Solving the Airline Crew Pairing Problem using Subsequence Generation

Matias Sevel Rasmussen · David M. Ryan · Richard M. Lusby · Jesper Larsen

Abstract Good and fast solutions to the airline crew pairing problem are highly interesting for the airline industry, as crew costs are the biggest expenditure after fuel for an airline. The crew pairing problem is typically modelled as a set partitioning problem and solved by column generation. However, the extremely large number of possible columns naturally has an impact on the solution time.

In this work in progress we severely limit the number of allowed subsequent flights, i.e. the subsequences, thereby significantly decreasing the number of possible columns. Set partitioning problems with limited subsequence counts are known to be easier to solve, resulting in a decrease in solution time.

The problem though, is that a small number of deep subsequences might be needed for an optimal or near-optimal solution and these might not have been included by the subsequence limitation. Therefore, we try to identify or generate such subsequences that potentially can improve the solution value.

Keywords Airline crew pairing · Crew pairing · Subsequence generation · Column generation · Limited subsequence

M.S. Rasmussen
Department of Management Engineering, Technical University of Denmark, Produktionstorvet, Building 424, 2800 Kgs. Lyngby, Denmark
Tel.: +45-45254442
Fax: +45-45933435
E-mail: mase@man.dtu.dk

D.M. Ryan
Department of Engineering Science, University of Auckland, 70 Symonds Street, Auckland 1010, New Zealand

R.M. Lusby
Department of Management Engineering, Technical University of Denmark, Produktionstorvet, Building 424, 2800 Kgs. Lyngby, Denmark

J. Larsen
Department of Management Engineering, Technical University of Denmark, Produktionstorvet, Building 424, 2800 Kgs. Lyngby, Denmark

1 Introduction

Crew costs are the second largest expenditure in the airline industry. Only fuel costs are larger, see [1]. Therefore, airline crew scheduling has received a lot of attention in the literature, and consequently, optimisation is heavily used by the airlines. The airline crew pairing problem which is dealt with in this work is a part of a larger series of optimisation problems that together produce the schedule for an individual crew member. In [1] a recent survey of airline crew scheduling can be found.

A *pairing* or a *tour-of-duty* is a sequence of flights which can be flown be a crew member. A pairing must start and end at the same crew base and comply with several rules and regulations in order to be feasible. The *airline crew pairing problem* then finds the set of pairings that exactly covers all flights at minimum costs.

2 Solution Method

The pairing problem is modelled as a set partitioning problem. Each row corresponds to a flight and each column corresponds to a pairing. Let m be the number of rows and n be the number of columns, and let c_j be the costs of column $j \in \{1, \ldots, n\}$. The entries of $\mathbf{A}$, a_{ij}, are one if column $j \in \{1, \ldots, n\}$ covers row $i \in \{1, \ldots, m\}$ and zero otherwise. The decision variables x_j for $j \in \{1, \ldots, n\}$ are binary. The mathematical programme can be written as

$$\begin{array}{ll} \text{minimise} & \mathbf{c}^\top \mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{1} \\ & \mathbf{x} \in \{0,1\}^n \ . \end{array}$$

The number of possible pairings in the set partitioning formulation is very large, so the pairings are typically only enumerated implicitly by column generation. In the present approach we will, however, not perform column generation, but *subsequence generation*.

The *subsequences* for a flight f are the set of subsequent flights that can follow f in a feasible way in a pairing. In general terms for a zero-one matrix $\mathbf{A}$, the *subsequence count*, $\mathrm{SC}(s)$, for any row s is given by

$$\mathrm{SC}(s) = |\{t : [a_{sj} = 1, a_{ij} = 0 \text{ for } s < i < t, a_{tj} = 1], j = 1, \ldots, n\}| \ .$$

Matrices with $\mathrm{SC}(s) \leq 1$ for all $s \in 1, \ldots, m$ are said to have *unique subsequence*, and such matrices are balanced, see [2]. Exploiting results from graph theory, we know that the LP relaxation of an SPP with a balanced $\mathbf{A}$ matrix has an integral optimal solution. Also shown in [2], the closer we get towards unique subsequence, the closer we get to naturally integral LP solutions.

Therefore, we severely limit the subsequence count for each flight when generating pairings. This results in significantly fewer possible pairings and, as mentioned, fewer fractions when solving the LP relaxation. The disadvantage, however, is that we might exclude some optimal subsequences. To remedy this, we use the information in the dual vector to identify missing subsequences. The dual vector is passed on to one or several column generators that produce negative reduced costs columns on a richer set of subsequences. These columns are analysed in order to identify potentially “good” subsequences.

The goal is, of course, to be able to, as early as possible, identify the subsequences that will end up in the optimal or near-optimal solution. Whenever a subsequence is identified as a potentially "good" subsequence, the whole set of columns which include the new subsequence are added to the LP. Furthermore, to prevent the LP from growing too big, subsequences can be removed from the LP, i.e. the set of columns containing the subsequence are removed.

3 Computational Results

In order to gain better understanding of the method, we have generated a set of set partitioning instances with a cost structure reflecting the cost structure from crew pairing problems. The results from the generated instances indicate that we can identify the missing subsequences in reasonable time.

Currently, we are in the process of performing tests on a set of real-world crew pairing problem instances.

4 Future Work

The results this far clearly justify further development. Firstly, as mentioned, real-world crew pairing problems will be tackled. Secondly, the subsequence identification process has room for improvements. Thirdly, the method is based on the dual vector, therefore dual stabilisation is likely to speed up the method, as dual stabilisation would make the duals more reliable. Lastly, the column generators can be run in parallel on different processors.

References

1. Gopalakrishnan, B., Johnson, E.L. (2005). Airline crew scheduling: state-of-the-art. *Annals of Operations Research*, 140, 305–337
2. Ryan, D.M., Falkner, J.C. (1988). On the integer properties of scheduling set partitioning models. *European Journal of Operational Research*, 35, 442–456