

Optimization Model and Algorithm for Irregular Flight Recovery

Quansheng Lei

Automation school. Beijing University of Posts and
Telecommunications
Beijing, China
e-mail: leiqshz@163.com

Peng Zhao

Automation school. Beijing University of Posts and
Telecommunications
Beijing, China
e-mail: zhaopengali@126.com

Abstract—Flight disruptions disturb airlines' routine, causing a lot of troubles to passengers as well as bringing on huge economic losses to airlines. So the irregular flight recovery is made to solve this problem. Irregular flight recovery is a large-scale and integer programming problem, which has complex variables and constraints. In this research, we model in terms of real situation and a column generation algorithm combined with Bellman-Ford algorithm is presented to release the influences resulted from disruptions. Computational experiments show that the proposed algorithm is feasible.

Keywords—flight disruptions; airlines; column generation; network optimization;

I. INTRODUCTION

Severe weather, crew sickness, airport curfews, aircraft mechanical problems and security are the common problems that force an airline to delay or even cancel their regular published flights. We customarily call the delayed or cancelled flights as irregular flights. Irregular flights always mean that passengers can't reach the destination on time and airlines will have to stand the loss due to delayed and cancelled flights. Generally, delaying or cancelling a flight should be costly for airlines. Thus irregular flight recovery should minimize the number of delayed or cancelled flights, and reassign aircraft to the flight so that loss of airlines can be cut down as much as possible.

Airline irregular operations are just airline operations under irregular situations, meaning that some problems that have happened or are expected to happen have disrupted or will disrupt the regular operations of airlines. Airlines that can do a good job during irregular operations probably lead the market and make more money, and there will be more passengers that would like to select these airlines. Aircraft schedule recovery is a key part in irregular operations, which has always been the central issue in OR research. We can define aircraft schedule recovery problem as this: Given aircrafts' initial position, regular flight schedule, the length of recovery period, then a discrete time band network should be calculate; then a model should be modeled on the base of the network; In the last, we solve the model to obtain a 'best' assignment between aircrafts and flights so that the delay and loss can be cut down. On the other hand, airlines should also operate their regular published flights after the recovery period. But aircraft schedule recovery is a very complex real time network optimization problem, belonging to NP-

hard problem. Some rapid and effective algorithms should be applied to solve optimization problem for obtaining enormous economic benefits. Several algorithms are invented to large scale NP-hard optimization problems, for instance, BCP (Branch, Cut and Price), dynamic programming and Lagrangian relaxation, column generation, simplex method etc.

Several works study on irregular flight recovery problem. Jarra (1993) introduces two cost flow models considering cancelled flights and delayed flights and then solves it with shortest path algorithm [1]. Teodorovic and stojko (1995) propose a Lexicographic dynamic programming model to minimize the number of cancelled flights and delay of passengers [2]. Wei et al. (1997) introduce a recovery method for crew management based on a multi-commodity integer network flow and also develop a heuristic branch-and-bound search algorithm [3]. In Argüello et al. (1997), the authors propose a fast heuristic based on randomized neighborhood search to solve irregular flight recovery problem [4]. Argüello et al. (2001) presents a heuristic based on an integral minimum cost flow on the time-band network [5]. Shavell (2001) studies the economic impact of schedule disruptions on airline companies [6]. Kohl et al. (2007) gives an overview of the literature on airline scheduling and discuss different aspects of the problem when irregular events occur [7]. Khaled F. Abdelghany, Ahmed F. Abdelghany, Goutham Ekollu (2008) propose an integrated decision support tool for irregular operations of airlines; they also model an optimized model on the base of cost and refer that the irregular flight recovery is a dynamic progress [8]. In china, Zhao xiuli and Gao qiang also contribute to this field.

A model that minimizes the airlines' cost is presented in this paper, and we apply column generation with Bellman-Ford [9] to solve the model. The generation of the time-space network refers to reference [10].

This paper is organized as follows. In the next section, a mathematical model of irregular flight recovery is presented. In section 3, we introduce column generation algorithm. Section 4 presents several computational experiments. Related conclusions were given in the last section.

II. MATHEMATICAL MODEL OF IRREGULAR FLIGHT RECOVERY

Disruption displeases passengers, and at the same time airlines suffer a great loss from it. So an efficient model

that can minimize the loss of airlines or can shorten passengers' delay is extremely needed. In this paper, we model on the base of time space network and multi commodity flow theory. There are two commodities, namely all the aircrafts in the recovery period and all flights that may be cancelled due to disruptions. In this paper, we regard each aircraft as a commodity and the cancelled flight is carried out by virtual aircraft. So irregular flight recovery is actually a minimum cost flow with integer multiple commodity capacity constraints. To ensure regular flight schedule after the end of recovery period, it is necessary to clarify the number and type of aircrafts needed in each airport.

We assume that there are i flights, j airports and k available aircrafts in the recovery period. Define the following sets, parameters, variables:

Sets:

F : set of flights;

P : set of planes;

S : set of stations;

R : set of feasible routes;

Indices:

f : indice of flight;

r : indice of route, or recovery plan;

p : indice of plane;

s : indice of station;

Decisions:

$$y_f = \begin{cases} 1 & \text{cancel flight } f \\ 0 & \text{otherwise} \end{cases}$$

$$x_r^p = \begin{cases} 1 & p \text{ covers } r \\ 0 & \text{otherwise} \end{cases}$$

Parameters:

z : minimized cost;

c_f : cancellation of flight f ;

d_r^p : delay cost when p covers route r ;

m_r^p : additional cost when p covers r ;

$$a_r^f = \begin{cases} 1 & f \in r \\ 0 & \text{otherwise} \end{cases}$$

$$b_s^r = \begin{cases} 1 & r \text{ ends at } s \\ 0 & \text{otherwise} \end{cases}$$

h_s : number of planes needed by airport s normally

Irregular flight mathematical model:

$$\text{minimize } \sum_{f \in F} c_f y_f + \sum_{p \in P} \sum_{r \in R} (d_r^p + m_r^p) x_r^p \quad (1)$$

Subject to:

$$\sum_{p \in P} \sum_{r \in R} a_r^f x_r^p + y_f = 1, \quad \forall f \in F \quad (2)$$

$$\sum_{r \in R} x_r^p \leq 1, \quad \forall p \in P \quad (3)$$

$$\sum_{p \in P} \sum_{r \in R} b_s^r x_r^p = h_s, \quad \forall s \in S \quad (4)$$

$$x_r^p = 0, 1 \quad \forall r \in R, p \in P \quad (5)$$

$$y_f = 0, 1 \quad \forall f \in F \quad (6)$$

Each recovery plan or feasible route r has a cost which is made up of delay cost and some potential cost and is associated with a binary variable x_r^p that equals one if the flight f is in the set of r , or simply that recovery plan is taken into the solution, 0 otherwise. A recovery plan is also described by the binary constants b_s^r, a_r^f, x_r^p . Those constants take value one if the plan r ends with the expected final state s , covers the flight f , and is serviced by plane p , respectively. A binary variable y_f is associated with each flight and it equals one if the flight f is canceled with cancellation cost c_f . Constraint (2) is a constraint about flight coverage, ensuring that each flight is either serviced or cancelled. Constraint (3) constraints that one plane can occupy one feasible route at most. Constraint (4) ensures that original schedule can be recovered at the end of recovery period. Constraint (5) and (6) enforce the binary constraints on the variables.

III. ALGORITHM INTRODUCTION

A. Column Generation Algorithm

For the class of linear program problems which has large number of variables but relatively few of constraints, if simplex algorithm is used to solve, a large capacity for saving coefficient matrix is needed and computation time is long for obtaining the matrix. On the other hand, it is always unnecessary on the ground that basic variables are much fewer than non-basic ones, large quantity of computation for obtaining the coefficient vectors of non-basic variables is wasted.

The general idea of column generation is that optimal solutions to large LP's can be obtained without explicitly including all columns (i.e., variables) in the constraint matrix (called the Master Problem or MP). In fact, only a very small subset of all columns will be in an optimal solution and all other (non-basic) columns can be ignored. The initial linear program is divided into master problem and sub problem. Because of the large scale number variables in master problem, we usually select part variables to construct the restricted master problem. Solve the restricted master problem, and pass the dual variables of master problem to sub problem; then sub problem check whether there are negative reduced cost column, if there is, add the column with negative reduced cost into master problem, and solve the master problem again till there are not negative columns generated in the sub problem.

The procedure of this algorithm is presented as following: Given one solution of LP for a restricted set of variables, if the current solution can still be improved, we need to find another solution to be added into our restricted problem; otherwise, the current solution is optimal for the problem. In case of minimization problem, addition of one variable will only improve the solution if its reduced cost is negative. It means that if all variables' reduced costs are non-negative, then the optimal solution for the entire problem has been found.

To better present column generation, we give a simple model as following.

$$\text{minimize } \sum_{p \in P} c_p x_p \quad (7)$$

Subject to:

$$\sum_{p \in P} a_p x_p \geq b \quad (8)$$

$$x_p \geq 0 \quad \forall p \in P \quad (9)$$

As we can see from the model, if there are many parameters in set P while it is restricted to only one constraint, then it is difficult to solve it through simplex method. Thus we can solve it with column generation algorithm. We give the solving procedure flow chart following Fig1.

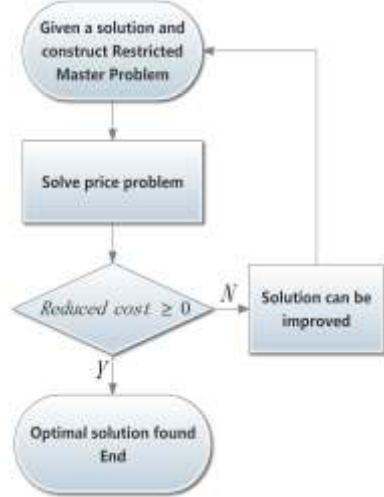


Figure 1. Column generation flow chart.

B. Sub-Problem in Column Generation Algorithm

The sub-problem is a process that generates new column. In our irregular flight recovery problem, the column generated by sub-problem is actually one new recovery plan for one aircraft. In the sub-problem, we need to calculate the reduced cost. In irregular flight recovery problem, we expect the minimized cost at last, so if all the variables' reduced cost is negative, then the optimal solution is found. Thus in the sub problem, we can apply shortest path algorithm to calculate the minimum cost for each feasible route/plan. In this research, we solve the sub-problem with Bellman-Ford algorithm. On the ground that the cost in each arc may be negative, thus we cannot solve it through Dijkstra's algorithm. It is an algorithm that computes shortest paths from a single source vertex to all of the other vertices in a weighted digraph. Bellman-Ford algorithm simply relaxes all the edges, and does this $|V| - 1$ times, where $|V|$ is the number of vertices in the graph. The detail procedure is presented in algorithm following.

Algorithm BellmanFord algorithm

```

1: procedure BF(listvertices, listedges, vertexsource)
2:   for each vertex  $v \in$  vertices do
3:     if  $v$  is source then
4:       distance[ $v$ ] = 0
5:     else
6:       distance[ $v$ ] =  $\infty$ 
7:     end if
8:     predecessor[ $v$ ] = null
9:   end for
10:  for  $i = 1$  to size(vertices) - 1 do
11:    for all edge( $u, v, w$ )  $\in$  edges do
12:      if distance[ $u$ ] +  $w$  < distance[ $v$ ] then
13:        distance[ $v$ ] = distance[ $u$ ] +  $w$ 
14:        predecessor[ $v$ ] =  $u$ 
15:      end if
16:    end for
17:  end for
18:  for all edge( $u, v, w$ )  $\in$  edges do
19:    if distance[ $u$ ] +  $w$  < distance[ $v$ ] then
20:      error : negative - weight cycle
21:    end if
22:  end for
23: end procedure
  
```

C. Solving Model

In this research, we solve our model following steps below.

Step 1: Obtain the restricted master problem and solve it with simplex method.

Step 2: Calculate duals of constraints,
 λ_i corresponds to formula(2),
 ξ_p corresponds to formula(3),
 η_s corresponds to formula(4);

Step 3: Modify arcs' cost in the network;

Step 4: Find the shortest cost route and then calculate reduced cost with following formula (10).

Step 5: Check if the minimum reduced cost is negative; if it is, go to step 1; otherwise, end the process.

$$reduced\ cost = d_r^p - \sum_{f \in F} a_f^f \lambda_f - \sum_{s \in S} b_s^r \eta_s - \xi_p \quad (10)$$

Several experiments are presented below to better describe this process.

IV. COMPUTATIONAL EXPERIMENTS

A program has been developed to implement our algorithm. The experiments were carried out on a PC with 2.53 GHz Intel(R) Core (TM) i3 CPU, 1GB of RAM. We did all this work in Microsoft visual studio 2010.

In this research, we consider an airline with 3 aircrafts and 12 flights, and the origin schedule is given in Table 1. We set MCT (Minimum Connection Time) of each airport 40 minutes; every airport's curfew time is set 2400 in the midnight. The airline will occupy a cost of \$20 once one flight is delayed by one minute.

Given that aircraft 3 should have some mechanical problems when checked at the origin, it couldn't fly the

subsequent flight that day. In this situation, if flight 31,32,33,34 flown by aircraft 3 were cancelled simply, airlines will occupy a loss of \$58,175. We would like to test our model and algorithm in this situation.

TABLE I. ORIGIN FLIGHT SCHEDULE

Flight	Origin	Destination	Departure	Arrival	Cancellations
Aircraft 1					
11	DAB	ORF	1410	1520	7350
12	ORF	IAD	1605	1700	10231
13	IAD	ORF	1740	1840	7434
14	ORF	DAB	1920	2035	14191
Aircraft 2					
21	ORF	DAB	1545	1700	11189
22	DAB	ORF	1740	1850	12985
23	ORF	IAD	1930	2030	11491
24	IAD	ORF	2115	2215	9581
Aircraft 3					
31	IAD	ATL	1515	1620	9996
32	ATL	IAD	1730	1830	15180
33	IAD	ATL	1910	2020	17375
34	ATL	IAD	2100	2155	15624

As Table 2 below shows, we only need cancel flight 33, 34 with a loss of \$45,999, saving about 21% for airlines comparing to cancelling all the flights flown by aircraft 3.

TABLE II. MODIFIED SCHEDULE

Flight	Origin	Destination	Departure	Arrival	Cancellations	Delays
Aircraft 1						
11	DAB	ORF	1410	1520		
12	ORF	IAD	1605	1700		
31	IAD	ATL	1740	1845		2900
32	ATL	IAD	1925	2025		2300
13	IAD	ORF	2055	2155		3900
14	ORF	DAB	2235	2350		3900
Aircraft 2						
21	ORF	DAB	1545	1700		
22	DAB	ORF	1740	1850		
23	ORF	IAD	1930	2030		
24	IAD	ORF	2115	2215		
Aircraft 3						
33	IAD	ATL			17375	
34	ATL	IAD			15624	

Departure difference of flight 32, 13 is 30 minutes, less than MCT. What leads to the shorter connection time is that we calculate each node's obtain time by selecting the earliest departure of all the flights departing from this node. Thus the delays calculated in this paper are actually less than the real delays. To minimize this inaccuracy deviation the length of discrete time band is suggested to be shortened. But the shorter discrete time band can unfortunately enlarge the network's scale, which could influence the solving efficiency. That is to say, the difference can be shortened to zero as much as possible, but in this paper, time-space network is not introduced for the length of this article.

V. CONCLUSIONS

In this research, a model on the base of multi-commodity flow is given, and it is solved through column generation with Bellman-Ford algorithm. We draw the

conclusion that our algorithm is feasible and efficient in solving the irregular flight recovery problem, where we solve sub problem by using Bellman-Ford algorithm instead of calculating all the feasible paths for the restricted master problem.

Since this is an ongoing project, several issues should be defined and extended. At the same time, we consider cancellations and delay in this paper, ignoring the swapping of resources. In the future, we would like to study the swapping resources intensively, and develop a system so that optimal decision is produced in time when a disruption happens.

ACKNOWLEDGMENT

The research has been supported by the National Science Foundation of China Grant No. 71171022, Engineering Research Center of Information Networks, Ministry of Education.

REFERENCES

- [1] Jarrah, A.I.Z., Yu, G., Krishnamurthy, N. and Rakshit, A:(1993) A decision support framework for airline flight cancellations and delays, *Transportation Science*, 27(3),266-280.
- [2] Teodorovic, D., Stojkovic, G., 1995: Model to reduce airline schedule disturbances, *Journal of Transportation Engineering* 121,324-331.
- [3] Wei, G., Yu, G. and Song ,M.(1997): Optimization model and algorithm for crew management during airline irregular operations, *Journal of Combinatorial Optimization* 1: 305-321.
- [4] Argüello, M.,Bard, J. and Yu, G. (1997):A grasp for aircraft routing in response to groundings and delays, *Journal of Combinatorial Optimization* 5: 211-228.
- [5] Argüello, M.,Bard, J. and Yu, G. (2001):Optimizing aircraft routings in response to groundings and delays, *IIE Transactions*33: 931-947.
- [6] Shavell, Z.:The Effects of Schedule Disruptions on the Economics of Airline Operations, USA/Europe 3rd Air Traffic Management Seminar (2000).
- [7] Kohl, N., Larsen, A., Larsen, J., Ross, A. and Tiourine, S.(2007):Airline disruption management - perspectives, experiences and outlook, *Journal of Air Transport Management*13(3): 149-162.
- [8] Khaled F. Abdelghany, Ahmed F. Abdelghany, Goutham Ekollu:An integrated decision support tool for airlines schedule recovery during irregular operations, *European Journal of Operational Research*, Volume 185, Issue 2, 1 March 2008, Pages 825-848, ISSN 0377-2217.
- [9] Awerbuch, Baruch (Massachusetts Inst of Technology,Cambridge, United States);Approximate distributed Bellman-Ford algorithms.Source: *IEEE Transactions on Communications*,v42,n 8,p 2515-2517,Aug1994.
- [10] Argüello, M., Bard J.F., and Yu G. (1998):Models and methods for managing airline irregular operations, in *Operations Research in the Airline Industry*, edited by G. Yu. Boston: Kluwer Academic Publishers.