Re-schedule flights after disruption by simple approach based on column generation heuristic

WSDM – Sichuan Airlines - Intelligent Flight Schedules

Challenge 2019

Min Yang
Dalian Maritime University
Dalian Liaoning China
yangmin0325@dlmu.edu.cn

ABSTRACT

WSDM - Sichuan Airlines - Intelligent Flight Schedules Challenge is an integrated aircraft and passenger recovery problem. The problem is to re-schedule flights, re-assign aircraft and re-accommodate passengers, with minimized impact on flight operations when extreme weather or aircraft faults causes large-scale delays in flight at multi-base. A foggy weather at Shuangliu Airport made all the outbound and inbound flight be delayed by 2 hours, some adjustment methods could be used to alleviated the delay of subsequent flights. To tackle this problem, a simple approach based on column generation heuristic be proposed to get a reasonable solution within one minute, which consists of three phases: aircraft routes generation, routes selection and passenger assignment. Because of insufficient time and mistake I made, the result compares to the champion`s work in the competition is not well enough. In the end, the work to be done in the future to improve the performance be discussed. And my team name is Xiaoshazhude.

KEYWORDS

aircraft recover problem, routes selection, column generation

ACM Reference format:

Min Yang. 2018. Re-schedule flights after disruption by simple approach based on column generation heuristic: WSDM - Sichuan Airline Challenge 2019. In *Proceedings of WSDM conference*, *Australia*, *February 2019 (WSDM'19)*, 3pages.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The rapid increasement of passengers and routes make lots of profit for Sichuan Airlines Co.,Ltd, but it brings many challenges for daily operations at the same time. One of the most difficult troubles is disruption, which have a large financial impact on the airline industry. But it provides a favorable environment for the application of operations research (OR) models and techniques[1].

For this problem, there are many efficient algorithms has been proposed, and I received my inspiration from Liang, Z et al., 2018[2].

WSDM'19, February 2019, Australia

https://doi.org/10.1145/nnnnnnn.nnnnnnn

With some problem relaxation, I build an aircraft route selection model routes to select a routes combination for each aircraft with minimal recovery cost.

The report is organized as follows: Section 2 introduces the details of the challenge task, followed by two mathematical models to define the problem; In Section 3, the method about routes generation and routes selection be described. Finally, the conclusion and the future work based on complete column generation be discussed in Section 4.

2 PROBLEM DEFINITION

2.1 Data description

The challenge provides seven tables, I convert them into three tables: *flights*, *airports* and *aircrafts*. There are 2150 flights thorough four days recovery horizon and 132 aircrafts involved. The first knowledge from data analysis is that only 14 aircraft routes have been disrupted directly. The second is that the impact of 2 hours delay can be absorbed by reducing the turnaround time of flights on the first day. Therefore, I decided to only consider the flights on the first day and the aircrafts start from airport 57. In addition, the aircrafts start from airport 50 be joined as well in order to improve the probability of aircraft swap. Unfortunately, that`s why my result is worse than the champion`s result.

2.2 Problem Model

The master problem is a linear relaxation of aircraft route selection model (ARSM). The sets, parameters and decision variables definition based on Liang`s model.

Sets

- C the set of aircraft
- F the set of flights
- L the set of routes
- D the set of days

Parameters

- α_f the weighted cost if flight f is canceled
- $\beta_{c,l}$ the cost if aircraft c flies route l

$$\delta_{f,l} = \begin{cases} 1 & \text{if flight f belongs to route } l \\ 0 & \text{otherwise} \end{cases}$$

γ_l^d the number of turnaround time decreasing used on day d in route l

N^d the number of flights on day d

φ maximum proportion of turnaround time violation

ω maximum proportion of canceled flights

Decision variables

$$x_{c,l} = \begin{cases} 1 \text{ if aircraft } c \text{ flies route } l \\ 0 \text{ otherwise} \end{cases}$$

$$y_f = \begin{cases} 1 \text{ if flight } f \text{ is canceled} \\ 0 \text{ otherwise} \end{cases}$$

Objective function and constraints

$$\min \sum_{f\in F} \alpha_f y_f + \sum_{c\in C} \sum_{l\in L} \beta_{c,l} x_{c,l} \quad (1)$$

Subject to:

$$\sum_{c\in C} \sum_{l\in L} \delta_{f,l}\, x_{c,l} + y_f = 1,\ \ \forall f \in F \quad (2)$$

$$\sum_{l\in L} x_{c,l} \le 1,\ \ \forall c \in C \quad (3)$$

$$\sum_{c\in C} \sum_{l\in L} \gamma_l^d x_{c,l} \le \varphi N^d,\ \ \forall d \in D \quad (4)$$

$$\sum_{f\in F} \rho^d y_f \le \omega N^d,\ \ \forall d \in D \quad (5)$$

$$x_{c,l} \in \{0,1\},\ \ \forall c \in C, \forall l \in L \quad (6)$$

$$y_f \in \{0,1\},\ \ \forall f \in F \quad (7)$$

The objective function minimizes the weighted sum of flights cancellation cost and the sum of route cost by each aircraft. And the route cost includes the costs caused by flight delay, flight swap, aircraft base change, passenger delay or cancellation.

Constraint (2) ensures that each flight is either canceled, or operated by one route only once. Constraint (3) ensures that each aircraft can only selects one route to fly. Constraints (4) ensures that the number of flight turnaround time on that day is less than or equal to the 5 percent of total number of flights on that day. Constraints (5) ensures no more than the 10 percent of total flights would be canceled. Constraints (6) and (7) are the binary constraints for the decision variables.

After the determination of aircraft schedules, we need to re-accommodate the passengers whose flight had been canceled or the number of seats in the replacement aircraft is insufficient. For this problem, I build a passenger assignment model for each group of canceled passengers from airport a to airport b.

Sets

P^{ab} the set of canceled passengers from a to b

F^{ab} the set of flights from a to b with spare seats

Parameters

$\lambda_{p,f}$ the cost if one of passengers p is assigned to flight f

c_f the number of spare seats for flight f

t_f the actual departure time for flight f

c_p the initial departure time of passenger p

t_p the initial departure time of passenger p

μ_p the reward for re-accommodating one passenger in p

θ the maximum delay time for transferred passengers

Decision variables

$z_{p,f}$ the number of passengers from p be accommodated to flight f

Objective function and constraints

$$\min \sum_{p\in P^{ab}} \sum_{f\in F^{ab}} \lambda_{p,f} z_{p,f} - \mu_p \sum_{f\in F^{ab}} z_{p,f} \quad (1)$$

Subject to:

$$\sum_{p\in P^{ab}} z_{p,f} \le c_f,\ \ \forall f \in F^{ab} \quad (2)$$

$$\sum_{f\in F^{ab}} z_{p,f} \le c_p,\ \ \forall p \in P^{ab} \quad (3)$$

$$z_{p,f} = 0,\ \ \forall t_p + \theta > t_f \text{ or } t_p > t_f \quad (4)$$

$$z_{p,f} \in \text{Integer}, \forall p \in P^{ab}, \forall f \in F^{ab} \quad (5)$$

The objective function minimizes the cost of re-accommodating passengers. Constraint (2) and (3) ensures that assignment capacity is satisfied. Constraints (4) ensures the departure time of re-accommodated passengers cannot earlier or 48 hours later than the departure time of flights. Constraints (5) is the integer constraint for the decision variables. But the model has not been used in my simplified problem.

3 APPROACH

3.1 Build connection network

In order to generate route for an aircraft, we need to build specific connection network for each aircraft. There are three limitation for aircraft to flies a flight, 1) routes including crossing water tasks; 2) airports aircraft type constraint; 3) routes aircraft type constraint. They can be regard as flight aircraft type li-limitation. With above limitation, the size of connection network for each aircraft could be reduced, which will cut down the size of potential routes. See algorithm 3.1 for details.

Algorithm 3.1 Build connection network

```
Input: aircrafts C and flights F
Output: connection network for each aircraft
for each aircraft c in C do
    filter the flights for aircraft c
    connect all flights and get successor list for each flight
    topological sort the connection network
end for
return connection network for each aircraft
```

3.2 Build aircraft routes

The After the establishment of the connection network for each aircraft, we need to generate routes as many as possible. A multi-label route generation algorithm abstracted from Zhe.Liang`s basic multi-label shortest path algorithm be used to tackle the problem.

The biggest difference with Zhe.Liang`s problem is we have chance to reduce the initial turnaround time as long as it longer than the 2/3 of minimal turnaround time. Therefore, when extend a flight to next one, we have three choice for turnaround: 1) initial turnaround time; 2) minimal turnaround time; 3) less than minimal turnaround time but more than 2/3 of minimal turnaround time. Violation of minimal turnaround time will cause penalty, reduction of initial turnaround time will occupy the quota. Although the first choice is better, but it may delay the following flight. As a result, I add the attribute '*cut*'. If the flight on the route reduced the turnaround time, it will increase by 1. And a table of which length is same as route will record the minutes of turnaround reduction if the flight`s turnaround time less than initial one.

Algorithm 3.2.1 Routes generation

Input: aircraft c and graph G(V, A)

Output: all no-dominated routes of aircraft c

Initialize the source node's label set as $\{\langle \mathbf{0}, \mathbf{0}, \mathbf{0}, \{\} \rangle\}$ and the other nodes` label sets as empty.

for each node i of network G in topological order **do**

    for each node j that is a successor of node i **do**

        Process flight-flight arc (i, j).(see Algorithm 3.2.2）

    end for

end for

Got the labels of the nodes connecting to sink node.

Delete the labels be dominated by any other label.

return all routes recorded by those labels.

The Algorithm 3.2.1 show the framework of routes generation, and more details about label extends function are given in Algorithm 3.2.2. And this algorithm just considers the flights of date 1.

Algorithm 3.2.2 flight-flight arc (i, j) labels extend

for each label $< \overline{\boldsymbol{\beta_i}}, \boldsymbol{\beta_i^d}, \boldsymbol{\beta_i^{cut}}, \{\mathbf{cuts}\} >$ in node i`s label set $\mathbf{B_i}$ **do**

    flight to flight with initial *Turntime.*

    flight to flight with minimal *Turntime.*

    flight to flight with turnaround time less than minimal *Turntime*.

end for

3.3 Solve master problem

After the generation of routes, those routes and their cost as the input of aircraft route selection model mentioned in Section 2.2. Then the commercial solver CPLEX 12.7 be used to get a routes combination with lowest recovery cost. And the result is 91352.

4 CONCLUSION

In this report, I introduced the approach how to tackle the integrated airline recovery problem of the Intelligent Flight Schedules Challenge of the WSDM Cup 2019. My final result is ranked the second place on the leaderboard. There is a large gap between the champion`s schedule and mine, the main reason is my radical assumption to simplify the problem. But the result also shows the reliability of column generation heuristic, which will get a better performance after subsequent refinement. And the idea of Bisaillon et al.(2011)[3] to handle the passengers recovery will be referred too. I am glad to share my code solution after the work finished, please contact me if you have any question.

ACKNOWLEDGMENTS

I wish to thank Sichuan airline, Baidu Inc and WSDM Conference give me a chance to test my ability in practice. Besides, I would like to thank professor Liang, Z for his online open course and excellent paper. Furthermore, thank my mentor Sun Zhuo, friend Yan Yu and my parents to support me.

REFERENCES

[1] Barnhart, C., Belobaba, P., & Odoni, A. R.. (2003). Applications of operations research in the air transport industry. *Transportation Science, 37*(4), 368-391.

[2] Liang, Z., Xiao, F., Qian, X., Zhou, L., Jin, X., & Lu, X., et al. (2018). A column generation-based heuristic for aircraft recovery problem with airport capacity constraints and maintenance flexibility. *Transportation Research Part B: Methodological, 113*, 70-90.

[3] Bisaillon, S., Cordeau, J. F., Laporte, G., & Pasin, F.. (2011). A large neighborhood search heuristic for the aircraft and passenger recovery problem. *4OR: A Quarterly Journal of Operations Research, 9*(2), 139-157.