Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Transportation Research Procedia 10 (2015) 900 – 909

Transportation Research Procedia

www.elsevier.com/locate/procedia

18th Euro Working Group on Transportation, EWGT 2015, 14-16 July 2015, Delft, The Netherlands

Using genetic algorithms to solve large-scale airline network planning problems

Katrin Kölker*, Klaus Lütjens

German Aerospace Center (DLR), Air Transportation Systems, Blohmstr. 8, 21079 Hamburg, Germany

Abstract

The airline planning process, including network planning and fleet assignments, is a complex and highly integrated strategic process involving multiple interrelated sub-problems that must be solved simultaneously. For example, analysing the effect of new technologies or changes in passenger demand or airline structures from an operational perspective requires the analysis and application of complex, structured airline planning resulting in extensive and complicated optimization problems. In this paper, we present an assessable formulation and approach to integrate network planning and scheduling as well as minor effects of aircraft rotation and passenger demand, which are solved by applying genetic algorithms. The objective function comprising direct operating costs, including costs of ownership and revenue, is based on heuristics that enhance the understanding of the airline's final profit composition.

Peer-review under responsibility of Delft University of Technology

Keywords: Strategic Airline Planning, Network Planning, Scheduling, Genetic Algorithm

1. Introduction

1.1. Motivation

The airline planning process is strategically performed by an airline prior to the day of flight operations to determine the most efficient operation for the existing fleet. It seeks to answer the questions of where to fly, when to fly, and with what aircraft to fly to gain the most profit. In quantitative research, we often ask for the influence of passenger attributes, such as time preferences, or of fleet attributes, such as speed on the network. Knowing the effect of fleet attributes on the airline net allows for a better understanding of the correlation between operational requirements and technical constraints in air transportation. The effect of new and innovative technologies can be assessed specifically from the airlines' perspective. The presented approach allows for the improved understanding of aircraft operations,

* Corresponding author.
E-mail address: katrin.koelker@dlr.de

Peer-review under responsibility of Delft University of Technology
doi:10.1016/j.trpro.2015.09.043

with fleet attributes as influencing factors. However, the technology can also be understood from the perspective of induced requirements, in which fleet attributes are adapted to the needs of airlines.

1.2. The network planning problem

Network planning problems are a main discipline in the configuration of an airline's planning process. Because a planning period is based on the airline's entire fleet and normally spans one week, it could easily comprise thousands of flights. The entire planning process consists of multiple sub-problems (Belobaba (2009)). From a given demand, the network planning process defines the routes to be operated by the airline and the number of flights operated in each of them. The next step, scheduling, defines when the flights are to occur, and the following fleet assignment fixes the operating fleet type for every flight. This allows for a defined flight plan for every sub fleet, for which the rotation planning establishes tail numbers. Later parallel processes include crew assignments, the revenue management process and fleet planning. However, these are not integrated into the modelling concept.

Therefore, to capture the effects of different fleet attributes appropriately, it is necessary to include scheduling effects and minor rotation aspects into the problem formulation as well as fleet assignments. However, this results in complex problems. If the standard approach using linear programming is applied, one quickly runs into problems that are not solvable in a reasonable runtime with today's techniques. There are integrative problem formulations that simultaneously capture network planning with scheduling and passenger paths. Lederer and Nambimadom (1998) performed an analysis comparing direct and hub-and-spoke networks based on a first network model, including the frequency of flights between airports. Barnhart et al. (2002) proposed an integrated method using an itinerary-based model to combine fleet assignments with a passenger flow model and stated the importance of capturing network effects within the airline planning process by a spill and recapture model to integrate the effects driven by passenger demand. To simultaneously model and optimize passenger flow, they used a passenger mix problem, spreading the passengers over flights given a fixed schedule. Lohatepanont and Barnhart (2004) extended this approach by integrating schedule planning. However, they did not build up a schedule from scratch but adjusted a given basic flight plan. This is similar to the methodology of Rexing et al. (2000), who addressed integrative airline planning by extending a fleet assignment problem using additional time windows to shift flights. Building up a flight plan from scratch is discussed by Yan and Tseng (2002), who developed an integrated model intended to include scheduling in the network planning process. They introduced a profit optimization program and an algorithm sufficient to solve smaller network planning problems. Their approach focuses on scheduling from scratch rather than adapting a given flight plan. Similar approaches including cargo networks based on integer and mixed integer linear programming have been made by, for example Tang et al. (2008) and Derigs et al. (2009).

Another main consideration of network planning is the impact of passenger paths and demand effects on the network. Jacobs et al. (2008) investigated a method to incorporate network effects into fleet assignments using dynamic demand based on revenue management. In this method, spill effects are integrated by iteratively applying both models and considering passengers of origin-destination markets instead of leg-based revenues. Sherali et al. (2013) proposed a similar approach including scheduling, fleet assignment and rotation using Benders' decomposition to solve the complex problem. Nevertheless, all of these approaches are based on linear programming and are highly depending on the performance of the solver.

Other studies investigated alternative solving methods to integrate more complex aspects into the optimization problem or allow faster routines. Hsu and Wen (2000) used grey theory and multi-objective programming to solve the network design problem considering passenger and airline objectives. They handled the uncertainty by initially applying Floyd's algorithm to choose route candidates and by applying grey theory afterwards to choose the final passenger routes. The airline net, including frequencies on the segments, was then derived on the given passenger routes. The approach was extended by the interaction of demand and supply in Hsu and Wen (2003). To do so, they modelled airlines' frequency determination and passengers' choices and iteratively optimized demand and supply until the chain of sub-models converged.

This paper presents an integrated model of network planning, including scheduling and some aspects of rotation planning based on the present models. Additionally, the passenger perspective is modelled by defining passenger itineraries. The difference is in the usage of genetic algorithms to solve the problem, which is mainly applied because of runtime issues but also to enhance the understanding of the structure of the optimization problem, which is necessary for an application of research questions.

1.3. Genetic algorithms

In this paper, we present a formulation in which a variable, which is a flight plan, is represented as a one-dimensional array to run genetic algorithms. These algorithms are based on the principle of evolution. The main idea of genetic algorithms is that an optimization variable is represented by an individual in the evolutionary process. From a set of individuals called a population, the algorithm randomly chooses individuals with probability given by their fitness, which is equal to the objective value of the optimization function. The chosen individuals form the parents and are recombined pairwise to produce new individuals, the so-called offspring, which comprise the next generation. The offspring undergo a mutation process in which single genes - i.e., entries in the array - are changed with a given probability. Multiple populations are sequentially produced, implying that the fitness of the population increases to find the fittest individual, which is equal to the best solution to the optimization problem (Holland (1992)). The main advantage of this optimization approach lies in the very few requirements on the search space itself and the knowledge of the general structure of the search space. Only a function that evaluates given variables by determining their fitness is required.

1.4. Introduction to the model formulation

Deriving the network from the array-based formulation is straightforward because the network is defined as the operated routes and the number of flights operated on these routes. The flight plan includes simple constraints to allow for rotation planning at a later time. Requirements of the flight plan could include the fact that the same number of aircraft at one airport at the beginning of every planning period should be equal the number at the end of the period to allow for recurrent flight plans. The fleet size is used to calculate the final capital cost. As an objective function, we use the profit function of the airline, which includes direct operating cost and revenues. Passenger movement is based on passenger path to allow for the analysis of point-to-point networks as well as hub-and-spoke-networks. Using the presented approach and genetic algorithms to solve it, we will perform a use case showing the application by means of one airline's network. This use case reveals the assumptions and data required to provide an overview for future assessments and analyses.

The rest of the paper is organized as follows. Section 2 presents the general model and the formulation of the solutions. Section 3 discusses the aircraft balance in the network. Constraints and feasibility are discussed in section 4, whereas section 5 discusses the recombination and mutation operators. Section 6 establishes the objective function as being equal to the airline's profit. Section 7 introduces the network for the use case, assumptions and preliminary results. Section 8 concludes the paper.

2. Model Formulation

2.1. General description

The presented model is an integrated optimization concept comprising network planning, scheduling and rotation planning. For simplification, this problem is reduced to one fleet type but will be expanded in the future. Because the model follows the operational airline perspective, the overall objective function of the optimization problem is the profit of the airline. The profit consists of the revenue generated by ticket prices from transported passengers minus the cost incurred by operation of the flights on the one hand and the capital costs of the aircraft on the other hand. Because the model focuses on the final flight plan, the profit is assessed from the set of eventually operated flights. The main idea of the optimization process is to generate a set of candidate flights that may include many more flights than those operated in the end and to choose a set of flights from these candidates. To generate all candidate flights, one time period is considered; this is set to one week in practice but can be changed in the model setup. However, the flight plan is configured such that it can be rolled over; i.e., the same flight plan can be operated seamlessly afterwards. This requires that the number of aircraft based at one airport at the beginning of one planning period equals the number of aircraft at the same airport when the period ends. Additionally, the number of aircraft necessary to operate the flight plan is assessed. Nevertheless, the fleet size can easily be fixed if required. Given passenger demand is the basis for building up the network. The demand defines the number of passengers travelling between airports, their possible itineraries (direct or stopover at a hub) and their willingness to pay. This allows for the network model to be merged with a passenger flow model when the next level of complexity is subsequently approached.

The formulation of a time-space network based on the airline network fosters the formal representation of the flights and the flight plan. Every node in the network is one airport at a specific point in time. Thus, a flight is a directed connection between two nodes in the time-space network, connecting the origin node at the departure time and the arrival node at the arrival time.

Modelling the time between the departure and arrival times is performed using the block time of the flights. However, each aircraft has a specific minimum ground time, which is the minimum time the aircraft must stay at the airport to execute the turnaround process. The block time of a flight plus the minimum ground time is the time span in which the aircraft is occupied with one flight and cannot operate another. This time is called flight time hereafter and is used when adding flights to the time-space network. Therefore, the minimum ground time is always satisfied and integrated in the model.

2.2. Mathematical formulation

The network planning problem and optimization method can be formally defined given a set of airports $A = \{a_i\}_{i=1,...,k}$ and a set of points in time $T = \{t_i\}_{i=1,...,l}$. These two sets enable the definition of the set of candidate flights $F = \{f_i\}_{i=1,...,n}$, where $f_i = (o_i, d_i, t_i^1, t_i^2)$; $o_i, d_i \in A$ are the origin and destination of the flight and $t_i^1, t_i^2 \in T$ are the departure and arrival times based on the flight time. To represent the passenger perspective, $M = \{m_i\}_{i=1,...,o}$ with $m_i = (a_i^1, a_i^2)$, where $a_i^1, a_i^2 \in A$ is a set of passenger markets; the set of passenger itineraries corresponding to this market is $I = \{itin_i^m\}_{i=1,...,r(m)}$, where $itin_i = (g_1, \ldots, g_s)$ is a connected multiple of flights; i.e., $g_j \in F$ and for all $g_j = (o_j, d_j, t_j^1, t_j^2)$ and $g_{j+1} = (o_{j+1}, d_{j+1}, t_j^1, t_j^2)$ we have $d_j = o_{j+1}$ for all $j = 1, \ldots, s-1$. Additionally, we define a revenue function assigning an average ticket price to every itinerary $rev(itin_i) \in \mathbb{R}^{\geq 0}$, a capacity for every flight $cap(f_i) \in \mathbb{N}$, the operating cost for every flight $coc(f_i) \in \mathbb{R}^{\geq 0}$, the capital cost for one aircraft $cc \in \mathbb{R}^{\geq 0}$ and the maximum demand in every market $demand(m_i) \in \mathbb{N}$. A time-space network is defined as a graph $G = (A \times T, F)$ consisting of nodes $(a, t) \in A \times T$ and edges $f \in F$. An edge $f_i = (o_i, d_i, t_i^1, t_i^2)$ connects nodes (o_i, t_i^1) and (d_i, t_i^2). Fig. 1 illustrates a simple example of the networks that comprise the network planning problem.

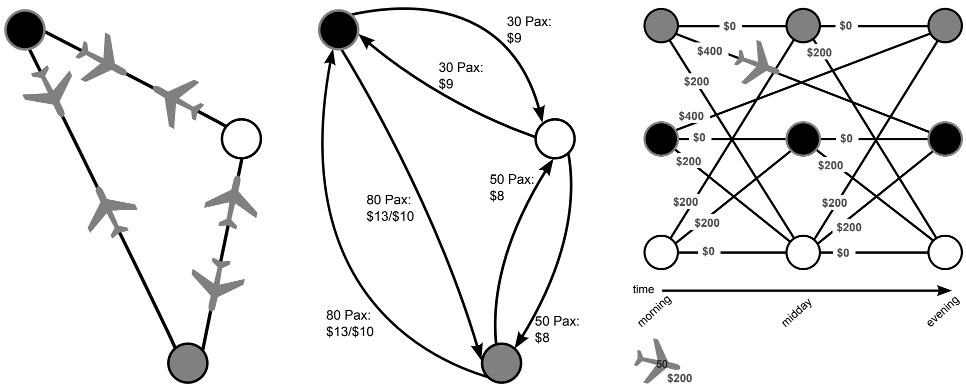

Fig. 1. (a) airline network consisting of three airports with one hub (white); (b) network with given passenger demand and ticket price ($13/$10 means $13 for the direct route and $10 for the route using white hub); (c) time-space network.

To define the optimization problem as a mixed integer program, we define three sets of decision variables. $x = (x_i)_{i=1,...,n}$ indicates a flight plan, which means that $x_i \in \{0, 1\}$ and $x_i = 1$ if and only if flight f_i is operated. The second decision variable represents the passenger flow. $p = (p_i)_{i=1,...,s}$ is the number of transported passengers belonging to itinerary $itin_i$, whereas p_i^m is the number of transported passengers belonging to itinerary $itin_i^m$ of market m. The last decision variable ac defines the number of aircraft required to operate the network. The airline network planning problem is defined as follows:

$$\max \sum_{i=1,\dots,s} rev(itin_i) \cdot p_i - \sum_{i=1,\dots,n} coc(f_i) \cdot x_i - ac \cdot cc$$

$$s.t.$$

$$\sum_{i=1,\dots,r(m)} p_i^m \leq demand(m) \qquad \text{for all } m \in M \tag{1}$$

$$\sum_{p_i : f \in itin_i} p_i \leq cap(f) \qquad \text{for all } f \in F \tag{2}$$

$$\sum_{f_i=(\cdot,a,\cdot,t)} f_i = \sum_{f_i=(a,\cdot,t,\cdot)} f_i \qquad \text{for all } t \in \{t_2,\dots,t_{l-1}\}, \text{ for all } a \in A \tag{3}$$

$$\sum_{f_i=(a,\cdot,t_0,t_1)} f_i = \sum_{f_i=(\cdot,a,t_{l-1},t_l)} f_i \qquad \text{for all } a \in A \tag{4}$$

$$\sum_{a \in A} \sum_{f_i=(a,\cdot,t_0^1,\cdot)} f_i = ac \tag{5}$$

$$itin_i \in \mathbb{Z}, x_j \in \{0,1\}, ac \in \mathbb{Z} \qquad \text{for all } i = 1,\dots,s \tag{6}$$

Note that constraints (1) ensure that the number of passengers does not exceed the maximal demand on the market. Constraints (2) guarantee that the number of passengers on every flight is at most the given capacity. Constraints (3) and (4) maintain aircraft balance by ensuring that every aircraft entering an airport at a specific time is leaving it as well and by ensuring that the number of aircraft at the end of the planning period equals that at the beginning. Finally, constraint (5) counts the number of aircraft by counting all aircraft at one specific point in time.

2.3. Optimization variables

The formulation of optimization variables uses the time-space network. An individual is consistent with a closed rotation of aircraft in the network in which the solutions do not have constant length and that describes a closed circle in the underlying graph. The definition is analogous to the formulation of the traveling salesmen problem presented by Michalewicz and Fogel (2004). Technically these are variables $y_i = (node_j)_{j=1,\dots,u}$, $node_j \in A \times T$ in which induced flights are flights that connect consecutive nodes. However, the operator definitions that adapt the variables of the genetic algorithm must guarantee the feasibility of the new individuals.

For example, a possible rotation in the network above would be $y_i = ((grey, morning), (black, evening), (black, morning), (white, midday), (grey, evening))$. Because the first and last nodes for an airport are exclusively connected, they do not have to be listed twice and the variable can be reduced to $y_i = ((grey, morning), (black, evening), (white, midday))$. The induced flights are the morning flight from grey to black and from black to white to blue as well as the midday flight from white to black. This example rotation leads to a final fleet size of two aircraft.

The following chapters explain how the variables are assessed within the implementation of the genetic algorithm.

3. Balancing Aircraft in the Network

Aircraft balancing is a main consideration in scheduling and fleet assignment. However, rotation planning is not fully integrated into the problem. Thus, there is a guarantee only for the option of a feasible aircraft rotation, which is induced by the fact that constraints (3) and (4) ensure that the number of flights entering an airport-time node is equal to the number of flights leaving an airport-time node. This includes the aircraft standing on the ground at a specific airport. Consequently, the constraints allow for a flight only when an aircraft is available at the departure airport. From this structure, the minimum number of required aircraft can be easily retrieved, which will be important in the following evaluation methods. To specify the structure and establish the balance mathematically, we define a matrix that represents the balance of the time-space network. For every point in time and airport, we count the number of arriving flights less the number of departing flights up to that point. This number is equal to the demand or surplus of aircraft at the airport at that time.

$$B = (b_{ij}) \in \mathbb{Z}^{k \times l}, \quad b_{ij} = |\{f_h \in F \,|\, f = (\cdot, a_i, \cdot, t \leq j)\}| - |\{f_h \in F \,|\, f = (a_i, \cdot, t \leq j, \cdot)\}|$$

The balance of the network is maintained if the number of arriving flights equals the number of departing flights at the end of the time period i.e., $b_{il} = 0\, \forall i = 1, \ldots, k$ which means that the number of aircraft at one airport at the beginning is equal to the number of aircraft at the end of the time period. The number of total aircraft required to operate the schedule is the sum of all aircraft at every airport at the beginning of the time period. Every airport must have aircraft available to operate all flights leaving that airport. The number of aircraft required at the beginning of the planning period at one airport is equal to the minimum negative value in the balance matrix for that airport. Thus, the sum of all required aircraft is

$$ac = \sum_{a=a_1,\ldots,a_k} \min\{0, b_{a1}, \ldots, b_{al}\}.$$

4. Meeting Constraints and Maintaining Feasibility

The model and the optimization variables are formulated such that the constraints are automatically met while generating the objective value of a variable. The constraints considering the passenger demand in one market (1) as well as the constraint ensuring that the capacity of every flight is not exceeded (2) can all be met by defining an appropriate number of passengers for every itinerary in the objective function. The heuristic to do so is presented in section 6. The constraint counting the maximum number of aircraft needed to operate the network (5) is generated by counting the number of aircraft as proposed in section 3 and applying this number directly to the profit. The third and fourth constraints ask for the balance of aircraft within the network, meaning that the number of aircraft leaving one time-space node must equal the number of aircraft entering that node. Because the formulation of individuals considers a full rotation from the aircraft perspective, this constraint is automatically met. However, the recombination and mutation operators are defined such that all generated variables are continuously feasible.

5. Operators of the Genetic Algorithm

A genetic algorithm requires recombination and mutation operators. Both are required to investigate the search space - i.e., the whole space of feasible solutions - without an overall knowledge of the structure of the space. The recombination process consists of selecting existing variables and combining them, which is called the exploration part of the algorithm. The mutation process covers the exploitation by concentrating on the local search starting with the existing variable (Eiben and Smith (2003)).

In this approach, the recombination operator combines two chosen parent variables to produce two offspring variables. Classical recombination operators are, for example, the one-point crossover presented by Holland (1992). From one generation, parents are chosen according to the wheel selection in which the probability to be chosen is proportional to the fitness of the individual. Two of these chosen individuals produce individuals by simply crossing over the array. A random crossover point in the array of individuals is chosen, and the first offspring is built up from the first part of the first parent and the second part of the other parent; the second offspring is built up via the reverse process. In this approach, the genes are based on rotations, meaning that consecutive genes are connected by the fact that a flight is offered between the two airport-time nodes represented by the two genes. Thus, the recombination operator must be restricted to generate feasible offspring.

The recombination operator requires that two rotations meet at two different points. One rotation represents the whole flight plan, meaning that multiple aircraft are included in it. Two complete rotations can be recombined only if they meet at as many as two different nodes in the time-space network. In the final flight plan, this means that two aircraft are at the same airport at the same time. Michalewicz (2004) provides an extensive overview of possible operators to use with the travelling salesman problem, which can be used here. Nevertheless, the time-space network is not a complete graph, and only some nodes are connected by flights. Therefore, the algorithm uses a modified 2-point crossover, which is adapted to the time-space network in a way in which only points that are equal in both parents can be crossing points. The two resulting offspring consist of the first and last part of one parent and the middle part of the second parent.

Following the recombination process, the mutation operator performs slight and random changes to the variables by altering one or more genes, which are the entries in the representing array. The genes of the so-defined offspring for the next generation undergo a simple mutation process in which a gene is flipped with a given probability. In the implementation presented in this paper, the mutation operator mutates one flight with a given probability such that a random other flight also departing from the node is chosen. This means that the rotation does not follow the original flight but follows another flight leaving at the same time at the same airport. A return flight is added so that the rotation returns to the original rotation as soon as possible. The variables can be recombined and rotated as illustrated in Fig 2.

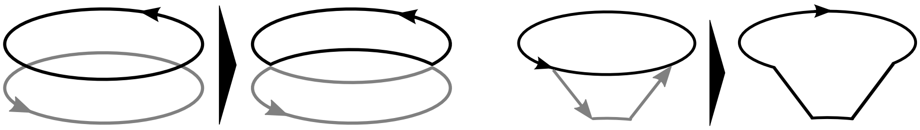

Fig. 2. (a) recombination operator (black and grey are two different rotations meeting at two nodes in the time-space network); (b) mutation operator for formulation based on aircraft rotation.

6. Objective Function

The objective function comprises two separated opposing modules: the costs and the revenue share. The negative part is generated by the cash operating costs and the costs of ownership in the objective function. Because the precise flights are given, the cash operating costs can be calculated straightforwardly. However, as shown in section 3, the number of required aircraft can easily be assessed by counting the aircraft at the beginning of the time period at every airport, which is equal to the minimum negative number in every row of the balance matrix. This can be accomplished because the final quantity of flights is known, and the balance matrix can thus be generated. Nevertheless, it is not even necessary to generate the whole matrix. It is sufficient to go along every row and identify the minimum negative value. This value equals the number of aircraft required to operate the network. The second part of the objective functions is generated by the revenue comprising the ticket prices. The maximum revenue that can be generated by the given flights can be assessed using linear programming or formulated as a minimum-cost flow problem. However, because this methodology must be applied for every new solution, these are too time consuming and are therefore replaced by a heuristic. The example above shows that it is not sufficient to sort itineraries and sequentially assign them. For example, the solutions contain the morning flights from the black to the white airport and the return flight at midday as well as the morning flight from grey to white and the return flight at noon. If the heuristic would sort the itineraries by ticket price and sequentially assign passengers, the revenue of the total flight plan would be

$$50 \cdot \$10 + 50 \cdot \$10 = \$1000.$$

If the value for the order of itineraries is the average ticket price per flight, which means that the longer itinerary has a modelled ticket price of \$5, and the same methodology is applied, the revenue is

$$30 \cdot \$9 + 30 \cdot \$9 + 50 \cdot \$8 + 50 \cdot \$8 = \$1340.$$

However, it can easily be observed that the optimum assigns 30 passengers to the four itineraries connected to the white airport and 20 passengers to the long itinerary. This solution gains a revenue of

$$30 \cdot \$9 + 30 \cdot \$9 + 30 \cdot \$8 + 30 \cdot \$8 + 20 \cdot \$10 + 20 \cdot \$10 = \$1420.$$

The definition of the objective function requires a heuristic that considers competing itineraries. To do so, we define for every itinerary additional artificial itineraries consisting of multiple itineraries so that the set of flights as the basis of the artificial itinerary is equal to the flights in the original itinerary. The ticket price of the artificial itinerary is the sum of all itineraries within the artificial itinerary. For the heuristic to determine passenger allocation, all itineraries including the artificial itineraries are sorted according to the sum of all ticket prices and sequentially assigned to the flights. If an artificial itinerary is picked, only the minimum number of passengers of all itineraries in the artificial itinerary is assigned to the underlying flights.

In the example, the itinerary (black,white,grey) with one stopover at the white airport and a ticket price of \$10 would generate one artificial itinerary consisting of two more itineraries ((black,white),(white,grey)) with a ticket price of \$8 + \$9 = \$17. Sorting all itineraries and choosing them according to their ticket price results in the artificial itineraries being selected first, and 30 passengers are assigned to each of the flights attached to the white airport. Afterwards, the long itinerary is picked, and 20 passengers are assigned to each flight. This results in the optimal solution. The example network has an optimal solution of \$320 profit in which six flights are offered with four aircraft and 260 transported passengers. The formalized algorithm to retrieve the revenue of a given set of flights is shown in Algorithm 1.

Algorithm 1 Revenue of flight plan

Input: set of flights $x = \{x_i\}_{i=1,\dots,n}$, where $x_i = 1$ iff f_i is operated
Output: $rev(x_i)$
$\hat{I} = I \cup \{i_g^{mult}\}_g$, where $i^{mult} = (itin_1^{mult}, \dots, itin_u^{mult})$ and $itin_j^{mult} \in I$
$revenue = 0$
while $\hat{I} \neq \emptyset$ **do**
take $it\hat{i}n \in \hat{I}$ so that $rev(it\hat{i}n) = \max_{\check{i} \in \hat{I}}\{rev(it\check{i}n)\}$ where $rev(itin^{mult}) = \sum_{j=1,\dots,u} rev(itin_j^{mult})$
assign $pax = \min_{f_i \in it\hat{i}n}\{cap(f_i)\}$ passengers to $it\hat{i}n$
$revenue = revenue + pax \cdot \sum_{itin_j^{mult} \in it\hat{i}n} rev(itin_j^{mult})$
remove $it\hat{i}n$
end while

7. Implementation and Use Case Study of the Model

This paper focuses on the formulization of the model. Nevertheless, the functionality is shown using a European airline net and one specific week to demonstrate the execution of the model as well as essential assumptions and data. Therefore, we take the A320-200 net of a European carrier operating two hubs in the first week of June 2014. Passenger data, including the passenger paths, and schedule data, including the exact time of departure, are taken from Airport Data Intelligence (2015). However, passengers are able to stop over only at the first hub. This means that flights are offered from airports all over Europe to both hubs; however, passenger itineraries either are direct or include a stopover at the main hub. Fig. 3 shows the geographical expansion of the net.

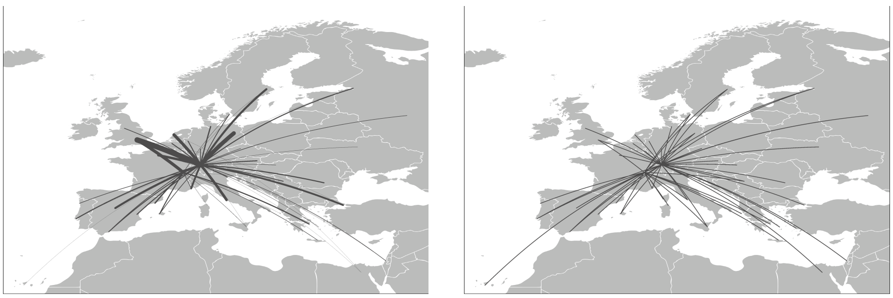

Fig. 3. (a) demand network of selected airline net, width of connection indicates amount of demand; (b) network of potential flights.

Because no demand model is available thus far, the model uses real passenger data. To obtain clean and utilizable passenger data, the following methodology is performed. First, all passenger paths are cut such that only segments in which at least one flight is operated by an A320 are considered. For passenger paths in which a segment is operated by more than one fleet type, we take the share of passengers transported on A320s. This is calculated as the total share of the A320 capacity compared to the capacity of all flights on this path. In the given network, 37.7% of all passengers are assumed to travel on an A320. Second, we assume that there are two hubs of the airline. Passengers are routed only via the main hub, whereas direct passenger paths may be connected to both hubs. This means that a possible flight is always connected to one of the hubs. To gain more insight into the model and clean the data, all other passenger paths are neglected. However, this affects 0.8% of all remaining passengers, which leads to the conclusion that all main passenger paths are included in the model. Third, because the passenger itineraries are automatically derived in the model and demand is given only between two markets, all itineraries can be grouped into markets. Additionally only markets with more than 200 passengers per month are considered because all others do not seem to be rentable a priori. Again, 5.1% of the remaining passengers but 88.5% of all markets are deleted, meaning that the most profitable markets are still considered. Finally, the monthly passenger volume is broken down to one week, which results in approximately 77 000 passengers divided into 122 markets. Table 1 displays the main descriptive figures of the modelled input demand network. Note that a passenger can occupy more than one seat if he is using a hub and is assigned to multiple flights.

Table 1. Demand and flight network of use case.

	Total Airline Net	Net of A320
Flights (per week)	2 585	933
Seats (per week)	403 371	156 744
Passengers (per week)	218 242 (935 327 per month)	77 501 (332 393 per month)

The block time for every flight is calculated using the great circle distance between the airports and a regression function calculated using the real flight plan of the reference week ($R^2 = 0.97$), and it is given by

$$blocktime[min] = 0.0737[min/km] \cdot distance[km] + 40.544[min]$$

Candidate flights are scheduled every 10 min considering curfews between 10 p.m. and 6 a.m. To estimate cash operating and ownership costs, the model of Swan and Adler (2006) is used. It defines the cash operating costs in terms of the distance and capacity of an aircraft. This approach suffices for estimating and distinguishes itself because different use cases of network planning are more comparable than other more complex models. It also provides an estimation for the costs of ownership.

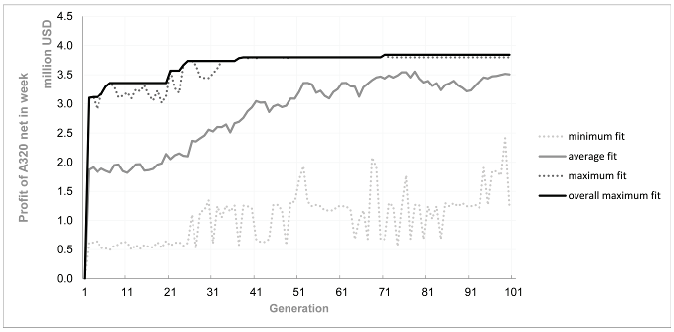

Fig. 4. development of the fitness for a run with the rotation based formulation.

The algorithm is completely implemented with JAVA to allow full access to and modification of the methodology and all operators. The runtime is approximately 15 min, half of which is consumed by the initialization and sorting. Standard software for solving linear programs is not capable of addressing the problem with standard computational power without extensive adaption and speed-up techniques. Fig. 4 shows a run with a formulation based on rotations. The maximum profit of the airline is approximately \$3.8 million per week, which seems reasonable for an airline of this size, as this considers only the operational profit. Indirect operating costs are not considered because they are not directly influenced by the network and are constant when changing the net with the given methodology. However, the final modelled fleet size of the airline is 8, approximately one-third of the real fleet size. The very high cost of ownership (approximately \$200 000 per week) compared to the direct operation costs might be one cause of this. Adaptations of the cost model would correct this. Additionally, the underlying airline is a full-service network carrier operating out of two hubs. This results in sub-networks belonging to different fleet types of the airline affecting each other. For example, there might be a demand for a feeder flight prior to a long-haul flight that is not considered in the model. These airline hubbing strategies require additional aircraft in the fleet, which is not yet modelled.

The figure also displays the development of the minimum, average and maximum fit of one generation throughout the optimization process. The steep slope in the beginning, which changes to a nearly asymptotic trend with a small slope, along with local unsteadiness suggests that the exploration and exploitation are well balanced in the methodology.

8. Conclusion and Outlook

In this paper, we introduced a model for integrated airline planning processes, including network planning and scheduling. The optimization variable is based on the aircraft rotations, in which the flights are given by the edge between two nodes in the time-space network. The objective function maximizes the overall profit of the airline while considering the costs of owning and operating the aircraft as well as revenues generated by ticket prices paid by transported passengers. The utilized optimization method is a genetic algorithm based on the evolutionary concept that uses standard as well as new recombination and mutation operators. The integrative approach and methodology allows for the assessment of complex research topics such as the influence of passenger preferences or fleet attributes on the network of a single airline. The latter can be already analysed by quantifying the effect of varying speed, minimum turnaround time, capacity or operating costs. However, the assumptions and simplifications in the use case will be discussed and improved in future studies to generate a more detailed and accurate application. Additionally, the passenger perspective has not yet been completed so future work will discuss the integration of a market model into this concept.

References

Belobaba, Peter; Odoni, Amedeo R.; Barnhart, Cynthia. 2009. The global airline industry. Chichester, U.K: Wiley (Aerospace series).

Derigs, Ulrich; Friederichs, Stefan; Schfer, Simon, 2009. A New Approach for Air Cargo Network Planning. Transportation Science 43 (3), pp. 370-380.

Holland, John H., 1992. Adaptation in natural and artificial systems. An introductory analysis with applications to biology, control, and artificial intelligence. 1st MIT Press ed. Cambridge, MIT Press (Complex adaptive systems).

Hsu, Chaug-Ing; Wen, Yuh-Horng, 2000. Application of Grey theory and multiobjective programming towards airline network design. European Journal of Operational Research 127 (1), pp. 44-68.

Hsu, Chaug-Ing; Wen, Yuh-Horng, 2003. Determining flight frequencies on an airline network with demandsupply interactions. Transportation Research Part E: Logistics and Transportation Review 39 (6), pp. 417-441.

Jacobs, Timothy. L.; Smith, Barry C.; Johnson, Ellis L., 2008. Incorporating Network Flow Effects into the Airline Fleet Assignment Process. Transportation Science 42 (4), pp. 514-529.

Eiben, Agoston E.; Smith, James E., 2003. Introduction to evolutionary computing. Springer (Natural computing series), NewYork.

Barnhart, Cynthia; Kniker, Timothy S., Lohatepanont, Manoj, 2002. Itinerary-Based Airline Fleet Assignment. Transportation Science 36 (2), pp. 199-217.

Lederer, Phillip J.; Nambimadom, Ramakrishnan S., 1998. Airline Network Design. Operations Research 46 (6), pp. 785-805.

Lohatepanont, Manoj; Barnhart, Cynthia, 2004. Airline Schedule Planning: Integrated Models and Algorithms for Schedule Design and Fleet Assignment. Transportation Science 38 (1), pp. 19-32.

Michalewicz, Zbigniew; Fogel, David B., 2004. How to solve it: modern heuristics. Second Edition, Springer, Berlin.

Rexing, Brian; Barnhart, Cynthia; Kniker, Tim; Jarrah, Ahmad; Krishnamurthy, Nirup, 2000. Airline Fleet Assignment with Time Windows. Transportation Science 34 (1), pp. 1-20.

Sabre, Airport Data Intelligence (ADI), 2015. Sabre Airline Solutions.

Sherali, Hanif D.; Bae, Ki-Hwan; Haouari, Mohamed, 2013. An Integrated Approach for Airline Flight Selection and Timing, Fleet Assignment, and Aircraft Routing. Transportation Science 47 (4), pp. 455-476.

Swan, William M.; Adler, Nicole, 2006. Aircraft trip cost parameters: A function of stage length and seat capacity. Transportation Research Part E: Logistics and Transportation Review 42 (2), pp. 105-115.

Tang, Ching-Hui; Yan, Shangyao; Chen, Yu-Hsuan, 2008. An integrated model and solution algorithms for passenger, cargo, and combi flight scheduling. Transportation Research Part E: Logistics and Transportation Review 44 (6), pp. 1004-1024.

Yan, Shangyao; Tseng, Chich-Hwang, 2002. A passenger demand model for airline flight scheduling and fleet routing. Computers & Operations Research 29 (11), pp. 1559-1581.