

不正常航班恢复的一种改进的列生成算法

吴 刚¹ 严 俊²

(1. 南京航空航天大学民航学院, 南京, 210016; 2. 中国民航科学技术研究院科技处, 北京, 100028)

摘要:分析了由飞机资源短缺造成的航班不正常情况, 将该问题视为带有容量约束的多个商品的整数最小费用流问题, 建立了多商品网络流数学模型。采用列生成算法框架求解该大规模整数规划问题。一般的列生成算法是每次迭代过程只加入一个列, 造成迭代次数过多, 因而降低了算法效率低。文中采用每次迭代过程中加入多个列, 并对加入的多个列应该满足的条件进行了分析。最后给出的算例验证了该方法的正确性和有效性。

关键词:不正常航班; 多商品网络流; 列生成

中图分类号: U8

文献标志码: A

文章编号: 1005-2615(2014)02-0329-06

Improved Column Generation Algorithm for Disrupted Airline Schedules Recovery

Wu Gang¹, Yan Jun²

(1. College of Civil Aviation, Nanjing University of Aeronautics & Astronautics, Nanjing, 210016, China;

2. Science and Technology Division, China Academy of Civil Aviation Science and Technology, Beijing, 100028, China)

Abstract: The disrupted airline schedules caused by the shortage of aircraft resources is considered as a multi-commodity integer minimum cost flow with the side constraints, and a mathematic model of multi-commodity network flow is established. The column generation scheme is introduced to solve the integer programming problem. Generally, only one column is added in each iteration for the column generation algorithm, which results in excessive number of iterations, thereby reduces the efficiency of the algorithm. The method of adding more columns in each iteration process is set up, and the requirements which the added multiple columns should meet are analyzed. Finally, an example is given to validate the correctness and efficiency of the method.

Key words: disrupted airline schedules; multi-commodity network flow; column generation

不正常航班是指不能按照航班计划执行的航班。造成航班不正常的原因常有航空器故障、恶劣天气、航空管制、机场保障以及旅客自身不能及时登机等。当航班发生不正常时, 航班公司运行调度人员需要对航班计划进行实时调度, 使得航班在最短的时间内恢复正常, 同时还要确保公司为此付出的代价最小。随着航班数量的增长, 依靠运行调度人员的经验进行航班调整变得越来越困难。最优化方法与计算机技术相结合的自动化恢复技术能够快速生成调度方案并为航空公司获取巨大的经

济效益和社会效应, 有关这方面的研究越来越为航空公司所关注。

不正常航班调度是一个非常复杂的实时网络优化问题, 属于非确定性多项式时间(Non-deterministic polynomial-time, NP)难问题。国外在这方面研究较早: 文献[1]针对飞机资源短缺提出了以旅客总延误最小化为目标函数的模型, 采用分支定界法求解。文献[2]在不考虑调机情况下以取消的航班数量和旅客总延误时间最小作为目标函数, 采用字典序优化技术进行建模求解。文献[3]研究

基金项目: 国家自然科学基金委员会与中国民用航空局联合基金(61079014)资助项目; 国家软科学研究计划(2011GX4B023)资助项目。

收稿日期: 2013-01-19; **修订日期:** 2013-11-21

通信作者: 严俊, 男, 高级工程师, E-mail: yanj@mail.castc.org.cn。

了暂时短缺飞机的实时扰动问题。在研究中允许采用调机策略,分别对延误和取消策略建立了最小费用网络流模型,但两个模型只能分开使用。文献[4]研究了由于机务故障导致飞机临时短缺情况下的航班计划恢复问题,以旅客延误时间最小作为目标函数,针对飞机路径的重构提出了一种贪婪随机自适应搜索算法(Greedy randomized adaptive search procedure, GRASP)。文献[5]提出了一种基于模拟退火与贪婪随机自适应搜索相结合的算法,可同时解决机务故障与机场关闭两种情况下多机型航班恢复问题。文献[6]提出了在飞机资源短缺情况下,采用时空网络建立模型,模型的求解转化为在网络中为每一架飞机寻找一条总费用最小的航班执行路线。

本文采用多商品网络流的数学模型来描述同一类型飞机的不正常航班恢复问题。针对飞机路线数量极多的情况,提出一种飞机资源短缺情况下的飞机路线构造方法,采用列生成算法框架,对入选飞机路线进行优化选择。最后给出了算例,验证了本文提出的方法的正确性和有效性。

1 多商品网络流模型

本文采用文献[3]中基于边的流变量的多商品网络流模型。每架飞机都视为一种商品,将取消的航班视为由虚拟飞机执行。单个飞机的恢复方案就变成寻找延误成本最小的可飞航班集合。为确保恢复期结束后后续航班的正常执行,各机场所需的飞机数量必须得到保证。

设在恢复期内涉及的航班数为 n ,机场数为 m ,可利用的飞机数为 k ,并引入以下符号:

(1) 集合

F : 航班集合; K : 可利用飞机集合; A : 机场的集合; P : 飞机可行路径集合。

(2) 上下标号记号

i : 航班下标, $i \in F$; k : 飞机上标, $k \in K$; a : 机场下标, $a \in A$; j : 飞机路径下标, $j \in P$ 。

(3) 参数

a_{ij} : 当航班 i 包含在路径 j 中时等于1,否则等于0;

b_{ja} : 当路径 j 在机场 a 终止时等于1,否则等于0;

c_i : 取消航班 i 的成本;

h_a : 恢复期结束时,为执行正常航班计划机场 a 需要的飞机架数;

d_j^k : 将飞机 k 执行路径 j 的成本,包括航班延

误成本和飞机交换成本。

(4) 决策变量

x_j^k : 将飞机 k 执行路径 j 时等于1,否则等于0;

y_i : 航班 i 取消时等于1,否则等于0。

应用上述记号,建立如下多商品网络流模型

$$\min z = \sum_{k \in K} \sum_{j \in P} d_j^k x_j^k + \sum_{i \in F} c_i y_i \quad (1)$$

$$\text{s. t. } \sum_{k \in K} \sum_{j \in P} a_{ij} x_j^k + y_i = 1, i \in F \quad (2)$$

$$\sum_{k \in K} \sum_{j \in P} b_{ja} x_j^k = h_a, a \in A \quad (3)$$

$$\sum_{j \in P} x_j^k \leq 1, k \in K \quad (4)$$

$$x_j^k = 0, 1, j \in P, k \in K \quad (5)$$

$$y_i = 0, 1, i \in F \quad (6)$$

式(1)为目标函数,表示航班延误成本和取消成本之和最小;式(2)为航班覆盖约束,表示每个航班要么被取消,要么被执行;式(3)为飞机流平衡约束,表示恢复期结束后各机场拥有的飞机架数必须等于后续航班正常执行需要的架数;式(4)为飞机可行路径的流量约束,表示每架飞机最多只能执行一条可行路径;式(5,6)为变量取值约束。

多商品网络流模型的特殊性,其解本身具有整数性。但是该问题是一个严重退化的线性规划问题。在大规模情况下将所有飞机路线放入模型进行求解将不可行,因此选用列生成算法求解该问题是可行的方法。

2 列生成算法

列生成算法^[7](Column generation algorithm, CGA)将原线性规划问题分解为主问题(Main problem, MP)和子问题(Sub-problem, SP)。由于MP的变量数目巨大,选择其中部分变量(至少包含一个可行解)构建一个限制主问题(Restrain MP, RMP)。求解至最优,并将此时RMP的对偶变量值传递给SP。求解SP生成具有负减约成本的列(MP目标函数求最小),将负减约成本最小的列加入到RMP中,继续求解RMP。重复上述过程,直至SP无法生成具有负减约成本的列为止,此时原问题达到最优。

2.1 限制主问题

将式(1~6)作为CGA的MP模型。但是对这个模型直接求解很困难,对于上百架飞机的恢复问题,变量的规模将达到上亿个。因而需要对MP的规模进行限制。选择飞机路线集合 P 中的部分飞机路线 P' 构建不正常航班恢复模型的RMP。

对于该问题,将原有的飞机的路线进行顺序延迟就可以得到一个可行解,也是一个最小的 RMP。通过 RMP 求解获得其对偶变量。航班约束、机场约束以及飞机约束的对偶变量的值,分别用 λ_i, η_a 与 μ_k 表示。依据单纯型理论,如果存在一条飞机路线 $j \in P'$,其减约成本小于零,则将其替代原基中的变量将使目标函数减少。

2.2 子问题

CGA 中的 SP 是一个生成“列”的过程,在不正常航班恢复问题中,RMP 数学模型中的每一列都对对应实际问题中的某架飞机的一个新的执行方案,于是 SP 求解是一个构建新的执行方案的过程。在 RMP 中,第 j 列的减约成本计算公式为

$$RPL_j = d_j^k - \sum_{i \in F} a_{ij} \lambda_i - \sum_{a \in A} b_{ja} \eta_a - \mu_k \tag{7}$$

RMP 目标函数求最小,SP 需选择负的减约成本的飞机路线加入到 RMP

$$\min_{j \in P} RPL_j \tag{8}$$

该 SP 的数学模型可以转化为标准的最小费用流模型

$$\begin{aligned} \min \quad & z = \sum_{m \in V} \sum_{n \in V} d_{mn} x_{mn} \tag{9} \\ \text{s. t.} \quad & \sum_{n \in \text{adj}(m)} x_{mn} - \sum_{n: m \in \text{adj}(n)} x_{nm} = \begin{cases} 1, & m = s \\ 0, & \text{其他} \\ -1, & m = t \end{cases} \tag{10} \\ & x_{mn} = 0, 1 \quad m, n \in V \tag{11} \end{aligned}$$

式(9)中: d_{mn} 表示时空网络中每条边的减约成本, x_{mn} 表示边的流量取值是单位流,取值 1 实际表明了此边构成飞机路线中的边。一条飞机路线的减约成本(RPL_j)即为构成飞机路线的所有边的减约成本之和($\sum_{m \in V} \sum_{n \in V} d_{mn} x_{mn}$)。式(10)表示了网络中流量平衡。

当存在减约成本为负的飞机路线 d_j^k 时,将其加入 RMP 并不能保证其在基本解中取值为 1,即不能使目标函数值下降。每一次迭代过程在 RMP 中只加入一条负减约成本的飞机路线的代价就显得太大。同时在航空公司实际应用过程中可供制订恢复方案的时间很短,计算过程可能随时终止,如果终止时目标值一直不变化,即可行方案无变化则会给应用带来困难。所以考虑一次在 RMP 中加入多条飞机路线既可提高效率又能保证恢复制定过程不断产生新的可行方案。

假如加入的负减约成本的飞机路线最终成为最优解中的基解,且其取值为 1,那么飞机路线 x_j^k

的形式,最多只有 4 种:第一种是 $x_j^k \subset x_{j_{\text{基}}}^k$ 且 $b_{ja} = b_{ja_{\text{基}}}$;第二种是 $x_j^k \subset x_{j_{\text{基}}}^k$ 且 $b_{ja} \neq b_{ja_{\text{基}}}$;第三种是 $x_j^k \not\subset x_{j_{\text{基}}}^k$ 且 $b_{ja} = b_{ja_{\text{基}}}$;第四种是 $x_j^k \not\subset x_{j_{\text{基}}}^k$ 且 $b_{ja} \neq b_{ja_{\text{基}}}$ 。

第一种情况时, d_j^k 自然满足机场约束式(3),航班约束式(2)也必然满足,只是对于属于 x_j^k 而不属于 $x_{j_{\text{基}}}^k$ 的航班 i ,以 $y_i = 1$ 的形式出现。

第二种情况时, x_j^k 不满足机场约束式(3),则在 RMP 的飞机路线 P' 中必须存在另一飞机 k' 飞机路线 j ,使得其 $b_{ij_0} = b_{ij_{\text{基}}}$ 。该飞机路线须满足飞机约束式(4),那么这样将使得基中其对应的飞机 k' 的飞机路线取零值或出基。

第三种情况时, d_j^k 自然满足机场约束式(3),但航班约束式(2)不满足,则在 RMP 的飞机路线 P' 中必须存在至少一个飞机 k' 的飞机路线 j ,使得构成其飞机路线的航班属于基 k' 而不属于 j 。那么这样将使得基中其对应的飞机 k' 的飞机路线取零值或出基。

第四种情况时, d_j^k 不满足机场约束式(3),航班约束式(2)。这种情况是第三种情况的扩展,同样将使得基中至少另一条飞机路线取零值或出基。

综合上述 4 种情况,一般而言,当将一条减约成本为负的列加入到 RMP 中时,RMP 基发生改变,如果该列取值为 1,则基中相应还应至少存在 1 个相应列取值为 1。

对于求解而言,在 SP 产生的列加入 RMP 后,使其进入基并且使得目标函数的值下降是最好的情况。但是要保证其目标值下降,则必引起具有有限性的一些列也进入基,但这些列的减约成本不一定为负值,称这样的列为伴随列或伴随飞机路线。但是如果加入 RMP 的所有飞机路线减约成本之和小于零,则一定使目标函数值下降。

因此不正常航班恢复子问题的 SP

$$\min_{j \in s} \sum_j RPL_j \tag{12}$$

$$\text{s. t.} \quad \min_{j \in P} \{RPL_j, 0\} < 0 \tag{13}$$

$$s \in S \tag{14}$$

式中: s 为包含减约成本为负入基飞机路线组合,该组合中的飞机路线满足约束(2~4),且当其减约成本之和为负时,将其加入。

2.3 入基飞机路线组合的构造

对于入基飞机路线某个组合 s ,如果在原 RMP 中已经包含了 s 中除负减约成本的所有飞机路线,则 s 不需要构造。但是 RMP 为了求解方便,一般严格控制飞机路线的数量,因此这种情况很少出现,构造入基伴随路线必不可少。

构造步骤如下:

- (1)从负减约飞机路线集的第一条飞机路线开始,根据构成该路线的航班确定 S 中所涉及的飞机路线(飞机路线在基中取值为 1)、航班集 V 、停驻机场集 T 。
- (2)从该航班集中去除包含在最小负减约飞机路线中的航班集 V' 。
- (3)航班集 V' 中所有航班延误成本减去其对偶变量值。

(4)求以 S 中不包含减约飞机路线所对应飞机之外的飞机 S' 中飞机在航班集 V' 以停驻机场集 T 为飞机路线终点的一个最小延误成本覆盖问题。

3 算例分析

某航空公司航班运行状态如表 1 所示。设取消一个航班的成本等价于该航班延误 480 min(国际民航业一般认为取消一个航班等同于该航班延误 8 h)。

表 1 初始航班计划与飞机状态
Tab. 1 Initial flight schedule and aircraft status

飞机号	就绪机场	就绪时刻	航班号	出发机场	到达机场	计划起飞时刻	计划到达时刻	飞行时间/min
1	ZSPD	23:50	11	ZSPD	VHHH	12:30	15:10	160
			12	VHHH	ZSPD	16:35	19:10	155
2	ZUUU	12:30	21	ZUUU	ZSPD	12:30	15:10	160
			22	ZSPD	VDPP	20:00	23:50	230
3	ZSQD	11:10	31	ZSQD	ZSPD	11:00	11:55	55
			32	ZSPD	ZUUU	14:55	18:20	205
			33	ZUUU	ZSPD	19:15	21:35	140

从表 1 中可知,1 号飞机因机务故障使得计划中的 11 航班与 22 航班将发生长时间的延误,以顺延方式形成一个初始解。

第一次迭代,根据初始解得到对偶变量,找到最小负减约成本路线为 2 号飞机所对应的 21—32—33。其组合中伴随飞机路线只有 3 号飞机路线 31—22,入基飞机路线组合 RPL 之和大于零,所以加入 RMP 只有飞机路线 21—32—33。

第二次迭代时,找到最小负减约成本路线为 3 号飞机所对应的 31—11—12—22。其组合中伴随飞机路线有 1 号飞机路线,其路线是无航班,2 号飞机路线 21—32—33(其已经存在于 RMP 中),入基飞机路线组合 RPL 小于零,所以加入 RMP 有两条飞机路线。

第三次迭代时,目标函数值变为 100,此时已经找不到负减约成本路线,表明此时已经找到最优解。在最优解中的各飞机路线见表 2。

如果按照参考文献[8]中的方法,每次只确定一条进入限制性主问题的飞机路线,其迭代过程见表 3。

4 算法效率分析

8 组国内某航空公司不同规模算例的基本信息,在 Inter(R) Core(TM) 2 Duo CPU T5670 1.80 GHz,2 GB 内存电脑上,使用 CGA 和文献[5]中的贪婪随机模拟退火算法(Greedy randomized search Annealing, GSA)进行对比,结果见表 4。

表 2 最优解中的飞机路线
Tab. 2 Aircraft routes in optimal solution

飞机号	就绪机场	就绪时刻	航班号	出发机场	到达机场	预计起飞时刻	预计到达时刻	飞行时间/min	延误时间/min
1	ZSPD	23:50	21	ZUUU	ZSPD	12:30	15:10	160	0
			32	ZSPD	ZUUU	15:50	19:15	205	55
2	ZUUU	12:30	33	ZUUU	ZSPD	19:55	22:15	140	40
			31	ZSQD	ZSPD	11:00	11:55	55	0
3	ZSQD	11:10	11	ZSPD	VHHH	12:35	15:15	160	5
			12	VHHH	ZSPD	16:35	19:10	155	0
			22	ZSPD	VDPP	20:00	23:50	230	0

表 3 只选择最小负减约成本的迭代过程
Tab. 3 Iterative process of only selecting the minimum negative reduced cost

迭代 次数	基中取值非零的 飞机路线	目标 函数值	最小负减约成本 所对飞机号	飞机路线	减约成本	延误 成本
1	1:11—12	1 315	2	21—32—33	—1 700	95
	2:21—22					
	3:31—32—33					
2	1:11—12	1 315	3	31—11—12—22	—1 340	5
	2:21—22					
	3:31—32—33					
3	1:11—12	1 315	3	31—11—12	—1 795	0
	2:21—22					
	3:31—32—33					
4	1:11—12	1 315	1	无航班	—1 215	0
	2:21—22					
	3:31—32—33					
5	1:无航班	100	无			
	2:21—32—33					
	3:31—11—12—22					

表 4 算例及两种结果分析
Tab. 4 Calculation examples and two kinds of results analysis

序号	规模	算法	取消 航班数	延误 航班数	延误 成本	各延误时间段延误航班数分布			
						≤0.5 h	0.5~2 h	2~4 h	>4 h
算例 1*	4a_2t_14f(1d)	CGA	0	2	125	0	2	0	0
		GSA	0	2	125	0	2	0	0
算例 2*	9a_1t_21f(5d)	CGA	0	9	687	3	4	2	0
		GSA	0	9	687	3	4	2	0
算例 3*	11a_2t_35f(4d)	CGA	0	11	882	1	8	1	1
		GSA	0	10	1 053	2	5	2	1
算例 4*	16a_1t_59f(4d)	CGA	0	2	140	0	2	0	0
		GSA	0	4	290	0	4	0	0
算例 5*	18a_1t_69f(6d)	CGA	0	10	1 290	0	5	3	2
		GSA	0	21	1 855	4	11	6	0
算例 6*	26a_1t_90f(3d)	MSP	0	2	46	2	0	0	0
		GSA	0	3	81	3	0	0	0
算例 7	30a_2t_112f(3d)	CGA	0	1	70	0	1	0	0
		GSA	0	5	295	1	4	0	0
算例 8	46a_3t_173f(5d)	CGA	0	5	260	0	5	0	0
		GSA	0	12	620	3	8	1	0

注：* 表示 CGA 的解已证明最优；规模列中的数字字母组合表示日航班计划中的飞机、机型、航班以及发生故障的飞机数目，例如：4a_2t_14f(1d)表示 4 架飞机，2 种机型，14 个航班，有一架飞机发生故障。

表 4 中算例分别用 CGA 算法和 GSA 算法进行求解,运算时间均在 10 min 之内,符合不正常航班恢复的实时性要求。由表 7 的结果对比可知,与 GSA 算法相比,CGA 算法的求解质量更高。由 8 组数据测试的结果统计可知,对 75%的算例,CGA 算法比 GSA 算法的延误成本更小;对 62.5%的算例,CGA 算法比 GSA 算法的延误航班数更少(若

延误航班数相同,CGA 在延误时间的分布上也更为合理,延误多为短时间(2 h 之内)和中等时间(2~4 h))。但是在时间效率上 GSA 算法比 CSA 算法要高效。随着规模不断增长,GSA 算法运行时间变化增长很缓慢,所有算例都在 3 min 内完成,而 CSA 算法运算时间却增长较快,最后两个例子接近 10 min 才完成。

5 结束语

本文研究了 CGA 框架下的求解大规模不正常航班恢复问题的方法。该方法针对该问题是严重线性规划退化问题的特点,提出了一种改进的列生成方法,该方法确定了以入基列的减约成本之和小于零作为伴随列进入 RMP 的判断依据,通过该方法可以减少迭代次数,提高了求解效率。通过一个实例验证了方法的正确性与有效性,并以一组算例对比分析了算法的效率。本文存在的不足之处是没有提供一个能够快速求解具有负值边的最短路有效的算法,这是进一步研究的方向。

参考文献:

- [1] Kohl N, Larsen A, Larsen J, et al. Airline disruption management—Perspectives, experiences and outlook[J]. *Journal of Air Transport Management*, 2007, 13 (3): 149-162.
 - [2] Teodorovic D, Guberinic S. Optimal dispatching strategy on an airline network after a schedule perturbation[J]. *European Journal of Operational Research*, 1984, 15 (2): 178-182.
 - [3] Jarrah A I Z, Yu G, Krishnamurthy N, et al. A decision support framework for airline flight cancellations and delays[J]. *Transportation Science*, 1993, 27 (3): 266-280.
 - [4] Arguello M F, Bard J F, Yu G. A GRASP for aircraft routing in response to groundings and delays[J]. *Journal of Combinatorial Optimization*, 1997, 1 (5): 211-228.
 - [5] 朱金福, 高强. 载运工具受扰路线恢复方法[P]. 中国专利:200910264134, 2010-06-23.
 - [6] Yan S, Yang D. A decision support framework for handling schedule perturbations[J]. *Transportation Research, Part B: Methodology*, 1996, 30 (6): 405-419.
 - [7] Barnhart C. Branch-and-price: Column generation for solving huge integer programs[J]. *Operations Research*, 1998, 46 (3): 316-329.
 - [8] 白凤, 朱金福, 高强. 基于列生成法的不正常航班调度[J]. *系统工程理论与实践*, 2010, 30 (11): 2036-2045.
- Bai Feng, Zhu Jinfu, Gao Qiang. Disrupted airline schedules dispatching based on column generation methods[J]. *Systems Engineering—Theory & Practice*, 2010, 30(11): 2036-2045.