qquad (1)$$

$$\sum_{i \in I} \zeta_{f,i} pax_i \leq \sum_{k \in K} \sum_{r_k \in R_k} \delta_{f,r_k} cap(k) x_{r_k} \qquad \forall\, f \in F \qquad (2)$$

$$pax_i \leq \sum_{t \in T} p_{ti}(m,t,i) p_t(m,t) demand(m) \qquad \forall\, i \in I_m, \forall m \in M \qquad (3)$$

$$\sum_{i \in I_m} \eta_{i,t} pax_i \leq \sum_{\tilde{t} \in T} alloc(m,\tilde{t},t) \qquad \forall\, t \in T \qquad (4)$$

$$\sum_{\tilde{t} \in T} alloc(m,t,\tilde{t}) \leq p_t(m,t) demand(m) \qquad \forall\, t \in T \qquad (5)$$

$$x_{r_k} \in \{0,1\}, pax_i \in \mathbb{Z}_{\geq 0}, alloc(m,t,\tilde{t}) \in \mathbb{Z}_{\geq 0} \qquad \forall\, r_k \in Rot_k, \forall\, k \in K, \forall\, t, \tilde{t} \in T \qquad (6)$$

The objective function represents the profit of the airline. Therefore, the first part is the revenue gained by the tickets sold to passengers on itineraries. The second part reflects the operating costs of all operated flights. Constraint (1) ensures that every aircraft operates only one feasible rotation. The second constraint (2) forces the number of passengers on every flight to be less the capacity. Here, the number of passengers is calculated by adding all passenger itineraries that contain the specific scheduled flight.

Constraints (4) to (6) are monitoring the passenger flow based on the given flight network. The first constraint (3) handles the overall number of passengers on one itinerary. It must not exceed the sum of the passenger from all time windows that are willing to be on the specific itinerary. The next two constraints, (4) and (5) ensure that the allocation of passengers to other time windows is limited. Constraint (4) guarantees that the number of passenger allocated to all itineraries in one time window don't exceed the demand for this time window. Whereas the next constraint limits the passenger that are allocated from one specific time window to all other time windows. This number is not allowed to exceed the demand for the specific time window.

Constraint (4) and (5) guarantee that the sum of the passenger of the itineraries on one market does not exceed the overall demand on this market i.e. that the following equation holds

$$\sum_{i \in I_m} pax_i \leq demand(m)$$

Proof.

$$\sum_{i\in I_m} pax_i = \sum_{i\in I_m}\sum_{t\in T}\eta_{i,t}pax_i = \sum_{t\in T}\sum_{i\in I_m}\eta_{i,t}pax_i \leq \sum_{t\in T}\sum_{\tilde{t}\in T} alloc(m,\tilde{t},t)$$
$$= \sum_{\tilde{t}\in T}\sum_{t\in T} alloc(m,\tilde{t},t) \leq \sum_{\tilde{t}\in T} p_t(m,\tilde{t})demand(m) = demand(m)$$

3 Market Modell

3.1 General approach

The main idea of the market model is to determine the passenger flow for a given airline schedule based on specific passenger preferences. Here, it is assumed that the passenger has already decided to fly with the analysed airline in order to neglect considerations on market share for the moment. However, the passenger can decide not to fly with this airline if the schedule does not suffice his needs or if there is no capacitiy on the required flights left. This means, that this passenger is spilled and not recaptured by the very same airline. In general, the passenger flow has to be calculated based on a given and fixed schedule implying possible passenger itineraries. Therefore, the topological structure of the passenger flow is given by the schedule and arising itineraries.

The main task of the market model is to define the exact passenger flow that is the scheduled itinerary for every passenger. The input is given as origin-destination demand meaning that only the origin and desired destination of the passenger is known. Additionally, passenger preferences are given to define passenger flow. For the purpose of this model, the market model is used to solve the constraints (3) to (5) of the model formulation. The model formulation is based on time windows. In this model the time windows comprise one hour in the week. Every passenger has an assigned time window which is the preferred departure time of the passenger. Specifically, for every passenger it is known in which hour of the week he wants to depart. Nonetheless, a certain proportion of spill is considered as well. For example, the total amount of all passengers that want to travel from London to Paris in one week is 5000. Thereof, 100 passengers favour to depart between 9 a.m. and 10 a.m. on Wednesday. However, 60 % of them would also be willing to depart between 10 a.m. and 11 a.m. and 20 % are also willing to start after 11am. However, the preferred departure time for all of them is still 9 a.m. to 10 a.m. Only their departure time is allocated to another time window.

The general concept of spill and recapture requires a demand of passengers given based on itineraries (Barnhart et al. (2002), Jacobs et al. (2008)). In the presented model, the demand is given on the level of market and split onto different time windows. This means that there is no traditional spill and recapture concept. However, passengers are distributed between different itineraries. Even passengers that originally prefer another time window are allocated to an itinerary in another time window. In this sense, demand is spilled and recaptured between time windows.

3.2 Passenger Preferences

Two main passenger preferences are revealed in the market model. On the one hand, the preference considering the desired time of departure is determined and applied in the model. This value corresponds to the variable $p_t(m,t)$ in the model formulation. To retrieve the

underlying preference curves, an approach similar to the procedure of Lehner et al. (2014) was applied. For this purpose, real passenger booking data and airlines' schedules have been merged to generate passenger distribution for one week. A backwards engineering process has been applied thereafter to simulate the spreading of the passengers to other time windows reversely. The detailed procedure will be presented in another study and is not important for this study as the distribution function operates as input values solely. The distribution represents the desired departure time of the passenger for the origin airport. Therefore, a curve for every airport expressing the weekly passenger share has been generated. Figure 1 illustrates the curve for different airports.

On the other hand, the share of passengers willing to deviate from their desired time windows needs to be fixed i.e. the value $p_{ti}(m, t, i)$ in the market formulation. To do so, an adapted curve of the function provided by Koppelman et al. (2008) is used (see Figure 2). The function indicates the share of passengers that is not willing to deviate from their preferred departure time by a specific amount. However, this function also acts as comparison between itineraries with different travel time on the same market. It states the willingness of passengers to take an itinerary with a given travel time compared to the best available travel time. If there exits multiple itineraries differing by a certain amount of minutes to the best available itinerary, the passenger share of every non-optimal itinerary is reduced by the share given by the function.

For example, one market has two assigned itineraries. The first departs Mondays at 9:20 a.m. and requires two hours and the second Tuesdays at 10:30 a.m. taking three hours because of a transfer. First, the weekly distribution is applied resulting in the fact that the time window of the first itinerary has a passenger share of 0.4 % whereas the time window of the second one has a share of 0.2 %. Second, this numbers have to be reduced by the share of passengers rejecting the itinerary because of a longer travel time. Here, the second itinerary is 60 minutes longer than the best available meaning that the passenger share of is reduced by 8 % from 0.2 % to 0.184 %.

Finally, passengers that prefer other time windows can be allocated to these itineraries as well. For example, the time window of Monday 1 p.m. - 2 p.m. has a passenger share of 0.1 %. According to the block time distribution, only 55 % of these are willing to deviate by three time windows to the first itinerary. Therefore, the share is decreased to 0.055 %. The final share of this time window is added on top resulting in a share of 0.455 %. However, none of these passengers want to travel on Tuesdays as it is more than six hours from their original time window. This procedure is repeated for every other time window to gain the final passenger share.

In the meantime, when passenger share is assigned to the longer time window, it has to be reduced by the amount of passenger rejecting because of longer itineraries as well. For example, the time window of Tuesday between 12 a.m. and 1 p.m. has a passenger share of 0.15 %. Of these, 30 % are not willing to deviate by two hour to the itinerary. Of the remaining 0.105 % another 8 % are not willing to take the longer itinerary as there is an itinerary available being 60 minutes shorter. Consequently, 0.0966 % passengers are allocated to the itinerary on Tuesday as well.

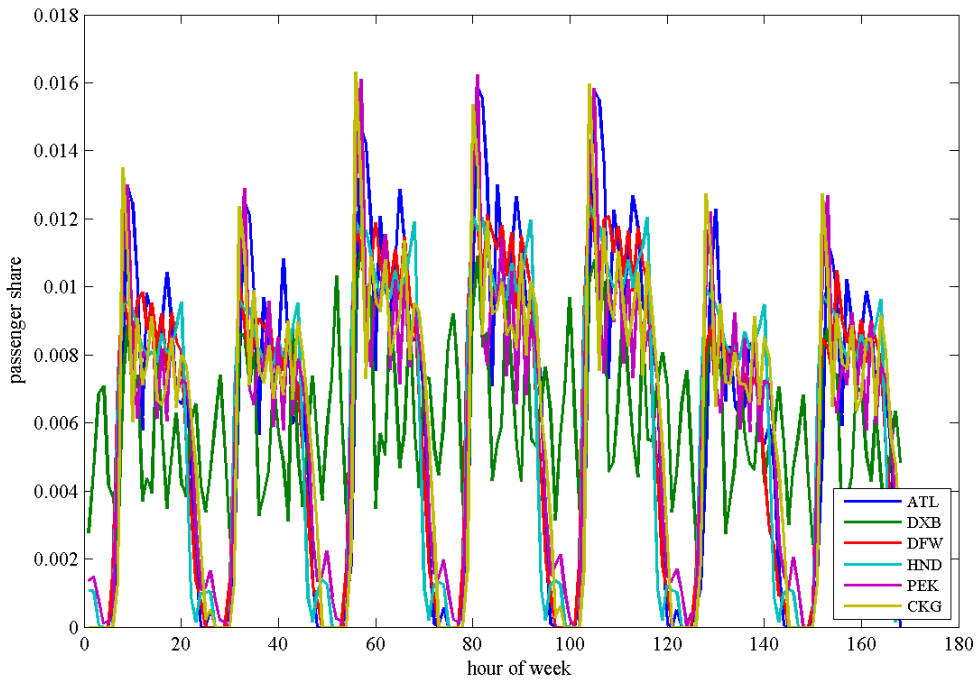

Figure 1 Weekly passenger share per hour representing the preferred departure hour of passengers for the given originating airport

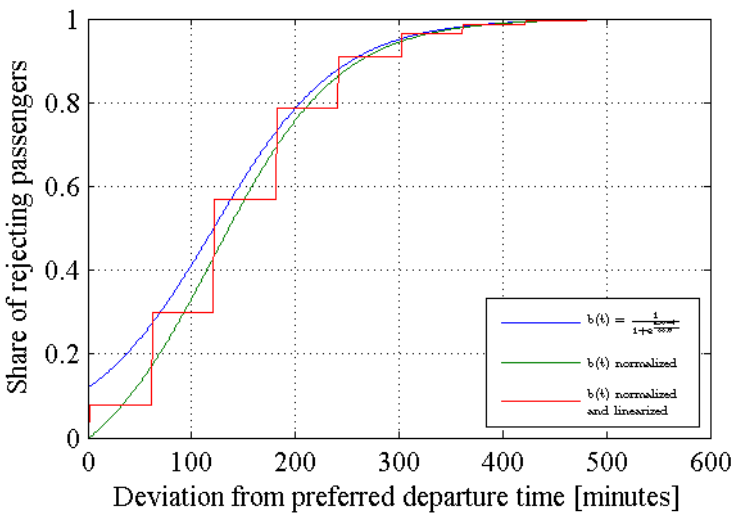

Figure 2 Block time distribution; adapted from Koppelman (2008)

3.3 Implementation

The implementation of the market model is embedded in the airline network model. Accordingly, possible passenger itineraries and flights are given as well as the market demand and curves of the market model. The procedure described in the previous section is applied for every itinerary and every time window. However, it is necessary to keep track of the overall passenger share of a time window i.e. to make sure that not more passengers are allocated to other time windows than available in this time window (constraint (5)) and to make sure that the number of passenger on all itineraries in one time window does not

exceeds the available demand of the window (constraint (4)). Formally, the procedure looks as follows.

1. $\forall\, m \in M$
   $\forall\, i \in I_m$
   $pax_i := \sum_{t \in T} p_t(m,t,i) p_{ti}(m,t) demand(m)$
   $\forall\, t \in T, \forall\, \tilde{t} \in T$
   $alloc(m,t,\tilde{t}) := \sum_{i \in I_m} \eta_{i,\tilde{t}} p_t(m,t,i) p_{ti}(m,t) demand(m)$
2. $\forall\, m \in M$
   $\forall\, t \in T$
   IF $\left(\sum_{\tilde{t} \in T} alloc(m,t,\tilde{t}) > p_t(m,t) demand(m)\right)$
   THEN
   $alloc(m,t,\tilde{t}) := \frac{p_t(m,t) demand(m)}{\sum_{\tilde{t} \in T} alloc(m,t,\tilde{t})} alloc(m,t,\tilde{t})$
3. $\forall\, m \in M$
   $\forall\, t \in T$
   IF $\left(\sum_{i \in I_m} \eta_{i,t} pax_i > \sum_{\tilde{t} \in T} alloc(m,\tilde{t},t)\right)$
   THEN
   $pax_i := \frac{\sum_{\tilde{t} \in T} alloc(m,\tilde{t},t)}{\sum_{i \in I_m} \eta_{i,t} pax_i} pax_i$

The three steps above define the share of passengers that want to travel at time t preferably but do travel at $\tilde{t}$. It is possible that both time windows are the same meaning that the passengers travel at their preferred time window. The first step sets the number of passengers per itineraries such that it is equal to the amount of passengers from all time windows that are willing to travel with this itinerary which is represented by constraint (3). The resulting variable $alloc(m,t,\tilde{t})$ defines the number of passengers that want to travel time t preferably but do travel at $\tilde{t}$ and is calculated by summing up the share of passengers from travel window t that travel on all itineraries departing in $\tilde{t}$. The second step guarantees that not more passengers from a time window are allocated to other time windows and, therefore, ensures that constraint (5) is met. The variable $alloc(m,t,\tilde{t})$ is reduced by the critical amount. The third step ensures the compliance of constraint (4), which is the requirement that the sum of all passengers on the itineraries in one specific time window does not exceed the number of passengers that are allocated to this time window. The number of passengers on the itineraries is reduced by the difference if necessary

With this approach it is possible to assign passengers to flights of a given flight schedule according to their timely preferences

4 Optimization Approach

The general approach to solve the network planning model with passenger preferences is based on genetic algorithm where one individual is representing the schedule of the airline. One individual or schedule in the context of the genetic algorithm is one possible solution of the problem. The basic approach was presented in Kölker and Lütjens (2015). The enhancement of the model presented in this paper includes an improvement of the

optimization approach and the detailed integration of passenger preferences in the network planning process.

It is combined with a local search algorithm and the market model to determine passenger behaviour. The objective function, which is the fitness of the individual, is equal to the possible profit achievable with this schedule i.e. individual. The profit consists of the operating costs of the flights within the schedule and the generated revenue by ticket prices. To determine the revenue and passenger flow, the market model is applied.

The flow chart of the procedure is visualized in Figure 3. A random generation of individuals i.e. schedules is produced at the beginning and a genetic algorithm consisting of choosing best individuals, combining and mutating is applied. Finally, a local search algorithm is performed on the individuals. To define the objective value of an individual throughout the evolution, the flights are scheduled to slots of ten minutes and possible itineraries. Based on these flights and itineraries, the market model calculates the passenger flow. Real data on schedule and passenger routes is used as input for the market model.

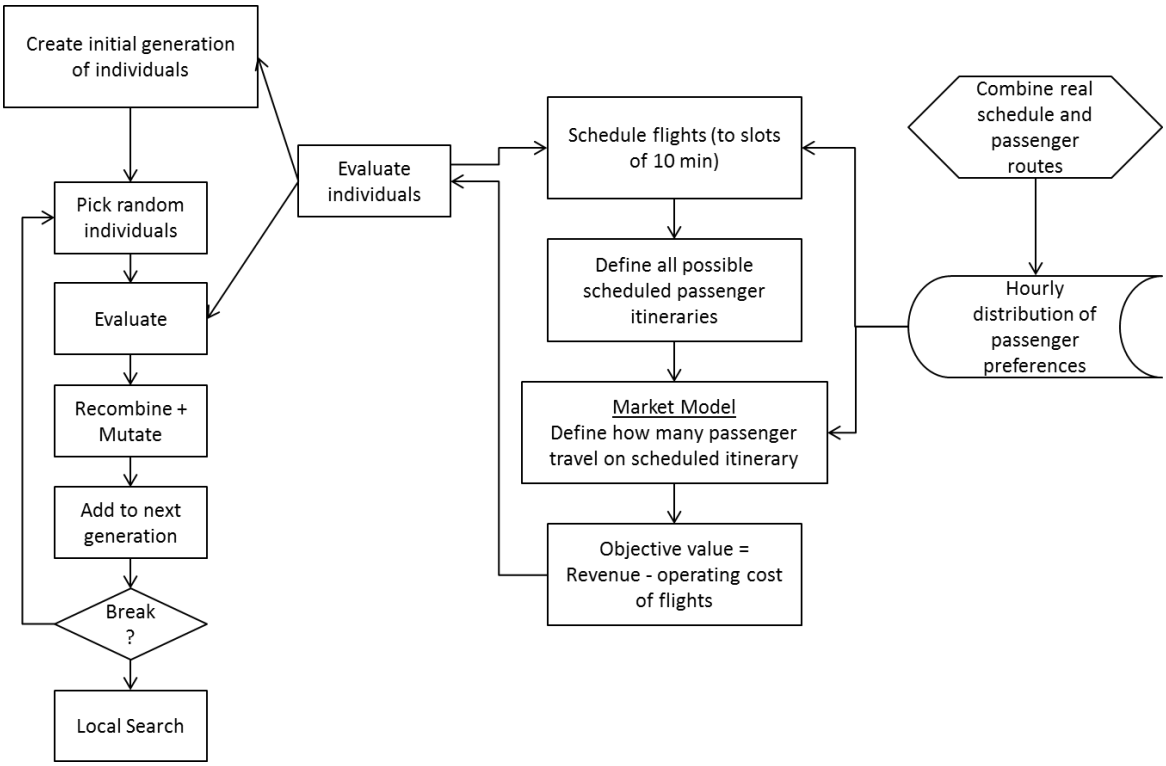

Figure 3 Flowchart of network planning model combined with market model

4.1 Genetic Algorithm

The genetic algorithm consists of several steps that are presented in the following. First of all, the encoding of the individuals is represented. Next, the combining and mutating operator are defined.

4.1.1 Encoding

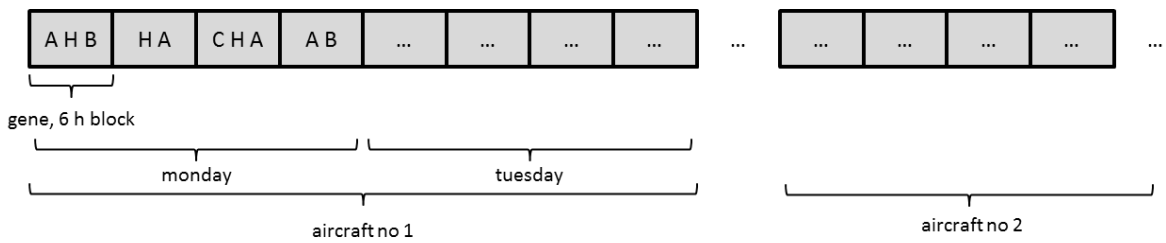

Figure 4 Schemata of one individual / schedule in the genetic algorithm

One individual in the genetic algorithm has to represent a weekly schedule of multiple aircraft. To do so, the individual is formulated as a string, where every part of the string represents the rotation of one aircraft. The rotation is represented by a sequence of airports. Figure 4 illustrates an example of one individual. The first aircraft starts at airport A and flies to airport H followed by airport B and H and so forth. The rotation of the second aircraft is represented by a sequence of airports in the next part of the same string directly after the rotation of the first aircraft. Therefore, the flights of the whole fleet and consequently the whole schedule of the airline can be deduced from this individual represented by sequences of airports.

However, there is no departure time associated to the flights yet. Due to the size of the problem, it is unfeasible to schedule every flight to slots of five or ten minutes. To overcome this problem, the sequences are grouped to blocks. Each block spans a time frame of six hours meaning that four blocks represent one day and 28 blocks are the rotation of the aircraft in one week. One block is one gene in the individual and a block can have different numbers of airports. Grouping airports and, therefore, flights to blocks of six hours, reduces the number of potential genes of the individual which reduces the number of solutions and decreases the solving time of the optimization approach. All flights in one block can be scheduled to any departure time in the block as long as all other flights are operable in this block. A flight can be between blocks, as it is the case in the example. Here, the first flight from B to H starts in the first block but ends in the second block. This means, when scheduling flights within blocks, the following rules have to be applied.

- The aircraft has to operate all given flights and turnaround times. No other flights are allowed.
- The aircraft has to be at all airports completely within the given six hours e.g. aircraft 1 has to start from A after 0 a.m. and arrive at B before 6 a.m.
- Flights in-between different blocks are allowed but start has to be in first block and arrival in second block e.g. aircraft 1 departs at airport B before 6 a.m. and arrives at H after 6 a.m.

4.1.2 Algorithm

The evolutionary process proceeds as follows. First, an initial of 100 individuals is created randomly from a list of feasible blocks and the objective value is calculated for each. In every following generation, 50 pairs of individuals are chosen randomly with the probability given by their fitness i.e. objective value. Every pair is combined to create two offspring, which are mutated afterwards. The mutation is performed for every gene with a probability of 1% meaning that this gene is exchanged with a random neighbouring gene. This procedure is applied to create a new generation of 100 individuals and repeated for 100 generations. The average objective value in the generation improves as time passes.

4.1.3 Infeasibility

Individuals can result in three different types of infeasible schedule. First, a flight can be infeasible if it is within the curfews of the origin and destination airport. Secondly, a flight can be infeasible because there is no candidate flight. This is the case when the last airport of the first block and the first block of the second airport do not match. This may be the case, when for example the airline is operating a hub-and-spoke network and both airports are spokes which shall not be connected with a flight. In both cases, the individual is penalized by decreasing the objective value by $50,000 for each infeasible flight. Third, a flight can be pushed partially into curfews because prior flights are too long. If this is the case, the individual is punished with $100 for every minute the flight is pushed into curfews. However, there exist routines fixing these infeasibilities up to a certain extent that are presented in section 4.2.

4.1.4 Neighbourhood

To apply mutation during the genetic algorithm and for the local search algorithm, it is important to define the neighbourhood of a solution. Two solutions are neighboured if they only differ by one gene and if the different genes are neighboured as well. Genes are neighboured if one of the three following requirements is met. Either both genes have the same number of airports in the same order and differ by exactly one airport; or one gene results from the other by deleting the first or last airport. In this sense genes are neighboured if they differ in one airport and the same applies for individuals. Therefore, neighbouring individuals differ either in size by one or in one airport.

4.1.5 Operator

For recombining two parent individuals, the single-point-crossover operator is used. Here, the offspring are created by taken two parent individuals and choosing a random gene as crossing point. The first offspring consists of the first part of the first parent and the second part of the second parent and the other offspring is build up vice versa. The mutation operator mutates every gene with a given mutation rate to a feasible neighbour.

4.1.6 Objective Function

To retrieve the objective value for one solution, the objective function adds the operating costs and the revenue. Calculating the operating costs is straightforward as the operated flights are known and the costs are based on the costs. The authors are aware of the fact that a schedule causes more costs as for example maintenance based costs which are influenced by

the amount of cycles an aircraft operates. For simplification, these costs are not added here but will be in further studies.

Furthermore, the revenue is generated by the ticket prices. To define the specific passenger flow as described in chapter 3, the flights have to been scheduled to slots of ten minutes first.

The implementation orders all itineraries by the following metric that is supposed to represent a first estimation of the possible profit on the itinerary. To define the value for one itinerary in the metric, the overall demand of the market is multiplied with the share of the time window and the ticket price on the market. Therefore, the capability reflects some sort of expectation of the itinerary. The itineraries are considered according to the order given by their capabilities. This procedure ensures that the itineraries with the higher ticket prices and highest demand amount are considered first.

Next, the itineraries are considered in the given order. If an itinerary is operable because the flights could be operated in the given schedule, the flights are fixed within the block. However, other flights in the block have to remain feasible if one flight is fixed. If one itinerary is scheduled, the metric of all itineraries on this market has to be adapted because the remaining demand on this market is reduced. Therefore, the metric is decreased by the capacity of the scheduled itinerary multiplied with the share of the time window and the ticket price. This procedure is performed for all itineraries. Eventually, all flights will be scheduled and a set of all available scheduled passenger itineraries is created.

Finally, the market model is applied to generate the final and actual passenger flow. This is necessary as last step because scheduling of flights or generating new itineraries can change the final passenger flow, as itineraries on same markets are competing for passengers when the time windows that they are allocated to are close. The revenue and finally the profit and objective value of the individual solution can then be retrieved by adding up all ticket prices of flying passengers.

4.2 Resolving Infeasibility

Different strategies are applied to improve the algorithm and presented in the following. However, three approaches are used to make an individual more feasible and are implemented straightforwardly. First of all, the flights operated within curfews have to be deleted. It is straightforward to check whether a flight starts or lands within curfews. Whenever this is the case, the departure airport is deleted from the specific block. Second, if two blocks are infeasible due to non-existing flights, a feasible replacement for the first block is sought. To do so, all genes in the neighbourhood of this are scanned and the current block is replaced by the best block in the neighbourhood which is feasible in terms of flights and curfews. Third, flights that partially reach into curfews are treated in the same manner as those being in curfews completely. These three fixing routines are applied every ten generations as they consume runtime.

4.3 Evolution of Champions

Another procedure of improving the optimization approach is utilizing the concept of genetic algorithms. As several runs have to be performed anyway, the best solutions of every run can be taken to form a new generation and start another evolution. Here, 100 runs are performed

at the same time. To form an evolution of the champions, the ten best solutions are taken every ten generation from every run and form a new initial generation of 100 individuals. The original approach of genetic algorithm is performed on these generations as well. This leads to ten additional runs that contain the best individuals of every run and improve themselves as time passes. However, the best of this champion runs is probably the last run, as this starts with a generation build up with the ten best solutions of the final generation of all runs.

4.4 Local Search and Improvement of the algorithm

To exploit the solution locally, a local search algorithm has been applied. For this purpose, a hill-climbing algorithm has been applied. It searches all the neighbours available for one solution and moves towards the best neighbour. For this purpose, the neighbourhood around a solution as defined in section 4.1.4 is applied. This approach conducted until no neighbour is available that is better than the current solution. This final solution is the local optimum in a small area of the solution space around the original solution.

To explore multiple regions of the solution space, the local search algorithm is applied to all individuals of the final generation as well as the best solution of all generations. This application analyses the local space of several solutions to increase the search space for the local search algorithm.

5 Results

A use case consisting of an artificial European network based on the network of the A320 of Swiss has been calculated. The use case comprises 46 airports with two hubs that can be connected with 15 aircraft. For every airport the curfews are assumed to be between 10 p.m. and 6 a.m. The passenger preferences are given with the market model as described above and the overall demand is 77,904 passengers for one week. The schedule time is taken from the original schedule whereas the turnaround time is assumed to be 30 minutes. In one week 75,572 candidate flights are possible resulting in 109,872 potential scheduled passenger itineraries. The ticket prices are modelled according to be $0.134 per kilometre based on the great circle distance. For more details see Kölker and Lütjens (2015).

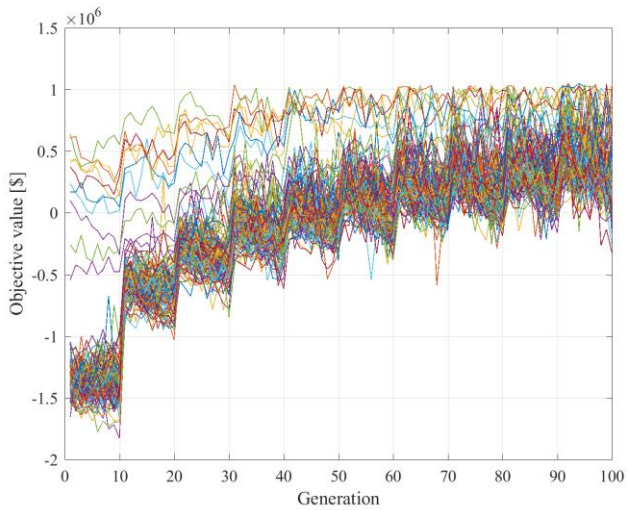

Figure 5 Progression of all 100 runs + 10 champion runs; one line represents one run

The algorithm has been performed 100 times. Figure 5 visualizes the progression of the different runs, where for every generation and every run the fitness of the best solution in the current generation is illustrated. The distinct improvement over time is clearly visible as well as the different steps every ten generations that occur from the application of the fixing routines. The best runs on top of all of the others are the champions' runs.

The final network has 472 flights per week with a capacity of 64,192 seats. The costs to operate this schedule are estimated with $2.51m. The overall passenger flow on this schedule consists of 33,461 passengers spread on 510 itineraries resulting in revenue of $3.57m. This yields an overall profit of $1.06m. The load factor of 52 % seems low but as this airline is using the network of A320 as feeder flights, it is quite close to reality. However, nearly 60 % of the demand is spilled and not transported, which is not reasonable. The algorithm has to be changed to overcome this problem which will be done in future studies. Figure 6 illustrates the influence of the different improvements of the algorithm.

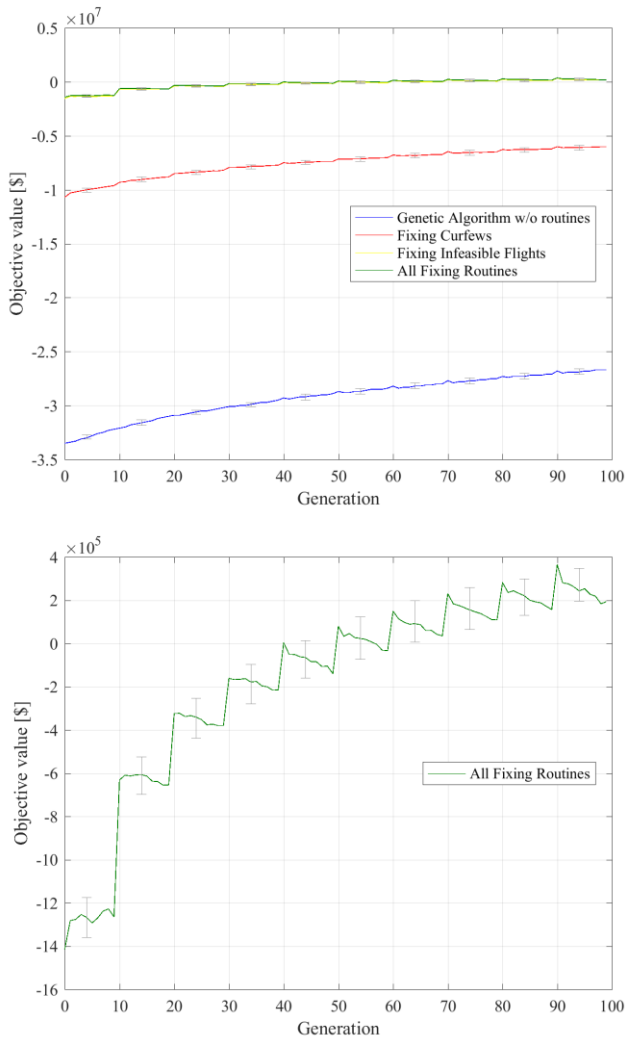

Figure 6 Maximum objective value in generation; median of all runs is plotted, whiskers represent upper and lower quartile

The improvement of the different fixing routine is clearly visible as well as the steps every ten generations. The local algorithm is performed after this approach and increases the value of

the solution by around ten percent meaning that it is quite appropriable and necessary to use this algorithm.

6 Conclusions and Outlook

The paper introduces a network planning model based on genetic algorithm that utilizes blocks of airports to represent aircraft rotation. One individual in the genetic algorithm is equal to a weekly schedule for the whole fleet of one airline. The objective value of one solution is the profit of the airline generable with the specific schedule. Therefore, the operating costs of the schedule are considered as well as the revenue gained with sold ticket. The model uses genetic algorithms to exploit the solution space and a local search algorithm subsequently to gain the local optimum in the neighbourhood of the best solutions.

The second aspect of the model considers the integration of passenger preferences into the network planning step. Here, timely preference's according the departure time in the week and the overall travel time are used to spread the given demand on the flights of the schedule. In the meantime, passengers have preferred travel options but are willing to deviate from this to a certain degree if no better flight is available.

The model formulation is suited to calculate larger airline networks with a weekly schedule, where flights are scheduled to blocks of ten minutes. However, the main purpose is not to optimize the airline network in terms of maximizing profit but rather to calculate realistic networks representing a typical network. This is necessary for various research questions that assess different parameters like aircraft characteristics or changes in passengers' behaviour. Additional studies considering these aspects will be performed utilizing the presented model.

7 References

Adler, N., Berechman, J. (2001), "Evaluating optimal multi-hub networks in a deregulated aviation market with an application to Western Europe", *Transportation Research Part A*, vol. 35, pp. 373–390.

Barnhart, C., Kniker, T. S., Lohatepanont, M. (2002), "Itinerary-Based Airline Fleet Assignment", *Transportation Science*. vol. 36, part 2, pp. 199–217.

Braun, M., Koch, A., Dahlmann, K., Grewe, V., Gollnick, V. (2010), "An Airline Network Design Approach Considerung Environmental and Economical Targets", *27th International Congress of the Aeronautical Sciences*, Nice, France.

Büdenbender, K., Grünert, T., Sebastian, H.-J. (2000), "A Hybrid Tabu Search/Branch-and-Bound Algorithm for the Direct Flight Network Design Problem", *Transportation Science*, vol. 34, part 4, pp. 364–380.

Cadarso, L., Marín, Á. (2013) "Robust passenger oriented timetable and fleet assignment integration in airline planning", *Journal of Air Transport Management*, vol. 26, pp. 44–49.

Derigs, U., Friederichs, S., Schäfer, S. (2009), "A New Approach for Air Cargo Network Planning", *Transportation Science*, vol. 43, part 3, pp 370–380.

Erdmann, A., Nolte, A., Noltemeier, A., Schrader, A. (2001), “Modeling and Solving an Airline Schedule Generation Problem”, *Annals of Operations Research*, vol. 107, pp. 117–142.

Grosche, T., Heinzl, A., Rothlauf, F. (2001), “A Conceptual Approach for Simultaneous Flight Schedule Construction with Genetic Algorithms”, *Applications of Evolutionary Computing - EvoWorkshops 2001: EvoCOP, EvoFlight, EvoIASP, EvoLearn, and EvoSTIM Como, Italy, April 18–20, 2001 Proceedings*, pp. 257–267.

Hsu, C.-I., Wen, Y.-H. (2003) “Determining flight frequencies on an airline network with demand–supply interactions”, *Transportation Research Part E: Logistics and Transportation Review*, vol. 39, part 6, pp. 417–441.

Jacobs, T. L., Smith, B. C., Johnson, E. L. (2008), “Incorporating Network Flow Effects into the Airline Fleet Assignment Process”, *Transportation Science*, vol. 42, part 4, pp. 514–529.

Kölker, K., Lütjens, K. (2015), “Using Genetic Algorithms to Solve Large-scale Airline Network Planning Problems”, *Transportation Research Procedia*, vol.10, pp. 900-909.

Koppelman, F. S., Coldren, G. M., Parker, R. A. (2008), “Schedule delay impacts on air-travel itinerary demand”, *Transportation Research Part B: Methodological*, vol. 42, part 3, pp. 263–273.

Lederer, P. J., Nambimadom, R. S. (1998), “Airline Network Design“, *Operations Research*, vol. 46, part 6, pp 785–805.

Lehner, S., Kölker, K., Lütjens, K. (2014), “Evaluating Temporal Integration of European Air Transport”, *29th Congress of the International Council of the Aeronautical Sciences (ICAS)*, St. Petersburg, Russia.

Lohatepanont, M., Barnhart, C. (2004), “Airline Schedule Planning: Integrated Models and Algorithms for Schedule Design and Fleet Assignment”, *Transportation Science*, vol. 38, part 1, pp. 19–32.

Mashford, J. S., Marksjö, B. S. (2001), “Airline Base Schedule Optimisation by Flight Network Annealing”, *Annals of Operations Research*, vol. 108, pp. 293–313.

Mercier, A., Soumis, F. (2007), “An integrated aircraft routing, crew scheduling and flight retiming model”, *Computers & Operations Research*, vol. 34, part 8, pp. 2251–2265.

Pita, J. P., Barnhart, C., Antunes, A. P. (2013), “Integrated Flight Scheduling and Fleet Assignment Under Airport Congestion” *Transportation Science*, vol. 47, part 4, pp. 477–492.

Rexing, B., Barnhart, C., Kniker, T., Jarrah, A., Krishnamurthy, N. (2000), “Airline Fleet Assignment with Time Windows”, *Transportation Science*, vol. 34, part 1, pp. 1–20.

Sherali, H. D., Bae, K.-H., Haouari, M. (2010), “Integrated Airline Schedule Design and Fleet Assignment: Polyhedral Analysis and Benders' Decomposition Approach”, *INFORMS Journal on Computing*, vol. 22, part 4, pp 500–513.

Sherali, H. D., Bae, K.-H., Haouari, M. (2013), “A benders decomposition approach for an integrated airline schedule design and fleet assignment problem with flight retiming, schedule balance, and demand recapture”, *Ann Oper Res*, vol. 210, part 1, pp. 213–244.

Tang, C.-H., Yan, S., Chen, Y.-H. (2008), “An integrated model and solution algorithms for passenger, cargo, and combi flight scheduling”, *Transportation Research Part E: Logistics and Transportation Review*, vol. 44, part 6, pp. 1004–1024.