Crew Recovery Using Machine Learning and Optimization

Ahmet Esat Hızır[a], Cynthia Barnhart[a], Vikrant Vaze[b]

[a]Massachusetts Institute of Technology, [b]Dartmouth College

Abstract

Due to the irregular nature of flight operations, airlines need to take a range of actions to recover their aircraft and crew schedules. The limited time frames prevent airlines from using a full-scale optimization approach. Consequently, airlines typically apply recovery solutions that can be far from optimal. This study proposes a practical method that combines machine learning and optimization to find improved recovery solutions. Our procedure is based on the idea that the most effective constraints to add to the recovery models without sacrificing the solution quality, can be determined in advance by leveraging the similarities between disruptions. Our experiments show that, this approach can reduce solution time significantly while still achieving high-quality solutions.

Introduction

Airline scheduling is one of the areas to which operations research methods have been applied successfully. Airlines optimize their aircraft and crew resources but the optimized schedules are rarely operated exactly as planned due to irregularities on the day of operations, such as inclement weather conditions. Airlines monitor their operations and take actions to recover their schedules and minimize the effects of irregularities. This process is generally called *airline recovery* or *disruption management*. While the recovery problems are smaller in size than their planning counterparts, the limited time availability makes it more challenging to use optimization approaches. Hence, airlines usually rely on heuristic solutions or expert judgments (Hassan, Santos, and Vink 2021). One of the primary concerns for researchers studying recovery problems, is keeping the solution run times within practical time limits. A common approach, in both research and practice, is to reduce the problem size by only including a limited number of flights, aircraft and crew to the solution space but this strategy affects the solution quality significantly. An important underutilized opportunity is in leveraging the previous or offline solutions. By detecting the similarities between different disruptions with the help of machine learning (ML) tools and guiding the optimization accordingly, it is possible to outperform the existing strategies to accelerate the optimization without sacrificing the solution quality.

The research presented in this study is one of the first investigations into how ML methods and optimization tools can be combined for airline recovery problems. The contributions can be divided into four main categories. First, we develop a general framework and propose fast and practical solution methods to find better solutions for the crew recovery problem within the limited time frames of the flight operations than other practical approaches by combining ML techniques with optimization tools. Second, the proposed methods have the ability to adapt to the available solution time limit. Third, the proposed methods can yield solutions that are robust to uncertainties. And lastly, the trained classifiers provide practical insights about the flight operations.

Problem Statement

The crew recovery process, attempts to repair the disrupted crew schedules. The recovery decisions may include delaying or canceling flights, re-scheduling crew and calling in reserves (Barnhart 2009). The objective is to minimize the total disruption cost. It is assumed that the total delay of a flight has two components (Lan, Clarke, and Barnhart 2006): 1) *Propagated delay* which is the delay caused by the late arrival of the previous flight operated by the same aircraft or crew, and 2) *Non-propagated delay* (NP) which is the delay caused by other exogenous factors, such as the air traffic conditions. This latter delay depends on the disruption characteristics of the current scenario. Hence, we define the feature set of a disruption scenario as the NP delay estimates of each flight. This feature set is sufficiently detailed to reflect the characteristics of different kinds of disruptions in the network and relatively concise to enable efficient implementation of ML methods.

Solution Methodology

A frequently used modeling approach for airline scheduling and recovery problems is called *string-based modeling* (Barnhart et al. 1998). A *string* is a sequence of flight legs operated by a single crew member. The recovery solutions can be characterized as the set of selected strings. Since a string is a sequence of flights, we can also characterize a recovery solution as the set of consecutive flights — called follow-on (F/O) pairs. The general framework in this study is based on the idea that, under similar circumstances se-

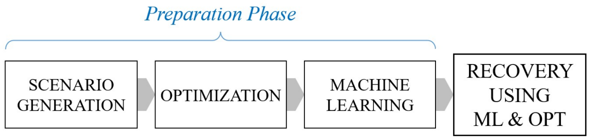

Figure 1: General Flowchart of the Solution Methodology

lected set of F/O pairs may also be similar. So, we can leverage the similarities between the current and the previous disruption instances to find high quality solutions in limited timeframes.

Figure 1 summarizes the general flowchart of the solution methodology. The preparation phase starts with the generation of a set of disruption scenarios based on historical data. Then, all scenarios are solved with a pre-specified optimality gap target. Using the generated database of solutions, binary classification models are trained for a subset of F/O pairs to predict whether they should be in the solution of the given disruption or not. During the day of operations, we fix some of the F/O pairs based on the classifier predictions by adding the corresponding constraints to the model. The remaining reduced problem is sent to the optimizer.

Computational Study

Problem instances and the generated disruption scenarios are based on the actual operational data of the domestic flight network of a major US carrier. Since airlines usually prefer to find a recovery solution within 1 or 2 minutes (Hassan, Santos, and Vink 2021), we focused on shorter timeframes. Figure 2 depicts the performance of the *ML*-based approach under different available time limits. *ML* ## corresponds to *ML*-based method with ## seconds of available time. Y axis is the cost difference with respect to the baseline, which is a 30-minute solution found by the default optimization approach with 0.1% optimality gap target for the full problem. *Threshold value* is used to evaluate the F/O pairs based on the classifier predictions. Lower threshold values imply that higher number of F/O pairs would be fixed. When the available time is longer, there is no need to use lower threshold values and fix too many F/O pairs in advance while for shorter solution time limits, it is better to fix relatively higher proportion of the F/O pairs. The solution quality curves converge into one when the threshold values are set to 0.93 or lower. That is because when a significant portion of the solution is fixed in advance, 30 seconds becomes sufficient to solve all disruption instances even if the available solution time limit is longer. Full scale optimization approach, which keeps the entire feasible solution space, does not generate acceptable or even feasible recovery solutions within the available time limits. The experimental results show that for a given problem and time availability there is a corresponding threshold value that helps to find the best solution.

The results presented above correspond to the cases where the recovery solution is created once for the recovery period and not modified again. It also assumes that the delay predictions are accurate and the airlines take their actions accordingly. But in reality, the recovery solutions are modified

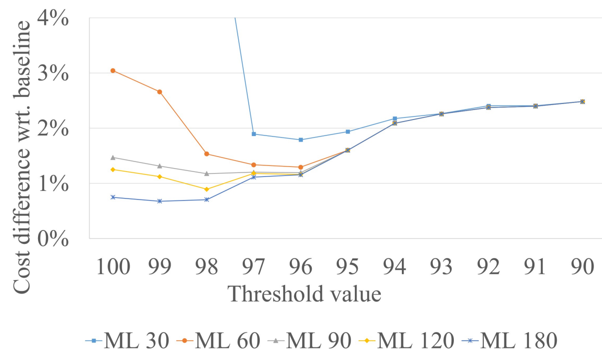

Figure 2: Solution quality curves for different time limits

several times a day and the predictions are never 100% accurate. In order to evaluate the performance of the proposed approach in a more realistic manner, we developed a simulation procedure which allows the recovery decisions to be altered several times during the day of operations while also relaxing the accurate delay predictions assumption. The day of operations is divided into 6 stages (4 hour intervals). An initial recovery solution is created based on the predictions available at the start of the day. Then at the end of each stage, airlines update their delay predictions for upcoming flights and modify the recovery solutions accordingly. It is assumed that airlines have more accurate delay predictions for the flights departing within a short time period and the accuracy of the predictions decreases for the flights beyond. The results after 5 recovery updates, show that ML-based approach has a slightly better overall performance when compared to the initial baseline recovery solution. This is despite the fact that the baseline solutions need significantly longer run times whereas the ML-based solutions are found within 1 minute.

It is important to note that, while the preparation phase is a time consuming task, there are many strategies, like using lower quality solutions for database generation, that can accelerate the preparation without affecting the overall performance of the approach significantly (results omitted due to space constraints).

Conclusion

In this study, we developed a practical approach which combines ML tools and optimization methods to solve crew recovery problems in limited time frames. The results demonstrate that for each problem instance and available solution time limit, there exists a set of size reduction strategies which provides the best solution quality and it is possible to find such effective strategies by incorporating ML methods into the recovery process. Our experiments also showed that the ML-based solutions are at least as robust to inaccurate flight delay predictions as traditional optimization-based solutions while requiring significantly less solution time. The next step in our research is to tackle the integrated aircraft and crew recovery problem where a straightforward application of the presented ideas is not possible due to the size of the integrated problem.

References

Barnhart, C. 2009. Irregular operations: Schedule recovery and robustness. *The global airline industry*, 23: 253–274.

Barnhart, C.; Boland, N. L.; Clarke, L. W.; Johnson, E. L.; Nemhauser, G. L.; and Shenoi, R. G. 1998. Flight string models for aircraft fleeting and routing. *Transportation science*, 32(3): 208–220.

Hassan, L.; Santos, B. F.; and Vink, J. 2021. Airline disruption management: A literature review and practical challenges. *Computers & Operations Research*, 127: 105137.

Lan, S.; Clarke, J.-P.; and Barnhart, C. 2006. Planning for robust airline operations: Optimizing aircraft routings and flight departure times to minimize passenger disruptions. *Transportation science*, 40(1): 15–28.