Airline Optimization

Kamelia Drenkova, Abhijit Joshi, Betzalel Moskowitz, Lipika Rana
MSDS460-DL: Decision Analytics
Northwestern University
Dr. Joe Wilck
March 10, 2024

Abstract

Optimizing flight scheduling for airlines is critical for maximizing profitability and operational efficiency. Airline companies consider market demands, operational costs, and logistical constraints among other factors as they make changes to improve their operations. This paper explores profit maximization and routing optimization of various routes through building two programming models based on fare costs and operating expenses modeled on Southwest airlines. The first model identifies profitable routes and highlights factors, such as market demand and operating costs. The second model is a mixed-integer model which aims to optimize flight schedules based on provided demand data. Several assumptions were made to simplify the problem. The model minimizes operating costs while meeting demand constraints and considering layover times. The computational experiments explored in this research demonstrate the effectiveness of the models in providing optimal solutions for routing and flight scheduling. Since increased complexity created challenges in executing the models, reasonable assumptions were made to minimize the scope of the problem to fit within the given resources and time-frame. Moreover, the complexity issues emphasize the need for more powerful computers in conjunction with heuristic algorithms to tackle larger-scale problems. Future research should focus on addressing the vast solution space and deriving realistic data-driven insights that are easily applicable to real-life problems. Overall, this study produced an optimal flight schedule which maximized profits.

Key Words: *airline, network, optimization, routing, scheduling*

Introduction

One of the most prominent challenges in the airline industry is creating an optimal routing schedule to maximize profit. In general, an airline's profitability depends on various factors, including market demand, aircraft availability and size, operating costs, and competition. The current market state significantly impacts an airline's profit margins. Identifying high-demand routes and adjusting pricing can offer airlines a competitive advantage. The size and availability of aircraft directly influence the potential for revenue generation and operational efficiency. Operational costs account for a substantial portion of the expenses that airlines incur. These costs encompass a wide range of expenditures, with fuel pricing, crew costs, and maintenance costs among the most important.

In addition to the considerations mentioned earlier for airline companies, optimizing passenger occupancy is another crucial factor. Each instance of a flight departing with unoccupied seats represents a missed opportunity for revenue generation (Subramanian et al. 1994). This phenomenon is known as "the problem of fleeting schedule". The core challenge for airlines lies in balancing the use of smaller capacity planes, which risk leaving potential passengers behind, against deploying larger capacity aircraft that may fly with empty seats, thus generating additional expenses. Therefore, airline schedules should be designed to capture as much business as possible while mitigating the potential costs (Subramanian et al 1994).

The objective for this project was to enhance the profitability and efficiency of airline operations. To achieve this goal, we formulated two main strategies: one focusing on resource allocation optimization and the other on demand-based flight scheduling. The hypothetical airline company discussed in this manuscript is modeled after Southwest Airlines. Various sources were utilized to gather accurate operations information including the Southwest and the

U. S. Department of Transportation websites. Lastly, the routing component of the problem was explored only by concentrating on the 15 largest airports in the United States.

In order to achieve the first strategy, we created a linear program that takes into consideration the per destination demand in number of passengers, the operational costs, and outputted For the second strategy, we created mixed-integer programming model that predicted demand data for a given day and determining the most optimal flights to operate at specified times in order to route passengers to their final destinations through either direct or one-stop connecting flights.

Literature Review

In the search to optimize profits and streamline operations, we sought inspiration from successful case studies in similar industries. The UPS air network optimization is a prime example of a meticulous optimization approach yielding significant cost savings and operational improvements. This literature review delves deeper into the UPS methodology, particularly focusing on the design sets utilized in their formulation, and explores how insights from UPS can inform optimization efforts in the context of Southwest Airlines.

UPS's optimization methodology centers around the composite-variable formulation applied to their next-day-air network design problem. This approach involves the construction of composite variables, incorporating various design sets such as fleets, routes, arcs, gateways, hubs, and origin-destination commodities. By systematically combining these elements and considering factors like route costs, package demand, and aircraft capacities, UPS achieves optimal fleet assignments, routes, and package flows. Implementing this methodology has led to substantial cost savings, improved planning processes, and organizational acceptance within UPS.

UPS leverages design sets for optimization. Central to UPS's optimization approach is the utilization of design sets to represent key elements of their air network. These design sets encompass a wide range of parameters, including fleet types (F), routes flown by fleet types (Rf), arcs in the time-space network (A), gateways (G), hubs (H), pickup routes (RP), delivery routes (RD), and origin-destination commodities (K). By carefully defining and manipulating these design sets, UPS formulates optimization models that accurately capture the complexities of their network design problem (Armacost et al. 2004).

UPS's emphasis on comprehensive design sets enables them to capture the intricate interdependencies within their air network, facilitating the generation of optimal fleet assignments, routes, and package flows. For Southwest Airlines, adopting a similar approach entails identifying and defining relevant design sets tailored to their specific operational context. By systematically analyzing fleet capacities, route options, demand patterns, and infrastructure constraints, Southwest can develop optimization models that effectively balance profitability and operational efficiency.

Next, UPS incorporates composite variables for optimization. In addition to individual design sets, UPS leverages composite variables to streamline the optimization process further. These composite variables encapsulate complex relationships between design elements, such as the demand between gateways and hubs, aircraft allocations for pickup and delivery routes, and route costs within composite routes. By integrating composite variables into their optimization models, UPS achieves a more holistic representation of their air network dynamics, facilitating more robust decision-making and cost-effective solutions (Armacost et al. 2004).

For Southwest Airlines, utilizing composite variables offers a similar opportunity to enhance the accuracy of their optimization models. Through synthesizing multiple design elements into composite variables, Southwest can capture complex interactions and trade-offs inherent in their routing and scheduling decisions. This approach enables Southwest to optimize fleet utilization, minimize operational costs, and maximize revenue generation, ultimately leading to improved profitability and competitive advantage.

Finally, there was a significant role of the VOLCANO System in UPS's optimization.The VOLCANO system, developed by UPS, plays a pivotal role in their optimization efforts. This system extends beyond simpler optimization modeling; it encompasses modules for handling various data types, including aircraft operating characteristics, customer service time windows, airport characteristics, and package volume. VOLCANO has a sophisticated network-design optimization problem solved using a composite-variable approach. This methodology enables UPS to efficiently assign aircraft fleets, determine optimal routes, and ensure feasible package flows, thereby enhancing the overall efficiency of their air network (Armacost et al. 2004).

For our project, we acknowledge that considering all variables similar to UPS's VOLCANO system is beyond our current scope. However, we draw inspiration from UPS for future large-scale projects. The VOLCANO system at UPS is a prime example of how advanced optimization techniques can transform airline operations. By delving into the design principles and methodologies underpinning VOLCANO, Southwest Airlines can extract valuable insights to shape their optimization framework. Incorporating elements such as composite-variable formulations and comprehensive data handling modules, Southwest can elevate their decision-making processes, refine resource allocation strategies, and ultimately drive profitability to new heights.

Methods

4.1 Data Collection

Our data collection process was thorough and well-structured, offering a comprehensive understanding of the factors influencing Southwest Airlines' operations. By focusing on key metrics such as airport hubs, flight duration, fare costs, passenger demand, and operational expenses, we delved into the intricate dynamics of the airline industry.

The identification of top airport hubs served as a fundamental step in understanding Southwest's network and operational priorities. By pinpointing these hubs, we laid the groundwork for analyzing the flow of passengers and resources within Southwest's network, enabling strategic decision-making.

The creation of matrices for each origin-destination pair demonstrated a meticulous approach to capturing the complexities of flight scheduling and resource allocation. By incorporating variables such as flight duration, fare costs, and passenger demand, we effectively mapped out the landscape of Southwest's route network. This granular analysis provided valuable insights into the demand-supply dynamics of various routes, informing decisions on flight frequency, pricing strategies, and resource allocation.

Flight duration and fare costs data were directly sourced from the Southwest Airlines website, ensuring accuracy and reliability. To mitigate the influence of seasonal fluctuations, information was collected for a specific date: February 19th. We obtained average prices and durations for flights from each origin to every destination, enabling a precise assessment.

Destination-specific demand data were acquired from the U.S. Department of Transportation's website. The website provided the number of passengers flying to particular airports with Southwest Airlines. For modeling purposes, we determined the daily average demand per airport from that information.

The operational cost came from a dashboard made by Trefis showcasing a breakdown of Southwest Airlines flight operating costs. Factors such as flight capacity, the distance in nautical miles between destinations, and the duration of flights were inputs on the dashboard to determine an estimation of the operational cost. The flight capacity was assumed at 180 seats. This operational cost encompasses various elements including fuel expenses, employee expenditures, rental fees, maintenance costs, and marketing expenditures.

4.2.1 Determining profitable routes

After studying how the UPS team optimized their network and collecting data from Southwest Airlines, it became clear that the optimization task was quite complex. We decided it would be best to approach the problem as a routing problem and initially to identify the efficient routes. This was motivated by the data taking on point values - fares and operating costs for each route. We combined the exploratory data analysis (EDA) phase and built a prototype routing model as one step. Obviously, Southwest was already flying its services to the 15 cities we selected. However, we wished to assess which routes were profitable and determine any limiting factors that may affect those profits. The routing problem was unique in that there were either direct connections from one city to another or connections enabling travel between the two cities via connecting flights. We assessed that direct connections between 15 cities would result in 210 unique routes, and while Southwest was not necessarily flying all of them, combining them with connecting routes would make it a large problem. For this reason, we decided to build the prototype for a reduced set of cities and routes, complete our analysis, and then scale up later.

4.2.2 Assumptions

The routes we selected for the model included 2 major hub cities for Southwest, Midway and Dallas. We combined routes that were long distance, for example, Baltimore to Los Angeles, and shorter commuter flights, such as Dallas to Houston and back. Demand for indirect connections, such as Boston to Oakland, would need to be included in the demand from the hub cities. We added routes from the two hub cities to all the other destinations, except kept a unique route from Midway to Boston and Dallas to Houston that it didn't share with the other hub. In reality, those routes exist, but we thought that would be a good routing exercise for the algorithm to tackle as it lends to the possibility of dropping or replacing less profitable routes for Southwest. We assumed the supply was 12000 passengers from the hub nodes and different demand levels at the destination cities. We did not include all flights returning to the source city as we analyzed that can be easily reversed from the model to model departure flights. However, we did include a smaller set of flights, especially on the commuter leg - Houston to Dallas and Dallas back to Houston. Those were modeled as separate routes. The two other assumptions we added were the airport capacity upper limit which would be determined by the amount of flights allowed and we set the minimum limit to 1 flight so that that route is exercised which allows the schedule to account for 1 flight per destination at a minimum. The other consideration was that this was a one-day operation.

```
routes = ['DAL_HOU', 'DAL_LAS', 'DAL_DEN', 'DAL_OAK', 'DAL_LAX','DAL_BWI',
          'MDW_LAS', 'MDW_DEN', 'MDW_OAK', 'MDW_LAX', 'MDW_BOS','MDW_BWI',
```

```
'LAS_DEN', 'DEN_BOS', 'LAS_HOU',
'LAX_BWI', 'LAX_DAL',
'BWI_MDW', 'BWI_DAL',
'HOU_DAL', 'HOU_OAK']
```

We assumed the average airplane capacity to be 180 as our operating costs were calculated using that value. Further, we assumed the daily flight limit per route to be 8. We could easily change this limit and rerun the model. We assumed the hub airports, MDW and DAL, would have the capacity to fly 8 flights per day to each of the connecting cities. For this problem, we assumed the airplanes ran at full capacity. While tackling the flight scheduling problem, we addressed this limitation.

4.2.3 Decision Variables

The routes we selected for the model included two major hub cities for Southwest – Chicago Midway (MDW), and Dallas (DAL) – along with seven other cities. We combined various types of routes, including long-haul flights like Baltimore (BWI) to Los Angeles (LAX), short-haul commuter trips such as Dallas (DAL) to Houston (HOU) round trip, and routes with indirect connections like Boston to Oakland, covering both directions. We added routes from the two hub cities to all the other destinations, but structured the network such that BOS and HOU only maintained connections to the rest of the network through MDW and DAL respectively. There were 23 routes in total, and the incoming and outgoing flights that created the network (see constraints) formed the major setup in the problem. In our formulation we designate X_r to denote the number of passengers flying from the source to the destination route r, where r is one of the routes defined in the 'routes' list. The decision variables can be represented as:

$$X_r \; \forall r \in routes, \; X_r \geq 0$$

4.2.4 Objective

The objective function is to maximize profits by calculating revenues per flight and reducing the operation cost per flight. The profit is calculated by summing the revenue from fares - operating costs per route:

$$Maximize \sum X_r \cdot R_r - (C_r / \mu_Q)$$

Where R_r= fare revenue for route r, C_r = operating cost of route r, and μ_Q representing the average flight capacity for aircraft.

4.2.5 Constraints

The constraints were set up to define the network of routes. The hub nodes were considered supply nodes and the connecting cities were demand nodes. This was required to get the network flow.

1. Supply from MDW and DAL - 6000 passengers each
2. Demand from cities - [DEN, HOU, LAS, OAK, BOS, LAX, BWI] as [1800, 1900, 2400, 1000, 2100, 2200, 2300] respectively. These values were a proportionate subset of the actual demand that was collected per destination.
3. Daily limit capacity per route per airport

4. Minimum limit of one flight for each route

The implementation was conducted in Python using Pulp's LpSolve package and using LpSum function from the LpMaximize module. We found an optimal solution based on the assumptions, routing setup (network), and the constraints defined above, which is explained further in results (section 5.1).

4.3.1 Flight Scheduling

With a solution to determine worthwhile flight routes to offer, we directed our attention to developing a model that could optimize flight schedules in a way that satisfies demand while minimizing operating costs. In particular, we were interested in building a mixed-integer programming model capable of ingesting a dataset of predicted demand data for a given day and determining the most optimal flights to operate at specified times in order to route passengers to their final destinations through either direct or one-stop connecting flights. Building a linear programming model for this problem is complex, so we decided to make some assumptions to simplify our problem.

4.3.2 Assumptions

Since we were building a scheduling model assumed to be used months before the day of flights under consideration, we were not concerned with aircraft routing. We believed that this problem was a separate routing problem that would be better solved closer to the operating day when more precise demand, weather conditions, and other operating conditions were known. Furthermore, accounting for aircraft routing would have introduced substantially more decision variables that would have made this problem far more difficult to solve in a reasonable amount of time. For these reasons, we deemed aircraft routing as beyond the scope of focus and assume that there will be an aircraft available for use at each scheduled time. Our approach thus did not intentionally take into account the complexities of multi-leg routing, aircraft rotation, crew scheduling, and maintenance requirements. Since we are not routing specific aircraft with potentially varying passenger capacity, we assume a homogeneous fleet - aircraft are considered to have the same capacity, operating costs, and flight durations for the designated flight. In this implementation, aircraft capacity is fixed at a maximum of 180 passengers.

We simplify "satisfying demand" by assuming that we meet demand if we manage to route all of our passengers from their origin to their destination by the end of the day. For this reason, we decided not to introduce demand constraints for various times of the day, but rather for the day as a whole (D_{ik}). While we had numbers for the average number of total departures from each airport per day, we did not have specific information on the demand from each origin to destination. For this reason, we decided to generate our demand matrix (D) randomly. Flight times are assumed to be fixed and known, ignoring potential variability due to weather, air traffic control, or other operational factors. We assume that the operating cost for a flight for each combination of origin i to destination k is known and fixed, ignoring the number of passengers aboard, variability in fuel prices, airport fees, maintenance fees, and other expenses.

Due to the temporal quality present in this problem, we simplify our problem by restricting flight departures times (t) to occur at discrete time intervals. In particular, we allow all potential plane departures to fly at the beginning of each hour from 7:00 AM to 9:00 PM local time. All flight times are expressed as integers in military time in Eastern Standard Time or EST (UTC-5). Since PDT (UTC-8) is three hours behind Eastern Standard Time, 9:00 PM local time would be 24:00 EST (00:00). Therefore, while airports in the EST will only have a possible range of t from $\{7,\ 8,\ ...,\ 21\}$) the range of all possible departure times are $\{7,\ 8,\ ...,\ 24\}$. We represent this range as $H\ =\ \{H_{min}, H_{min}\ +\ 1,\ ..., H_{max}\}$, where H_{min} and H_{max} refer to the earliest and latest possible departure time in EST respectively.

Since flight scheduling is complicated by the possibility of connecting flights, we assume that all passengers must either fly directly from origin (i) to destination (k) or connect via *one* stop (j) in which a minimum layover and a maximum layover period are observed as previously defined. The minimum layover parameter allows us to constrain the problem to enable enough time for passengers to make their connections should small delays and plane transfers occur. The maximum layover parameter allows for constraining the solution such that connecting passengers do not endure excessive layovers. To model the effect of "hub cities", we also allowed for the definition of the maximum number of flights per day for each unique flight route originating from a given airport (F_i).

4.3.3 Decision Variables

Three types of decision variables are needed for modeling the problem. The first variable X_{ikt} represents the number of passengers traveling directly from airport i to airport k at time t, where a direct flight exists and all ts in a set of times H ranging from H_{min} to H_{max}. X_{ikt} must also be a non-negative integer that does not exceed the maximum aircraft capacity (Q_{max}):

$$X_{ikt}\ \in\ 0\ \leq\ \mathbb{Z}\ \leq Q_{max}, \forall\ (i, k)\ \in\ Direct\ Flights,\ t\ \in\{H_{min}, H_{min}\ +\ 1,\ ..., H_{max}\}$$

The second variable Y_{ikt} is a binary variable indicating whether a flight is operated (1) or not (0) from airport i to airport k at time t. It is defined for all pairs of airports $(i,\ k)$ where a direct flight is possible and for all times t within the same specified set as above:

$$Y_{ikt} \in \{0, 1\},\ \forall\ (i, k)\ \in\ Direct\ Flights,\ t\ \in\ \{H_{min}, H_{min}\ +\ 1,\ ..., H_{max}\}$$

The third variable $X_{conn_{ijkt_1t_2}}$ represents number of connecting passengers who travel from airport i to airport j at time t_1 and then from airport j to airport k at time t_2. This variable is defined for all possible one-stop routes where a flight from i to j at time t_1 and from j to k at t_2 is feasible (also considering travel time from i to j (T_{ij}), minimum and maximum layover times (L_{min}, L_{max})). The variable must be a non-negative and not greater than the maximum capacity of a single aircraft (Q_{max}):

$$X_{conn_{ijkt_1t_2}}\ \in\ 0\ \leq\ \mathbb{Z}\ \leq Q_{max},\ \forall (i, j, k)\ \in\ OneStopRoutes,\ t_1,\ t_2 \in H$$
$$where\ t_2\ \geq\ t_1 + T_{ij} + L_{min}\ and\ t_2 \leq\ t_1 + T_{ij} + L_{max}$$

4.3.4 Objective

The objective is to minimize operating costs given a set of possible flights ($F = \{(i, j, t) | \; i, j \in Airports, \; t \in H\}$ and operating costs for each possible flight (C_{ikt}):

$$Minimize \sum_{(i, k, t) \in F} C_{ikt} \cdot Y_{ikt}$$

4.3.5 Constraints

The model was subject to the following constraints:

1. Do not exceed the daily maximum quantity of flights from departing airport for each destination (F_{max_i}):

   $$\sum_{t=H_{min}}^{H_{max}} Y_{ikt} \leq F_{max_i}, \; \forall i, k$$

2. Satisfy demand (D_{ik}) for reaching destination from each departing airport:

   $$\sum_{t=H_{min}}^{H_{max}} X_{ikt} + \sum_{j \neq i,k} \sum_{t_1 = H_{min}}^{H_{max}} \sum_{t_2 = H_{min}}^{H_{max}} X_{conn_{ijkt_1t_2}} \geq D_{ik}, \; \forall i, k$$

3. The total number of passengers (direct and connecting) on any flight cannot exceed the plane's capacity Q_{max} (180 in our implementation):

   $$X_{ijt} + \sum_{k \neq i,j} X_{conn_{ijktt_2}} + \sum_{l \neq i,j} X_{conn_{lijt_1t}} \leq Q_{max} \cdot Y_{ikt}, \; \forall i, j, t$$

4. A flight should not be operated if the total number of passengers (direct and connecting) on any flight is below the plane's minimum capacity to operate $Q_{operate}$ (60 in our implementation):

   $$X_{ijt} + \sum_{k \neq i,j} X_{conn_{ijktt_2}} + \sum_{l \neq i,j} X_{conn_{lijt_1t}} \geq Q_{min} \cdot Y_{ikt}, \; \forall i, j, t$$

Constraint (2) integrates direct (X_{ikt}) and connecting ($X_{conn_{ijkt_1t_2}}$) flight options in meeting overall passenger demand. Constraints (3) and (4) do more work than only ensuring that capacities are respected – the use of Y_{ikt} together with X_{ikt}, and $X_{conn_{ijkt_1t_2}}$ variables link the variables. This prevents a flight from being operated if the number of passengers doesn't meet the constraints while forcing the flight to be operated if the number of passengers meets the constraints.

The constraint ensuring that a connecting flight is feasible was not programmed into the model explicitly – a custom function was written to iterate over all of the combinations of one-stop connecting flights and filter out the infeasible connecting itineraries. The filtered connecting itineraries were thus the only connecting itineraries considered by the model and eliminated the need for the model to consider the infeasible connecting itineraries.

The implementation was conducted in Python using Google OR-Tools' SCIP solver. To demonstrate the effectiveness of our model, we generated random demand data between airports and used the previously collected data for flight durations, and operating costs as inputs to help define the model.

Computational Experiment and Results

5.1.1 Determining Profitable Routes

Table 1: Parameters provided to the Flight Routing Model

Parameter	Value
Average Flight Capacity	180
Daily Flight Limit Per Route	8
Airport Capacity Limit per route	<= Average Flight Capacity * Daily Flight Limit, per route
Hub cities	2
Destination cities	7

5.1.2 Solution:

We obtained an optimal solution for the routing problem. *Table 2* demand constraints fulfilled for three of the cities.

Table 2: Demand fulfillment for the three of the cities

City	Demand (<=)	Output
Denver	1800	1800
Oakland	1000	1000
Boston	2100	2100

5.1.3 Objective function

The objective function (*Figure 1*) shows the coefficients calculated by the LpSolve module after subtracting the operating costs from the total revenue, and normalizing it per passenger. If the coefficients were negative, it would mean the normalized operating costs per passenger were higher for the route. The were potentially loss making routes which could be addressed by either increasing the airplane capacity or increasing the fare for that route. This could be a key insight for the airline.

Figure 1. Objective Function

10.67*flyers_BWI_DAL + 32.3*flyers_BWI_MDW + 18.54*flyers_DAL_BWI + 15.61*flyers_DAL_DEN + 34.18*flyers_DAL_HOU + 26.57*flyers_DAL_LAS + 29.98*flyers_DAL_LAX + 20.81*flyers_DAL_OAK - 3.04*flyers_HOU_DAL + 17.08*flyers_HOU_OAK + 13.09*flyers_LAS_HOU - 3.6*flyers_LAX_BWI + 10.09*flyers_LAX_DAL + 15.46*flyers_MDW_BOS - 3.31*flyers_MDW_BWI + 28.79*flyers_MDW_DEN + 31.48*flyers_MDW_LAS + 37.15*flyers_MDW_LAX + 20.51*flyers_MDW_OAK

5.1.4 Results

The experiment conducted on the routing model required setting up two hub nodes, seven source and destination cities, a network of routes interconnecting the cities, and a given set of operating constraints. The model returned an optimal solution in which eighteen out of the twenty-one routes were profitable, and the total operating profit came out to be $441,188.80. In terms of other interesting statistics, the total passengers flown were 54,000 and 300 flights were taken. The routes that were found unprofitable based on the existing fares and operating costs for those routes were Houston (HOU) to Dallas (DAL), Midway (MDW) to Baltimore (BWI), and Los Angeles (LAX) to Baltimore (BWI). We experimented by increasing the airline average capacity per aircraft from 180 to 190 and reran the model which immediately made the routes profitable. Thus, it could be interesting for the airline to consider the benefit of larger planes for those routes or adjusting airfares to make the routes profitable.

5.2.1 Flight Scheduling

As a proof of concept, our team attempted to use our model to build a schedule to route our randomly generated demand while optimizing for operating costs. While the data collected included data for 15 cities, we decided to test out our model by routing between just three airports (BOS, BWI, DAL). The matrices for the flight duration, operating costs, and randomly-generated demand data are available in *Appendix A*. In these matrices, each row corresponds to an origin and each column represents a destination.

With the parameters defined in *Appendix A*, our model yielded 166 decision variables and 130 constraints. It took about a second to solve, and produced the output in *Figure 2*:

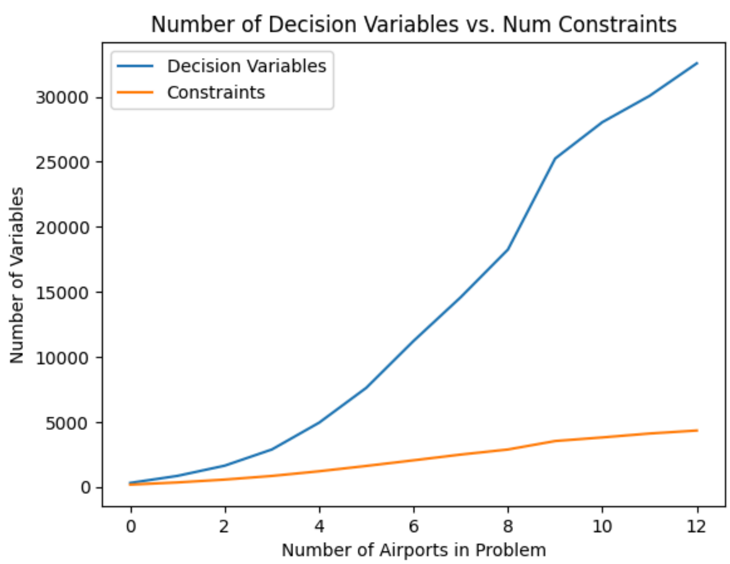
Figure 2: Initial Experiment Results

```
Decision Variables: 166
Constraints: 130
Optimal Solution Found:
==========================================================================================================================
Flight Schedule:
--------------------------------------------------------------------------------------------------------------------------
Flight BOS to BWI at 07:00 EST is operated, carrying 121 passengers ( 22 direct,  99 first-leg connectors,   0 second-leg connectors).
Flight BWI to DAL at 07:00 EST is operated, carrying 180 passengers (180 direct,   0 first-leg connectors,   0 second-leg connectors).
Flight BOS to BWI at 08:00 EST is operated, carrying 180 passengers (  0 direct, 180 first-leg connectors,   0 second-leg connectors).
Flight DAL to BWI at 09:00 EST is operated, carrying 180 passengers (  0 direct, 180 first-leg connectors,   0 second-leg connectors).
Flight BWI to BOS at 10:00 EST is operated, carrying 180 passengers (180 direct,   0 first-leg connectors,   0 second-leg connectors).
Flight DAL to BWI at 10:00 EST is operated, carrying 180 passengers (154 direct,  26 first-leg connectors,   0 second-leg connectors).
Flight DAL to BWI at 11:00 EST is operated, carrying 180 passengers (180 direct,   0 first-leg connectors,   0 second-leg connectors).
Flight BOS to BWI at 16:00 EST is operated, carrying 180 passengers (180 direct,   0 first-leg connectors,   0 second-leg connectors).
Flight BWI to DAL at 17:00 EST is operated, carrying 180 passengers (  0 direct,   0 first-leg connectors, 180 second-leg connectors).
Flight BWI to BOS at 20:00 EST is operated, carrying 180 passengers (154 direct,   0 first-leg connectors,  26 second-leg connectors).
Flight BWI to DAL at 20:00 EST is operated, carrying 180 passengers ( 81 direct,   0 first-leg connectors,  99 second-leg connectors).
Flight BWI to BOS at 21:00 EST is operated, carrying 180 passengers (  0 direct,   0 first-leg connectors, 180 second-leg connectors).
==========================================================================================================================
Connecting Itineraries:
--------------------------------------------------------------------------------------------------------------------------
Connecting flights from BOS to BWI at 07:00 EST and BWI to DAL at 20:00 EST are operated, carrying 99 connecting passengers.
Connecting flights from BOS to BWI at 08:00 EST and BWI to DAL at 17:00 EST are operated, carrying 180 connecting passengers.
Connecting flights from DAL to BWI at 09:00 EST and BWI to BOS at 20:00 EST are operated, carrying 26 connecting passengers.
Connecting flights from DAL to BWI at 09:00 EST and BWI to BOS at 21:00 EST are operated, carrying 154 connecting passengers.
Connecting flights from DAL to BWI at 10:00 EST and BWI to BOS at 21:00 EST are operated, carrying 26 connecting passengers.
==========================================================================================================================
Operating Cost:
--------------------------------------------------------------------------------------------------------------------------
$160,932.00
==========================================================================================================================
CPU times: user 1.11 s, sys: 10.9 ms, total: 1.13 s
Wall time: 1.17 s
```

Looking at *Figure 2*, we managed to obtain an optimal solution for routing passengers to their final destinations between BOS, BWI, and DAL. The output shows the flight schedule with the total number of passengers aboard each flight. It also displays how many of each type of passengers are on the flight, verifying that our constraints of maximum and minimum aircraft capacity are respected and the demand is met.

We also attempted to increase the complexity of our problem by adding in other airports into our problem. While the initial experimental model executed in one second, adding a fourth airport took several hours to execute until we gave up and terminated the code early. *Figure 3*. plots the number of decision variables vs constraints of the models intended to solve the scheduling problem for an increasing number of airports.

Figure 3. A Comparison of Decision Variables and Constraints with Increased Complexity



As the number of airports increase, the number of decision variables seemingly increases exponentially while the number of constraints increases in a more linear fashion. This increasing divide between decision variables and constraints means that the solution space grows rapidly as the complexity increases. The growing solution space originating from even a slightly more complex problem results in a sizable increase in computational resources needed to solve the problem. A model that factors in the vastly more complex realities of airline operations would likely be impossible under the current formulation to solve to optimality.

Future Work

The routing problem was a huge network modeling exercise with a total 210 routes, and forced us to quickly consider only a subset of the cities to make the problem manageable and yield an optimal solution. Each route is a directional arc connecting the source with the destination node, and therefore the connection back from the destination to the source city would be a new arc or route. Scaling up the network model to incorporate all the cities and setting up the routes would require programmatic handling and will deal with large amounts of complexity. The assumptions we had to make in order to simplify the model, such as assuming the fixed airplane capacity and fares, could be further expanded to take into account additional variations. Dynamic programming could be used to further stimulate a daily operation, or run seasonal demand scenarios, to examine how the model behaves under certain *stress* conditions.

For the scheduling problem, one might approach the massive solution space with algorithms that yield *near optimal results*. Such algorithms include heuristic methods like greedy algorithms, where we begin fulfilling the highest demand routes with the lowest operating costs, and gradually moving to less optimal routes. We might also consider applying simulated annealing, genetic, or other evolutionary algorithms. Experimenting with these algorithms could provide “good-enough” solutions for the more complex problems through iterative exploration of the solution space. If these methods are deemed effective, we might explore relaxing some of our assumptions to model more realistic operational considerations such as changing daily demand, fleet routing, crew routing, continuous flight schedules, and multi-stop connecting flights. Running sensitivity analysis could also be useful to help further identify key constraints and decision variables that are impacting the model.

Conclusion

This study aimed to optimize flight scheduling and routing for airlines, with a focus on maximizing profitability and operational efficiency. By developing two distinct models - a linear programming model to identify profitable routes and a mixed-integer programming model for flight scheduling optimization - we addressed critical challenges in airline operations. The models, which took into account fare costs, operating expenses, and market demand, demonstrated the potential for significant improvements in routing and scheduling strategies.

Our computational experiments, based on data from Southwest Airlines and the U.S. Department of Transportation, highlighted the models' effectiveness in identifying optimal solutions for the complex problem of airline scheduling. However, the research also underscored the increasing computational challenges associated with larger-scale problems, suggesting a pivotal role for heuristic algorithms in navigating the vast solution spaces of more complex scenarios.

Future research directions are clear: there is a substantial need to explore heuristic and metaheuristic approaches that can efficiently address larger-scale airline scheduling problems while accommodating more realistic operational factors.

This study contributes to the body of knowledge in airline operations optimization by providing a framework for airline profitability and efficiency enhancement through advanced programming models. It paves the way for further exploration of sophisticated algorithms capable of tackling the inherent complexities of airline scheduling and routing problems, with the ultimate goal of sustaining airline competitiveness and operational excellence in a rapidly evolving market landscape.

By integrating insights from successful case studies like UPS's optimization efforts and applying a systematic approach to data collection and model formulation, we have laid the groundwork for future advancements in airline operations optimization. Our research offers valuable methodologies and insights for airlines seeking to optimize their routing and scheduling strategies in pursuit of heightened operational efficiencies and profitability.

References

Armacost, Andrew P., Cynthia Barnhart, Keith A. Ware, and Alysia M. Wilson. "UPS Optimizes Its Air Network." *Interfaces* 34, no. 1 (February 2004): 15–25. https://doi.org/10.1287/inte.1030.0060.

"Breaking Down Southwest Airlines' Flight Operating Costs." n.d. Trefis. Accessed March 10, 2024.https://www.trefis.com/no-login-required/Ci4zhXAj?fromforbesandarticle=a-closer-look-at-southwests-per-flight-costs.

"Bureau of Transportation Statistics." n.d. U. S. Department of Transportation. Accessed March 10, 2024. https://www.transtats.bts.gov/Data_Elements.aspx?Data=4.

Deswal, Ashish. "Breaking down Southwest Airlines' Flight Operating Costs." Trefis. Accessed March 10, 2024. https://www.trefis.com/no-login-required/Ci4zhXAj?fromforbesandarticle=a-closer-look-at-southwests-per-flight-costs.

"Southwest." n.d. Southwest. Accessed March 10, 2024. https://www.southwest.com/air/low-fare-calendar/?slp=SALEPAID&departureDate=2024-04-01&returnDate=2024-04-04&s_tnt=1035431%3A0%3A0&adobe_mc_sdid=SDID%3D2CD4840D1704EC1D-398B35F2411B2039%7CMCORGID%3D65D316D751E563EC0A490D4C%40AdobeOrg%7CTS%3D1710126275&adob.

Subramanian, Radhika, Richard P. Scheff, John D. Quillinan, Steve Wiper, and Roy E. Marsten. 1994. "Coldstart: Fleet Assignment at Delta Air Lines." *Interfaces* 24, no. 1 (Jan): 104-120.

Appendix A: Sample Data and Parameters Used in Initial Experiment

Flight duration in minutes (T)

	BOS	BWI	DAL
BOS	0	100	0
BWI	90	0	215
DAL	0	165	0

Randomly-generated demand in number of passengers (D)

	BOS	BWI	DAL
BOS	0	202	279
BWI	192	0	114
DAL	206	171	0

Operating cost in USD (C)

	BOS	BWI	DAL
BOS	0	8291	0
BWI	8855	0	18415
DAL	0	18083	0

Parameter	**Value**
Minimum Layover Time in Minutes (L_{min})	60
Maximum Layover Time in Minutes (L_{max})	420
Plane Capacity (Q_{max})	180
Minimum Passengers to Operate Flight (Q_{min})	60
Number of Flights to Each Destination (F_{max}	{'default': 10, 'BWI': 15, 'MDW': 15, 'DAL': 12, 'DEN': 12, 'LAS': 15}
Earliest Departure in EST (H_{min})	7
Latest Departure in EST (H_{max})	24
Cities	Boston (BOS), Baltimore (BWI), Dallas (DAL)