Integrated Regional Airline Scheduling Via Column Generation

Alberto Santini[1] and Vikrant Vaze[2]

[1]Universitat Pompeu Fabra
[2]Thayer School of Engineering at Dartmouth

Keywords: Timetable development; fleet assignment; passenger choice; column generation

1 INTRODUCTION AND CONTRIBUTIONS

Flight schedules have a significant effect on airline profitability and competitiveness. A large number of airlines are classified as regional airlines. They are responsible for operating between 30% and 40% of the flights in the US domestic market. For example, in 2022, four of the 10 largest US airlines in terms of flight operations were regional airlines. They typically offer short-haul flights using smaller aircraft, leverage hub-and-spoke networks, and help passengers connect with flights operated by their mainline partner carriers. We jointly optimize frequency planning, timetable development, and fleet assignment, as well as some limited aspects of route planning, for a regional airline network, assuming that the mainline partner's schedule is fixed.

Our first contribution is a new integrated optimization formulation for medium-term planning decisions that incorporates a realistic passenger demand model. Unlike previous studies (Wei *et al.*, 2020, Yan *et al.*, 2022), which used flight-level decision variables, we use a single composite variable to model all non-stop flights between a pair of airports and their complex interdependencies. In contrast, the existing airline scheduling literature uses composite variables to model the itineraries of individual aircraft, crew, and passengers primarily for short-term planning. Second, the composite variables provide an extended formulation enjoying an extremely tight continuous relaxation. We exploit this tightness to obtain near-optimal solutions via column generation (CG) and a restricted master heuristic. We also propose a new acceleration approach based on Dynamic Programming (DP) to quickly generate promising columns. When combined with implicit dual smoothing, symmetry breaking, and subproblem aging, this acceleration approach allows us to solve large-scale real-world instances in 3 hours. Finally, through extensive computational analyses based on ablation studies, we demonstrate the effectiveness of our overall modeling and computational framework. We identify the main operational drivers of the profit improvements enabled by our approach. Furthermore, numerous sensitivity analyses confirm that our results are robust to relaxing key modeling assumptions. Ultimately, the proposed approach provides high-quality solutions to real-world instances within practically reasonable runtimes.

Recent studies develop timetabling and fleet assignment models that take into account the implications of choice-based demand. They either simplify the choice model or rely on heuristic solution approaches. In contrast, our approach finds provably near-optimal solutions, and it can also incorporate frequency planning and some limited aspects of route selection. Existing scheduling studies that use CG do so to generate either aircraft routes or passenger itineraries. In contrast, our approach uses CG to generate segment schedules, which is a new application of CG in airline scheduling. Finally, we develop and solve a new passenger mix model using DP to significantly speed up the pricing subproblem.

2 METHODOLOGY

2.1 Single-Segment Scheduling

We start by formulating the single-segment scheduling problem under a General Attraction Model (GAM) passenger allocation (Gallego *et al.*, 2015). The binary variable $\boldsymbol{z}$ indicates whether the schedule includes a flight departing in a specific period and operated by a specific aircraft type. The variables $\boldsymbol{x}$ and $\boldsymbol{x}_0$ are the number of passengers in a specific market who take, respectively, a flight at a specific time, or the outside option (OO). The objective function (1a) maximizes the profit. Constraint (1b) ensures that the segment schedule does not use more aircraft than available in any time window whose size equals the round-trip segment travel time of that aircraft type. (1c) defines the demand for each flight as proportional to its attractiveness. (1d) ensures that the sum of the number of passengers taking the various alternatives equals the total passenger demand in the market. (1e) ensures that the offered seats are enough to carry all passengers taking the flight. AC denotes aircraft, ACT aircraft types, SS segment schedules, and PAX stands for passengers.

$$\max_{\boldsymbol{x},\boldsymbol{z}} \sum_{\text{Aircraft-Types}} \sum_{\text{Time-Periods}} \sum_{\text{Markets}} (\text{Fare} \times \text{PAX} - \text{Cost} \times \text{Number of Flights}) \tag{1a}$$

$$\begin{aligned}
\text{s.t.} \quad & \text{Flights in Round-Trip Window} \leq \text{Available Aircraft} && \forall \text{ ACT, Periods} && \text{(1b)}\\
& \text{PAX Carried} \leq \frac{\text{Itinerary Attractiveness}}{\text{OO Attractiveness}} \times \text{OO Pax} && \forall \text{ Periods, Markets} && \text{(1c)}\\
& \text{OO PAX} + \sum_{\text{Periods}} \text{PAX Carried} = \text{Demand} && \forall \text{ Markets} && \text{(1d)}\\
& \text{PAX} \leq \text{Available Seats} && \forall \text{ Periods} && \text{(1e)}\\
& \boldsymbol{z} \text{ binary}, \boldsymbol{x}, \boldsymbol{x}_0 \geq 0 && \forall \text{ ACT, Markets, Periods} && \text{(1f)}
\end{aligned}$$

2.2 Network-Wide Scheduling

The network-wide schedule optimization problem combines individual single-segment problems and links them through fleet size and aircraft balance constraints. Binary decision variables $\boldsymbol{y}$ indicate whether the regional airline operates a specific segment schedule on a specific segment. Decision variables $\boldsymbol{w}$ represent the number of aircraft of a specific type available at a specific airport at the beginning of the planning period. The objective function (2a) maximizes the total profit from the set of chosen segment schedules. Constraints (2b) ensure that exactly one segment schedule is selected for each segment operated by the regional airline. (2c) ensures that no more aircraft of each type are used than are available. (2d) links the variables $\boldsymbol{y}$ and $\boldsymbol{w}$ that impose the conservation of flow at each airport in each period and for each type of aircraft. (2e) ensures that the aircraft are correctly positioned at the end of the planning horizon to ensure that the same schedule can be repeated after the end of the planning horizon. We note that for a given segment, the number of $\boldsymbol{y}$ variables is exponential in the number of periods and also exponential in the number of aircraft types. Furthermore, in the worst case, the number of feasible solutions of (2a)–(2f) is exponential in the number of segments.

$$\max_{\boldsymbol{y},\boldsymbol{w}} \sum_{\text{Segment Schedules}} \text{Schedule Profit} \tag{2a}$$

$$\begin{aligned}
\text{s.t.} \quad & \text{Number of Chosen Schedules} = 1 && \forall \text{ Segments} && \text{(2b)}\\
& \text{AC at Horizon Start} \leq \text{Available AC} && \forall \text{ACT} && \text{(2c)}\\
& \text{Remaining AC} = \text{Starting} - \text{Departing} + \text{Arriving} && \forall \text{ Airports, ACT, Periods} && \text{(2d)}\\
& \text{AC at Horizon Start} = \text{AC at Horizon End} && \forall \text{ Airports, ACT} && \text{(2e)}\\
& \boldsymbol{y} \text{ Binary}, \boldsymbol{w} \text{ Non-Negative Integers} && \forall \text{ SS, ACT, Airports, Periods} && \text{(2f)}
\end{aligned}$$

Table 1 – *Ablation study results. All gaps are calculated w.r.t. the true dual bound.*

	Base (All Activated)		**No Barrier**		**No Shortcut**		**No Noise**		**No DP**	
Instance	Gap %	Time (s)	Gap %	Time (s)	Gap %	Time (s)	Gap %	Time (s)	Gap %	Time (s)
Mesa Airlines (YV)	0.11	52.80	0.09	4602.45	0.11	277.76	0.10	1042.75	0.10	703.75
Republic Airways (YX)	0.20	884.05	0.20	1506.19	0.20	1696.42	0.20	2541.12	0.20	1883.55
Skywest Airlines (OO)	0.00	10800.00	0.00	10800.00	0.01	10800.00	—	10800.00	0.00	10800.00
YV + YX	0.17	1115.75	0.14	1742.86	0.19	1666.04	0.14	8481.55	0.13	6051.81
YV + OO	0.08	10800.00	—	10800.00	0.11	10800.00	—	10800.00	—	10800.00
YX + OO	0.12	10800.00	—	10800.00	—	10800.00	—	10800.00	—	10800.00
YV + YX + OO	0.09	10800.00	—	10800.00	—	10800.00	—	10800.00	—	10800.00

2.3 Column Generation Algorithm

The schedule for any given segment that appears in an optimal network-wide solution is not necessarily optimal for the single-segment problem and vice versa. Therefore, one cannot directly use the single-segment model to build the set of segments required to solve the network-level model. Enumerating all feasible schedules would be prohibitively expensive, due to the size of the set. Moreover, most feasible schedules are not attractive from a commercial point of view. These observations suggest that an appropriate solution method for the network-wide problem should use the single-segment model to identify promising schedules. At the same time, information from the network-wide model should be used at the single-segment level to ensure that the resulting schedules work well when used jointly. To achieve the above objective, we propose a column generation algorithm that uses dual information from the continuous relaxation of (2a)–(2f) to generate promising schedules using a variation of the single-segment model (1a)–(1f). Formulation (2a)–(2f) is challenging to solve due to the exponentially many variables $\boldsymbol{y}$ for each segment. However, this formulation enjoys an especially tight continuous relaxation. Furthermore, we develop an exact column generation procedure and a novel acceleration technique, that efficiently solve the continuous relaxation. Ultimately, a tight continuous relaxation and an efficient column generation procedure together yield high-quality solutions within short runtime budgets (of at most three hours) for planning problems at the scale of real-world airline networks.

2.4 Heuristic Column Generation

In principle, there is no need to solve the pricing subproblem to optimality, and any feasible solution with positive reduced cost, when added to the column pool, can improve the restricted master problem (RMP) objective. However, we prove that, if we used a suboptimal solution of the pricing subproblem to build a segment schedule as a new column in the RMP, its objective coefficient would be wrong, because it would be calculated using passenger figures that do not follow the GAM. To compute the objective coefficient for the network-wide scheduling model with GAM-consistent passenger allocation for a possibly suboptimal solution, one can fix the value of the $\boldsymbol{z}$ variables and re-solve the pricing subproblem to optimality (we call such a procedure "z-fixing"). Indeed, such a transition from a feasible solution that does not follow the GAM to one that does, by definition, cannot decrease the reduced cost.

We develop a novel approach to derive an upper bound on the reduced cost of a column by solving a relaxation of the pricing subproblem. This procedure can result in two cases. On the one hand, if the upper bound is negative, then no positive-reduced-cost columns exist. On the other hand, while producing the bound, the procedure usually builds a segment schedule that can be added to the column pool of the RMP. We solve this relaxed version of the pricing subproblem using DP. We find that DP usually produces high-quality columns and helps to improve the overall convergence of the algorithm.

Table 2 – *Results overview when optimizing the networks of Mesa (YV), Republic (YX), or both.*

	YV			YX			YV+YX		
	Revenue	**Cost**	**Profit**	**Revenue**	**Cost**	**Profit**	**Revenue**	**Cost**	**Profit**
Status Quo	1,664,272	696,246	968,026	1,723,135	839,696	883,439	3,387,407	1,535,941	1,851,465
Freq ±0	1,698,251 (+2.04%)	696,348 (+0.01%)	1,001,903 (+3.50%)	1,762,283 (+2.27%)	842,073 (+0.28%)	920,211 (+4.16%)	3,471,732 (+2.49%)	1,538,942 (+0.20%)	1,932,790 (+4.39%)
Freq ±1	1,805,247 (+8.47%)	742,565 (+6.65%)	1,062,682 (+9.78%)	1,869,700 (+8.51%)	924,730 (+10.13%)	944,970 (+6.96%)	3,713,894 (+9.64%)	1,670,419 (+8.76%)	2,043,474 (+10.37%)
Freq ±2	1,841,328 (+10.64%)	751,723 (+7.97%)	1,089,605 (+12.56%)	1,933,830 (+12.23%)	936,473 (+11.53%)	997,357 (+12.89%)	3,790,470 (+11.90%)	1,692,362 (+10.18%)	2,098,108 (+13.32%)

3 RESULTS AND DISCUSSION

Table 1 quantifies the impact of the main algorithm improvements: 1) Barrier: Using the barrier algorithm when solving the relaxed RMP to increase the likelihood that the optimal dual solution lies midface. 2) Shortcut: Cutting short some column generation iterations when the likelihood of producing new positive reduced cost solutions is low. 3) Noise: Adding a tiny amount of noise to the fares to break the symmetry in the subproblem. 4) DP: Using the DP algorithm and solving MIP (1a)–(1f) only if the DP does not produce any positive-reduced-cost column. In the "Base" configuration, all algorithm improvements are active. In the other four configurations, we disable each improvement to evaluate how much the performance of the corresponding algorithm differs compared to the Base configuration. We used Gurobi 9.0.0 as the black-box linear and mixed-integer optimization solver, and a time limit of 3 hours. Table 1 shows that each algorithmic improvement typically leads to a 2X-10X acceleration. In fact, the last two instances are not solvable until all four improvements are used, showing the effectiveness of our algorithm design.

Table 2 presents the effects of our modeling and algorithmic approach on the key financial metrics. First, to evaluate the "Status Quo" schedule, we fix all flights to their real-world departure times and operating aircraft. In this case, the algorithm reduces to applying the passenger allocation model by solving the subproblem (1a)–(1f) for all segments. The other three configurations are named "Freq $\pm$ t" and allow the flight frequency to vary by at most t flights on each segment. In particular, "Freq $\pm$ 0" only allows flight retiming and fleet reassignment, but not frequency change. When $t \in \{1, 2\}$, we still impose a minimum frequency of one flight per day to ensure that no destination is completely excluded. Table 2 provides several insights. First, the costs and revenues computed by our approach are both higher than those under the Status Quo in all cases. Furthermore, in each case, a profit higher than the Status Quo is obtained because the absolute revenue growth is greater than the absolute cost increase. As expected, in the Freq ±0 case, with the frequencies not allowed to change in any of the segments, the entire cost growth is driven only by the reassignment of aircraft corresponding to different operating costs. Consequently, the corresponding cost changes are small (in the 0.01%-0.28% range). On the other hand, revenue increases by 2.04%-2.49% leading to a profit increase of 3.50%-4.39%. These results demonstrate the potential for hundreds of thousands of dollars in daily profit improvements.

References

Gallego, Guillermo, Ratliff, Richard, & Shebalov, Sergey. 2015. A general attraction model and sales-based linear program for network revenue management under customer choice. *Operations Research*, **63**(1), 212–232.

Wei, Keji, Vaze, Vikrant, & Jacquillat, Alexandre. 2020. Airline Timetable Development and Fleet Assignment Incorporating Passenger Choice. *Transportation Science*, **54**(1), 139–163.

Yan, Chiwei, Barnhart, Cynthia, & Vaze, Vikrant. 2022. Choice-based airline schedule design and fleet assignment: A decomposition approach. *Transportation Science*, **56**(6), 1410–1431.