

选择子网络的飞机恢复问题模型改进研究

陈斌斌

同济大学经济与管理学院, 上海

收稿日期: 2024年11月20日; 录用日期: 2024年12月18日; 发布日期: 2024年12月27日

摘要

本文分析选择子网络的飞机恢复方法的缺陷, 即非目标不正常航班可能导致额外的航班取消。针对该缺陷, 本文对经典飞机恢复模型进行改进, 增加虚拟航班的概念以减小非目标不正常航班的影响。最后, 本文使用机器学习选择子网络并滚动求解, 对改进模型的效果进行测试。实验结果对经典模型和改进模型的结果进行对比, 显示出改进模型在减少航班取消数量上的显著优势。

关键词

飞机恢复问题, 不正常航班恢复, 子网络, 运筹优化, 机器学习

Research on Model Improvement for Aircraft Recovery Based on Sub-Network Selection

Binbin Chen

School of Economics and Management, Tongji University, Shanghai

Received: Nov. 20th, 2024; accepted: Dec. 18th, 2024; published: Dec. 27th, 2024

Abstract

This paper analyzes the deficiency of aircraft recovery methods that employ sub-network selection, namely, the potential for non-target irregular flights to result in additional flight cancellations. To address this deficiency, this paper proposes an improvement to the classic aircraft recovery model by introducing the concept of virtual flights to mitigate the impact of non-target irregular flights. Finally, this paper utilizes machine learning to select sub-networks and employs a rolling approach to recover the whole network, thereby testing the effectiveness of the improved model. Experimental results comparing the outcomes of the classic model and the improved model demonstrate

the significant advantage of the latter in reducing the number of flight cancellations.

Keywords

Aircraft Recovery Problem, Irregular Flight Recovery, Sub-Network, Operations Research Optimization, Machine Learning

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

民航运营管理主要包括航班计划、飞机排班计划、机组排班计划、不正常航班管理等部分[1]。其中，航班计划、飞机排班计划、机组排班计划都是计划部分，需要提前几个月或者几周完成。但是，既定的航班计划往往会因为各种原因无法按时起飞或者降落，这样的航班被称之为“不正常航班”或者受干扰的航班。当不正常航班出现后，为了避免对后续航班造成严重的影响，需要在短时间内(通常为几分钟内)对不正常航班进行重新规划，以使得整体航班网络受到的损失最小，这个过程被称为“不正常航班恢复”。重新规划的方式通常有航班延误、飞机交换、航班取消等。其中飞机交换是指由另一架飞机代为执行某些航班。不正常航班恢复问题是复杂度非常高的问题，需要考虑飞机、机组、旅客等等因素。通常，不正常航班恢复问题被划分成一系列的子问题，分别包括飞机恢复问题(Aircraft Recovery Problem, ARP)、机组恢复问题(Crew Recovery Problem, CRP)和旅客恢复问题(Passenger Recovery Problem, PRP) [2]。其中，飞机恢复问题是机组恢复问题和旅客恢复问题的基础，因此飞机恢复问题在整个不正常航班恢复问题中至关重要。

飞机恢复问题可以理解成，将所有计划运营的飞机，从其恢复期开始时所在的机场，经过一系列可衔接的航班、检修或者地面等待，最终到达恢复期结束后需要抵达的机场。因此，飞机恢复问题实际上是一种多商品流问题(Multi-commodity Flow Problem, MCP)：将来自多个产地的一种或多种商品，通过运输网络高效、及时地运输到一个或多个目的地[3]。多商品流问题是运筹学的经典问题之一，通常被建模成混合整数线性规划模型(Mixed Integer Linear Programming, MILP)。建模方式有两种，分别为基于弧的模型(arc based model)和基于路径的模型(path based model)。目前常用的飞机恢复问题框架中，也大多使用运筹学建模，并使用精确或启发式算法加快问题求解。相关研究可以参考 Rosenberger 等[4]、Eggenberg 等[5]、Liang 等[6]。

虽然现有的框架已经在实际应用中展现了良好的效果，但是当航班网络规模很大时，仍然难以满足在几分钟内得到高质量解的需求。并且，实际场景中往往只是突发出现一个或少数几个不正常航班，但是却对整个航班网络进行建模、求解。这种情况不仅会浪费计算资源，也可能导致原航班计划剧烈改动，进而带来其他方面的成本。因此，也有研究尝试选择航班网络中的部分资源(飞机、机组等)或者称之为子网络，并在子网络中恢复不正常航班。由于问题规模的缩小，求解时间得以大大缩短。较早的文献如 Vos 等的文章[7]，使用一种飞机选择算法对飞机进行选择，从而构成子网络。随着机器学习的兴起，也有研究尝试使用机器学习算法选择子网络，例如 Vink 等[8]使用分类的机器学习算法选择飞机、Eikelenboom 等[9]使用排序的机器学习算法选择飞机、Haider 等[10]对机器学习选择子网络的方法进行完善等等。选择子网络的飞机恢复问题求解框架如图 1 所示，针对某个不正常航班，选择最有可能参与恢复该不正常

航班的数个飞机，只对这些飞机和其航班计划所组成的子网络进行建模、求解，最终输出恢复方案。

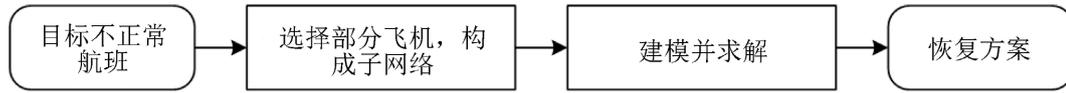


Figure 1. Framework of aircraft recovery based on sub-network selection
图 1. 选择子网络的飞机恢复框架图

但是，选择子网络的飞机恢复方法仍然存在缺陷。针对目标不正常航班选择的子网络中，也许会在其他不正常航班(后文中均称为非目标不正常航班)，并且这些不正常航班无法在当前的子网络中得到恢复。那么这些无法被恢复的航班往往会导致航班取消。如图 2 所示的简单航班子网络为例，图中每个方格代表航班，方格开始和结束的位置分别代表航班计划起飞和降落的时间，方格的宽度表示航班的计划飞行时间。方格的左上角(如 a1)代表出发机场，右下角(如 a2)代表降落机场，中间(如航班 1)代表航班编号。带有阴影的方格(航班 2 和航班 6)代表不正常航班。航班 2 与航班 1 之间的时间间隔太短，不满足最小连接时间；航班 6 与航班 5 的机场无法衔接。每一行的几个航班代表某个时间段内，某架飞机的计划飞行路径，例如航班 1、航班 2、航班 3 是飞机 1 的计划飞行路径。图 2 展示了一个非常简单的选择子网络的不正常航班恢复，目标不正常航班是航班 2，为了恢复航班 2 选择飞机 1 和飞机 2 的航班构成子网络。在该子网络中，航班 2 可以通过与航班 5 进行飞机交换的方式进行恢复，即航班 2 交给飞机 2 执飞，航班 5 交给飞机 1 执飞，其他航班不作调整。但是，在该子网络中存在另一个不正常航班即航班 6，航班 6 无法在该子网络中得到恢复。通过运筹学建模求解的方式，航班 6 与航班 7 只能取消。然而，航班 6 与航班 7 的取消只是因为子网络的解空间较小，并不代表在其他子网络中也无法得到恢复。图 2 是一个简单的案例，现实场景下，由于不正常航班的类型复杂和子网络的规模较大等原因，这种由于子网络的原因导致的额外航班取消也许会更加严重。

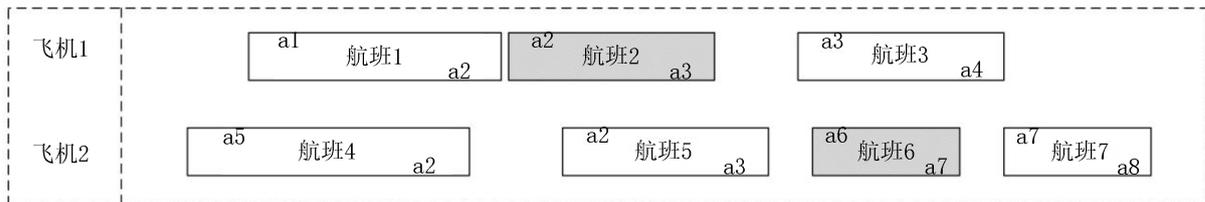


Figure 2. Example of flight recovery in a sub-network
图 2. 子网络航班恢复案例

本文尝试对选择子网络的不正常航班恢复问题模型进行改进，以减少上述额外的航班取消问题。并以机器学习算法选择子网络的飞机恢复方法为例，验证模型改进对额外航班取消问题的影响。本文使用的数据来自实际航班网络，所使用语言为 C++，求解器为 CPLEX12.10 版本，分类机器学习算法为梯度提升算法。

2. 飞机恢复问题模型改进

由于本文采用选择子网络的飞机恢复方法，子网络的网络规模能够得到有效的控制，基于路径的模型相较于基于弧的模型更加适合网络规模较小的场景。并且，基于路径的模型约束较少，很多约束条件可以在路径生成中被考虑到。因此，本文采用基于路径的模型对飞机恢复问题建模，相关模型更详细的描述可参考 Liang 等[6]、Zhang 等[11]。经典的飞机恢复模型如下：

$$\min \sum_{r \in R} \sum_{a \in A} C_{a,r}^O * x_{a,r} + \sum_{f \in F} FCcost_f * y_f \quad (1)$$

$$\sum_{r \in R} \sum_{a \in A} I_{f,r} * x_{a,r} + y_f = 1, \forall f \in F \quad (2)$$

$$\sum_{r \in R} x_{a,r} \leq 1, \forall a \in A \quad (3)$$

$$\sum_{r \in R} \sum_{a \in A} Main_{m,r} * x_{a,r} = 1, \forall m \in M \quad (4)$$

$$\sum_{r \in R} \sum_{a \in A} NumS_{s,r} * x_{a,r} \leq AvailS_s, \forall s \in S \quad (5)$$

$$x_{a,r}, y_f \in \{0,1\}, \forall a \in A, \forall f \in F, \forall r \in R \quad (6)$$

其中, F 是航班集合, f 为航班索引; A 是飞机集合, a 为飞机索引; M 是飞机检修集合, m 为飞机检修索引; S 是时隙(slot)集合, s 为时隙索引; R 是飞机的路径集合, r 为飞机路径索引; 参数 $C_{a,r}^O$ 是指飞机 a 执飞路径 r 的运营成本; 参数 $FCcost_f$ 指航班 f 的取消成本; 参数 $I_{f,r}$ 表示航班 f 是否在路径 r 中, 为 0~1 参数; 参数 $Main_{m,r}$ 表示飞机检修 m 是否在路径 r 中, 为 0~1 参数; 参数 $NumS_{s,r}$ 表示路径 r 使用时隙 s 的次数; 参数 $AvailS_s$ 表示时隙 s 的可用容量; $x_{a,r}$ 为 0~1 决策变量, 表示飞机 a 是否执飞路径 r ; y_f 为 0~1 决策变量, 表示航班 f 是否取消。式(1)为目标函数, 最小化各种成本, 包括飞机执飞路径的成本和航班取消的成本; 式(2)为航班覆盖约束, 表示每一个航班必须出现在某个路径中或者被取消; 式(3)为飞机覆盖约束, 表示每架飞机最多执飞一条路径; 式(4)为飞机检修覆盖约束, 表示每一个飞机检修必须出现在某个路径中; 式(5)为时隙容量约束, 每个时隙的使用次数不得超过时隙容量上限。

如前文所述, 选择子网络的飞机恢复问题方法中, 子网络内非目标不正常航班可能会导致额外的航班取消。针对这个缺陷, 本文对飞机恢复模型进行改进以适应选择子网络的飞机恢复方法。借鉴基于弧的模型中航班复制、超弧等概念[12], 将非目标不正常航班能够影响的最小范围内的航班组合成一个虚拟航班。仍然以一个简单的航班子网络为例, 如图 3 所示, 航班 6 是子网络中的非目标不正常航班, 其影响的最小范围航班组合为航班 5、航班 6, 因此将航班 5、航班 6 及其中间的航班连接时间组合成一个虚拟航班(图中为虚拟航班 8)。虚拟航班与真实航班共同纳入建模求解的过程。若不选择虚拟航班 8, 则航班 6 和航班 7 必须取消; 选择虚拟航班 8, 可以理解为由航班 6 引发的扰动被隐藏起来, 航班 6 是虚拟航班 8 的一部分从而无法在模型求解过程取消, 进而航班 7 也不必取消。

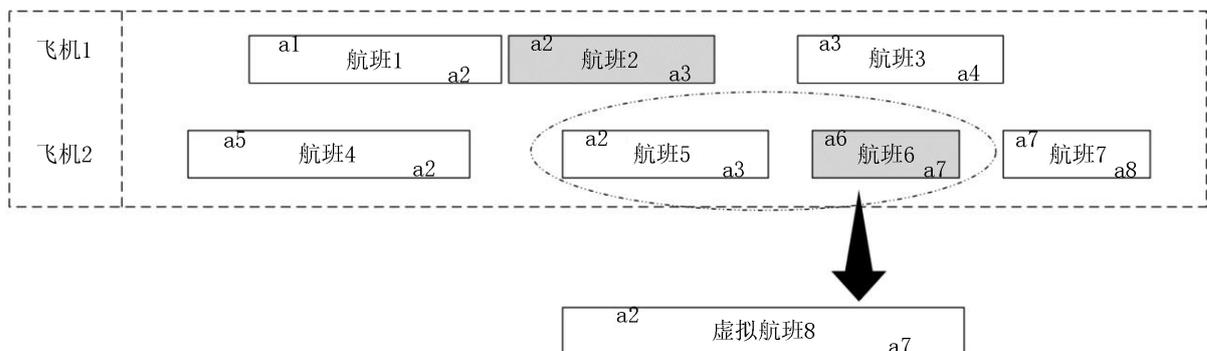


Figure 3. Example of the virtual flight
图 3. 虚拟航班示例图

由于虚拟航班由某些真实航班组成, 虚拟航班与真实航班的关系需要在模型中被考虑, 模型的结构

也需要改变。记虚拟航班的集合为 VF ，索引为 vf ；记组成虚拟航班的真实航班(后文均称为受影响的航班)集合为 CF ，索引为 cf ；其他航班集合仍为 F ，索引为 f 。为了避免额外航班取消的产生，非目标不正常航班都应以虚拟航班的形式存在，以期在后续恢复过程中得到恢复。因此，虚拟航班的航班覆盖约束应当如约束(7)所示。其中， $COM_{cf,vf}$ 为 0~1 参数，表示虚拟航班 vf 是否由航班 cf 组成； VR 为含有虚拟航班的路径集合，索引为 vr ； R 的含义与上文有所不同，表示不包含虚拟航班的路径集合，索引仍然为 r 。需要注意的是，任意受影响的航班 cf 只能出现在某一个虚拟航班 vf 中。约束(7)的含义是，受影响的航班 cf 只能单独出现在某个路径中，或者组成虚拟航班出现在某个路径中，或者被取消。

$$\sum_{r \in Ra \in A} I_{cf,r} * x_{a,r} + \sum_{vf \in VF} \sum_{vr \in VR} \sum_{a \in A} COM_{cf,vf} * I_{vf,vr} * x_{a,vr} + y_{cf} = 1, \forall cf \in CF \quad (7)$$

在实际场景中，航班取消的成本往往远远大于航班延误、飞机交换等运营成本，所以约束(7)能够保证在航班取消和含有虚拟航班的路径中，模型求解结果会选择含有虚拟航班的路径。成本设置方面，虚拟航班是由非目标不正常航班及其影响到的航班组成的一个航班，其内部的干扰成本暂不作考虑，即虚拟航班被视作一个正常航班，其在路径中的运营成本设置与正常航班相同。综上所述，得到改进的飞机恢复模型如下：

$$\min \sum_{r \in Ra \in A} C_{a,r}^O * x_{a,r} + \sum_{vr \in VR} \sum_{a \in A} C_{a,vr}^O * x_{a,vr} + \sum_{f \in F} FCcost_f * y_f + \sum_{cf \in CF} FCcost_{cf} * y_{cf} \quad (8)$$

$$\sum_{r \in Ra \in A} I_{f,r} * x_{a,r} + \sum_{vr \in VR} \sum_{a \in A} I_{f,vr} * x_{a,vr} + y_f = 1, \forall f \in F \quad (9)$$

$$\sum_{r \in Ra \in A} I_{cf,r} * x_{a,r} + \sum_{vf \in VF} \sum_{vr \in VR} \sum_{a \in A} COM_{cf,vf} * I_{vf,vr} * x_{a,vr} + y_{cf} = 1, \forall cf \in CF \quad (10)$$

$$\sum_{r \in R} x_{a,r} + \sum_{vr \in VR} x_{a,vr} \leq 1, \forall a \in A \quad (11)$$

$$\sum_{r \in Ra \in A} Main_{m,r} * x_{a,r} + \sum_{vr \in VR} \sum_{a \in A} Main_{m,vr} * x_{a,vr} = 1, \forall m \in M \quad (12)$$

$$\sum_{r \in Ra \in A} NumS_{s,r} * x_{a,r} + \sum_{vr \in VR} \sum_{a \in A} NumS_{s,vr} * x_{a,vr} \leq AvailS_s, \forall s \in S \quad (13)$$

$$x_{a,r}, x_{a,vr}, y_f, y_{cf} \in \{0,1\}, \forall a \in A, \forall r \in R, \forall f \in F, \forall cf \in CF, \forall vr \in VR \quad (14)$$

式(8)为目标函数，最小化总成本，包括飞机路径运营成本和航班取消成本。式(9)为非受影响航班(正常航班和目标不正常航班)的航班覆盖约束；式(10)为受影响航班的航班覆盖约束；式(11)~(14)与式(3)~式(6)的含义相同。

3. 数据实验

本文需要选择对恢复目标不正常航班有帮助的飞机，使用这些飞机及其计划航班构建子网络。飞机的标签有两类，一类是有帮助的飞机，另一类是没有帮助的飞机，选择飞机的过程被理解为分类问题。因此，本文采用分类的机器学习算法选择飞机。另外，飞机选择问题的特征数据适合用决策树类的机器学习算法进行特征分裂。例如飞机是否出现在目标机场的特征，出现在目标机场才有可能参与恢复问题；相反，未出现在目标机场则参与恢复的可能性低。综上所述，本文采用梯度提升(Gradient Boosting)的分类机器学习算法选择子网络。梯度提升算法将多个弱学习器的预测结果结合起来作为最终预测结果，弱学习器通常为决策树，“梯度”是指优化的方法。关于梯度提升的详细原理，可参考 Friedman [13]和 Chen 等[14]的文章。

本文训练集的数据来源于真实的航班网络数据，训练集数据共 25,941 条。特征选择方面，本文按照

飞机可能的恢复方法(延误、飞机交换)将特征分为体现延误可能的特征(如最长地面停留时间)和体现飞机交换可能的特征(如与机场相关的特征),以及航班网络基本信息的特征(如飞机数量、航班数量等)。经过特征筛选最终选出 69 个特征。训练集数据标注方面,先对整个航班网络进行运筹学建模求解,整理出参与恢复目标不正常航班的飞机,将这些飞机的标签记为 1,其他飞机的标签记为 0。本文使用 python 中 sklearn 框架训练模型,并使用 C++调用模型文件进行飞机预测。测试的案例中可能会存在多个不正常航班,因此采用滚动求解的方式对所有不正常航班进行求解,恢复框架图如图 4 所示。

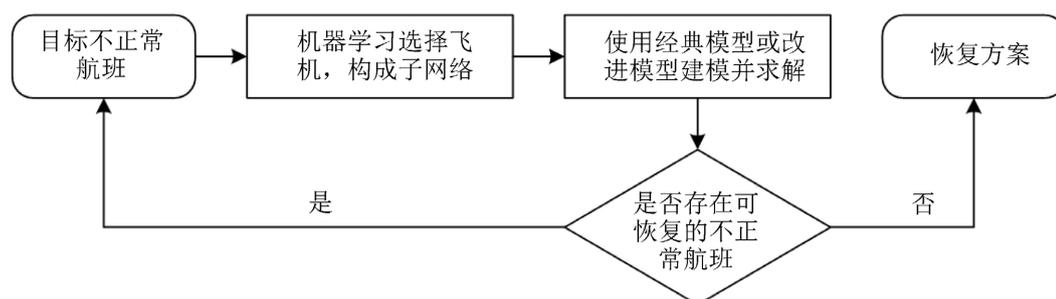


Figure 4. Framework of machine learning for selecting sub-networks and rolling recovery

图 4. 机器学习选择子网络并滚动恢复框架图

本文选择 2 个小规模案例(航班数小于 200)、2 个中等规模案例(航班数在 200 至 800 之间)、2 个大规模案例(航班数在 800 以上)进行实验结果展示。实验结果如表 1 所示。求解时间栏和总恢复成本栏中,改进模型下的百分比数据表示改进模型与经典模型相比,数值减少的比例($\frac{\text{改进模型} - \text{经典模型}}{\text{经典模型}} \times 100\%$)。

Table 1. Experimental results

表 1. 实验结果

案例编号	航班数	航班取消数量		求解时间(s)		总恢复成本	
		经典模型	改进模型	经典模型	改进模型	经典模型	改进模型
1	101	5	2	8.94	9.01 1.00%	601,712	484,594 -19.50%
2	181	4	1	10.22	9.98 -5.28%	1,776,512	1,184,461 -33.33%
3	771	35	13	57.67	62.25 8.12%	846,457	381,552 -54.92%
4	815	8	3	36.95	34.82 -5.76%	169,572	106,616 -37.13%
5	1052	22	10	47.17	52.64 11.60%	1,178,830	586,259 -50.27%
6	2487	23	8	143.49	146.16 1.86%	1,488,348	806,045 -45.84%

首先,表 1 的实验结果显示,航班取消数量上,使用改进模型滚动求解的结果显著优于使用经典模型滚动求解的结果。使用经典模型时,6 个案例的平均航班取消数量为 16.2 个;使用改进模型时,6 个案

例的平均航班取消数量为 6.2 个, 远低于 16.2 个。从结果上也可以看出, 航班取消数量的减少程度与网络图的规模没有直接的关系。例如, 包含 771 个航班的案例 3 中航班取消数量减少 22 个, 而包含 2487 个航班的案例 6 中航班取消数量减少 15 个。其次, 由于航班取消成本一般较大, 因此在总恢复成本方面, 同样是使用改进模型时显著更好。6 个案例中改进模型相对于经典模型的总恢复成本减少比例在 19.50% 到 54.92% 之间, 平均的减少比例为 40.17%。该结果也可以从侧面体现出, 选择子网络的飞机恢复方法中额外的航班取消对最终恢复质量有很大影响。最后, 求解时间上, 由于改进模型中虚拟航班可能会增加滚动求解的轮数, 因此部分案例的求解时间较使用经典模型时稍长。例如, 案例 5 中经典模型下的求解时间 47.17 s, 改进模型下的求解时间 52.64 s, 较经典模型下的结果增加 11.60%。但是时间的绝对值差距很小, 仅为 5.47 s, 符合实际场景的要求。类似的案例还有案例 3。其他案例中, 案例 1 和案例 6 中的求解时间差距很小, 分别为 1.00% 和 1.86%; 案例 2 和案例 4 中, 改进模型下的求解时间甚至略优于经典模型下的求解时间, 分别加快 5.28% 和 5.76%。综上所述, 改进模型在选择子网络的飞机恢复问题中能够有效减少航班取消的数量, 并且对求解时间的影响较小, 验证了改进模型的效果和可行性。

4. 结语

本文针对选择子网络的飞机恢复方法中, 非目标不正常航班可能导致额外航班取消的问题, 对模型进行相应的改进, 增加虚拟航班的概念以适应子网络滚动求解的恢复方法。本文使用实际案例数据对改进模型的效果进行测试, 在小规模、中规模、大规模的案例中, 改进模型均显示出在航班取消数量上的显著优势, 验证了改进模型的有效性。

参考文献

- [1] 朱金福, 等. 航空运输规划[M]. 西安: 西北工业大学出版社, 2009.
- [2] Su, Y., Xie, K., Wang, H., Liang, Z., Art Chaovalitwongse, W. and Pardalos, P.M. (2021) Airline Disruption Management: A Review of Models and Solution Methods. *Engineering*, 7, 435-447. <https://doi.org/10.1016/j.eng.2020.08.021>
- [3] Haghani, A. and Oh, S. (1996) Formulation and Solution of a Multi-Commodity, Multi-Modal Network Flow Model for Disaster Relief Operations. *Transportation Research Part A: Policy and Practice*, 30, 231-250. [https://doi.org/10.1016/0965-8564\(95\)00020-8](https://doi.org/10.1016/0965-8564(95)00020-8)
- [4] Rosenberger, J.M., Johnson, E.L. and Nemhauser, G.L. (2003) Rerouting Aircraft for Airline Recovery. *Transportation Science*, 37, 408-421. <https://doi.org/10.1287/trsc.37.4.408.23271>
- [5] Eggenberg, N., Salani, M. and Bierlaire, M. (2010) Constraint-Specific Recovery Network for Solving Airline Recovery Problems. *Computers & Operations Research*, 37, 1014-1026. <https://doi.org/10.1016/j.cor.2009.08.006>
- [6] 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. <https://doi.org/10.1016/j.trb.2018.05.007>
- [7] Vos, H.M., Santos, B.F. and Omondi, T. (2015) Aircraft Schedule Recovery Problem – A Dynamic Modeling Framework for Daily Operations. *Transportation Research Procedia*, 10, 931-940. <https://doi.org/10.1016/j.trpro.2015.09.047>
- [8] Vink, J., Santos, B.F., Verhagen, W.J.C., Medeiros, I. and Filho, R. (2020) Dynamic Aircraft Recovery Problem—An Operational Decision Support Framework. *Computers & Operations Research*, 117, Article ID: 104892. <https://doi.org/10.1016/j.cor.2020.104892>
- [9] Eikelenboom, B. and Santos, B.F. (2023) A Decision-Support Tool for the Integrated Airline Recovery Using a Machine Learning Resources Selection Approach. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4519717>
- [10] Haider, I., Sen, G., Arsalan, M. and Das, A.K. (2024) Subnetwork Prediction Approach for Aircraft Schedule Recovery. *Engineering Applications of Artificial Intelligence*, 133, Article ID: 108472. <https://doi.org/10.1016/j.engappai.2024.108472>
- [11] Zhang, D., Henry Lau, H.Y.K. and Yu, C. (2015) A Two Stage Heuristic Algorithm for the Integrated Aircraft and Crew Schedule Recovery Problems. *Computers & Industrial Engineering*, 87, 436-453. <https://doi.org/10.1016/j.cie.2015.05.033>

-
- [12] Wei, K. and Vaze, V. (2024) Efficient Construction of Arc-Based Transportation Network Models Using Graph Attention Networks. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4685714>
- [13] Friedman, J.H. (2001) Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, **29**, 1189-1232. <https://doi.org/10.1214/aos/1013203451>
- [14] Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 785-794. <https://doi.org/10.1145/2939672.2939785>