Les Cahiers du GERAD ISSN: 0711–2440

Combining Benders decomposition and column generation for integrated crew pairing and personalized crew assignment problems

V. Zeighami,
F. Soumis

G–2017–41

May 2017

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Combining Benders decomposition and column generation for integrated crew pairing and personalized crew assignment problems

Vahid Zeighami

François Soumis

GERAD & Department of Mathematics and Industrial Engineering, Polytechnique Montréal (Québec) Canada, H3C 3A7

vahid.zeighami@gerad.ca
francois.soumis@gerad.ca

May 2017

Les Cahiers du GERAD

G–2017–41

Abstract: The airline crew scheduling problem, because of its size and complexity, is usually solved in two phases: the crew pairing problem and the crew assignment problem. A pairing is a sequence of flights, connections, and rests starting and ending at the same crew base. The crew pairing problem consists of determining a minimum-cost set of feasible pairings such that each flight is covered exactly once. In the crew assignment problem, the goal is to construct monthly schedules from these pairings for pilots and copilots independently, while respecting all the safety and collective agreement rules. However, this sequential approach may lead to significantly suboptimal solutions since it does not take into account the crew assignment constraints and objective during the building of the pairings. In this paper, first, we propose an extension of the crew pairing problem that incorporates pilot and copilot vacation requests at the crew pairing stage. Second, we introduce a model that completely integrates the crew pairing and crew assignment problems simultaneously for pilots and copilots. To solve this integrated problem, we develop a method that combines Benders decomposition and column generation. We conduct computational experiments with real-world data from a major US carrier.

Keywords: Airline crew scheduling, Benders decomposition, column generation

Acknowledgments: This research was supported by the Natural Sciences and Engineering Research Council of Canada and a collaborative R&D grant from AD OPT, a division of Kronos.

1 Introduction

The airline planning problem is one of the most challenging problems in the field of operations research. Because of its size and complexity, this problem is typically solved sequentially in four main steps: *flight scheduling, fleet assignment, aircraft routing,* and *crew scheduling*. The flight scheduling problem determines a set of flights with specific departure and arrival times, with the goal of maximizing the expected profit. The fleet assignment problem determines the type of aircraft (Boeing 737-500, Airbus A320, etc.) to assign to each scheduled flight so as to maximize the profit, based on the different capacities and the number of available aircraft. The aircraft routing problem assigns individual aircraft to flights while satisfying maintenance requirements.

After these steps, the airlines solve the crew scheduling problem. Since the crew cost is the largest after the fuel cost, crew scheduling is one of the most important problems in airline planning. The crew scheduling problem determines crew schedules that cover all the scheduled flights and satisfy the constraints. This problem is usually solved in two steps: the pairing problem and the assignment problem.

A pairing is a sequence of flights, connections, and rests starting and ending at the same crew base. The crew pairing problem generates a set of pairings given the scheduled flights such that the cost of the pairings is minimized and all the flights are covered exactly once. The crew assignment problem builds monthly schedules for each crew member given the set of pairings such that every pairing is covered exactly once.

There are three different approaches for the assignment problem: the bidline approach, the personalized approach with seniority order, and the personalized approach with a global objective. In the past, North American airlines typically applied the bidline approach. This produces anonymous monthly schedules, and the (co)pilots select their schedules in seniority order. The personalized approach with strict seniority, which is becoming increasingly popular in North America, sequentially maximizes the satisfaction of the employees in decreasing order of seniority. In the personalized approach with a global objective, which is often employed by European airlines and is becoming more popular with American airlines, all the monthly schedules are produced simultaneously to maximize an objective function. This may be the sum of the personal satisfactions with the possible addition of a term removing any employee discrimination.

The crew pairing and assignment problems are usually formulated via set partitioning or set covering models with additional constraints. In the pairing problem, the variables represent the feasible pairings and the constraints ensure that each flight is covered. In the assignment problem, the variables are feasible schedules and the constraints ensure that each pairing is covered.

Various methods and models have been proposed for the crew scheduling problem. We refer the reader to the recent surveys by [22] and [17]. The literature mostly focuses on either crew pairing or assignment; there are a few studies of integrated models and solution methodologies. [26] and [5] proposed a solution technique for crew pairing based on Lagrangian relaxation and subgradient optimization. Solution methodologies based on column generation (CG) were developed by [13], [3], [39], [23], and [38]. Heuristic algorithms were developed by [16], [1], [2], [7], and [18]. [34] used a rolling-horizon approach based on CG to solve the pairing problem.

Several approaches have been proposed for the assignment problem. [20] proposed using a priori CG for the bidline assignment problem with the goal of minimizing the number of bidlines and maximizing the covered credit time. [9] developed a bidline generator system based on simulated annealing. To solve the bidline-generation problem at Delta Air Lines, [11] proposed a two-phase approach based on genetic algorithms. [40] presented a three-phase method using mixed integer programming for bidline generation. Using CG, [14] solved the preferential bidding problem with seniority to construct personalized monthly schedules for pilots and officers. [25] and [8] described two heuristic algorithms for the SPP-based bidline scheduling problem. [22] formulated the personalized crew scheduling problem via set covering and used CG as the solution method.

While solving the airline planning problem sequentially may lead to poor solutions, only a few researchers have investigated integrating two or more of these stages. Integrated fleet assignment, maintenance routing, and crew pairing was considered by [29]. Integrated aircraft routing and crew pairing was proposed by [12],

[4], [28], and [10]. Integrated flight scheduling, aircraft routing, and crew pairing was considered by [24]. [35] and [15] considered the integration of fleet assignment and crew pairing. Integrated aircraft routing, crew scheduling, and flight retiming was studied by [27]. [36] considered the integration of the fleet assignment, aircraft routing, and crew pairing problems. Integrated crew pairing and assignment was studied by [41]; they consider small problems with 59 to 210 flights and a few flights per pairing. They use a global formulation and solve it with CPLEX. [37] proposed three heuristic approaches based on the genetic algorithm to integrate crew pairing and assignment. [32, 33] studied integrated crew pairing and assignment where the objective is to minimize the total cost and the number of pilots. They consider the bidline approach, and they combine dynamic constraint aggregation with CG. [21] presented a heuristic algorithm for integrated crew pairing and personalized assignment. The algorithm alternates between the pilot and copilot problems, trying to obtain common duties and pairings.

In this paper, we consider the crew pairing and personalized crew assignment problems with a global objective, where each pairing requires one pilot and one copilot, and each (co)pilot requests two vacations per month. Using a sequential approach to solve these problems produces the same pairings for (co)pilots and reduces the propagation of the perturbation and the complexity of the process, but the pairings obtained can be far from optimal because the schedule constraints and objectives are not taken into account. Hence, it is difficult to maximize the satisfaction of crew preferences.

In summary, this paper makes the following contributions. We propose a novel extended crew pairing model that considers (co)pilots' vacation requests (VRs) to obtain better pairings for the crew assignment step. We present an aggregation rule to reduce the number of constraints. Based on the proposed model, we introduce a novel integrated model for crew pairing and personalized assignment. We develop a solution methodology based on Benders decomposition and CG that alleviates the computational difficulties arising from the large number of variables. The pairings are generated by the Benders master problem, which is composed of a CG master problem (MMP) and a set of CG subproblems (MSP), one for the pairings starting on each day at each base. The monthly schedules for (co)pilots are generated by the Benders subproblems, which consist of a CG master problem (SMP) and a set of CG subproblems (SSP), one for each crew member. The challenge is that the Benders cuts containing pairing variables must be placed in the MMP, and their dual variables cannot be transferred to the MSP to evaluate the cost values of the new column. To overcome this difficulty, we introduce the concept of weak Benders cuts that do not take into account the dual variable corresponding to the pairings. We prove that with these weak Benders cuts and the suggested dual solution of the SMP, combined Benders decomposition and CG reach optimality. We also use a warm-start strategy that relies on the solution process of Benders decomposition to speed up the convergence.

The remainder of this paper is organized as follows. Section 2 gives the problem statement, and Section 3 describes the extension of the crew pairing problem. Section 4 presents the integrated mathematical formulation, and Section 5 describes the solution methodology. Section 6 presents the data sets, experiments, and computational results. Finally, Section 7 provides concluding remarks and discusses future research.

2 Problem statement

In this section, we give detailed definitions of the crew pairing and personalized crew assignment problems. The definition and feasibility rules for pairings and schedules can differ from one airline to another. We use a subset of the feasibility rules commonly used in the literature, such as those of [34] and [22].

2.1 Crew pairing problem

Given a set of flights and crew bases, the crew pairing problem finds a set of minimum-cost pairings such that each flight is covered by exactly one pairing and each pairing starts and ends at the same base. A pairing is a sequence of one or more duties and overnight stops. A duty is a working day for a crew member and consists of a sequence of consecutive flights or deadheads separated by rest periods. A deadhead is a flight where the crew travels as passengers for repositioning purposes. A flight is defined by departure/arrival locations and fixed departure/arrival times. There is a maximum of 5 landings per duty. The briefing and debriefing times at the beginning and end of each duty are 60 and 30 minutes, respectively. The maximum pairing duration is 4 days, the maximum number of duties per pairing is 4, the minimum connection time between two consecutive flights in a duty is 30 minutes, and the minimum connection time between two consecutive duties is 9.5 hours. A crew member must work between 4 and 8 hours in a duty, and the length of a duty cannot exceed 12 hours. The cost of a pairing has a complicated structure with three components: the cost of waiting times, the deadheading cost, and the total cost of the duties in the pairing. We use the pairing cost of [30].

2.2 Personalized crew assignment problem

We assume that there is a fixed number of (co)pilots at each base. In our test, each (co)pilot requests two vacations per month. We consider the (co)pilot assignment problems as personalized assignment problems. We build monthly schedules for the (co)pilots that cover all the pairings and satisfy the maximum number of VRs. There is a maximum of 85 flying hours per month and a maximum of 6 consecutive working days. In each schedule, the pairings are separated by two different rest times: the day-off rest and the rest between two consecutive pairings. The minimum and maximum rest times between any two consecutive pairings are 8 and 12 hours, respectively. A penalty cost is associated with each unsatisfied VR and each uncovered pairing. The objective function minimizes the cost of the uncovered pairings and unsatisfied VRs.

3 Extension of the crew pairing problem

The basic crew pairing problem contains the flight covering constraints. We extend this model to consider several additional factors. We ensure that the number of active pairings does not exceed the number of available (co)pilots. Also, we ensure that it is possible to add the minimum rest period after each pairing and the required day-off without exceeding the number of available (co)pilots. Furthermore, we consider the VRs. In addition, the pairings contained in each schedule are separated by two different rest times. The first links the end of the pairing at the base to the start of the pairings whose departure times are greater than or equal to the arrival time of the relevant pairing plus the minimum rest time (8 hours in our tests). The second rest time links the end of the pairing to the first midnight at the base at the start of a day-off or vacation. We add a set of flow conservation constraints to the pairing problem, as shown in Figure 3. A flow equal to the number of available (co)pilots starts at the beginning of the month and travels in a network with the pairings, night rests, and vacations as its arcs. This network permits us to introduce the bonus for the VRs in the pairing problem.

3.1 Aggregation rules

To reduce the number of flow conservation constraints, we define an aggregation rule for the arrival and departure flights of each base and each day. A sequence of flights arriving at a base on a given day can be aggregated if there are no departing flights within 8 hours of the associated arrival times. Similarly, a sequence of departing flights can be aggregated if there are no arriving flights in the 8 hours before the associated departure times. Figure 1 shows 11 arriving flights (terminating at black nodes) and 12 departing flights (starting at red nodes) for two consecutive days at a base. Figure 2 illustrates the aggregated arrival and departure nodes. The arriving flights are aggregated into four groups (terminating at blue nodes). The departing flights are likewise aggregated into four groups (starting at orange nodes).

3.2 Network representation of flow conservation

Figure 3 shows the flow conservation constraints that consider (co)pilot availability and VRs during the construction of the pairings. We assume that the vacations start and end at the base, and the vacations occur between two midnights.

After the aggregation, the nodes are sorted in chronological order. Let A_b and D_b be the resulting ordered sets of aggregated arrival and departure nodes at base $b \in B$. A *vacation* arc links two midnight nodes if there is a corresponding VR. For a given node $i \in A_b$, a *midnight* arc links this node to the earliest node

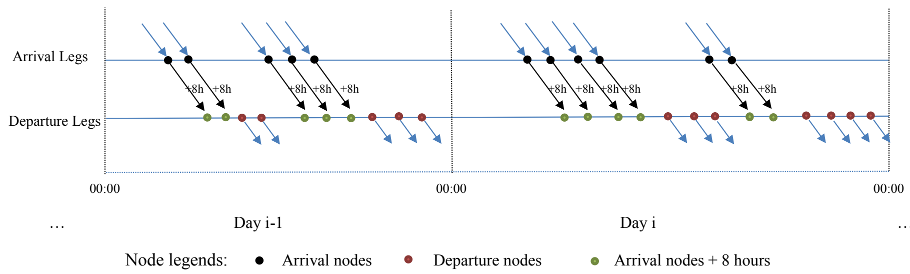

Figure 1: Example of arriving and departing flights at a base.

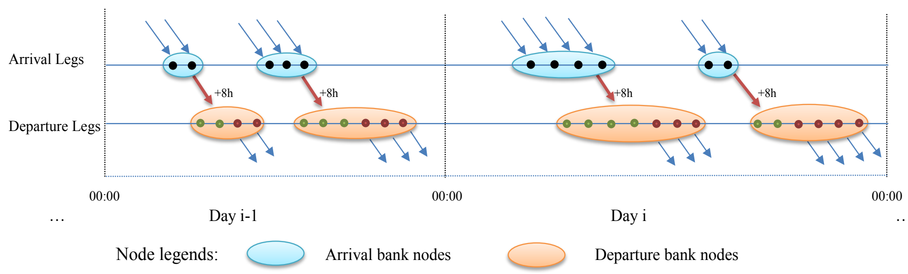

Figure 2: Example of node aggregation at a base.

in D_b, and a *start of vacation* arc links this node to the first midnight node if there is at least one vacation arc that starts from this midnight. Each node in D_b is linked to *waiting* arcs. For a given midnight node $i \in \{1, \ldots, 31\}$, if there is at least one vacation arc that ends at this midnight, an *end of vacation* arc links this node to the earliest node in D_b. There are two depots in the network. The first corresponds to the pilots and the second corresponds to the copilots at each base.

4 Mathematical formulation

This section introduces the notation and describes the integrated crew schedules model that includes the extended crew pairing and personalized assignment models. To solve these problems we propose three approaches, which is described in Section 5. They are called the sequential (SEQ) approach, the sequential approach with the extended crew pairing problem (SEQEP), and the integrated (INT) approach.

4.1 Crew Pairing Notation

Let B be the set of all bases, P_b be the set of feasible pairings in base $b \in B$, and F be the set of scheduled flights to be covered. We denote the flights contained in pairing p by F_p. Let L_b^j be the set of pilots ($j = 1$) or copilots ($j = 2$) at base $b \in B$. Let V_l be the set of VRs for (co)pilot $l \in L_b^j$ at base $b \in B$. The variable y_p is 1 if pairing $p \in P_b$ is chosen and 0 otherwise. The variable e_f is 1 if flight $f \in F$ is not covered and 0 otherwise. Let c_f be the penalty cost for uncovered flight $f \in F$. The parameter ρ_f is the duration of flight f. The parameter e_f^p is 1 if flight f is covered by pairing p and 0 otherwise. The parameter E_b is the maximum credited flying time for base $b \in B$. We apply the negative cost (bonus) c_v for the coverage of $v \in V_l$ to help maximize the number of satisfied VRs. The cost of pairing $p \in P$ is c_p.

4.2 Flow Conservation Notation

We denote by A_b and D_b the sets of aggregated arrival and departure nodes in base $b \in B$. The set of midnight nodes is $I = \{1, 2, \ldots, 31\}$. Let E_b^j be the set of outgoing arcs from midnight nodes I in base $b \in B$ to nodes in D_b that could start pairings for pilots ($j = 1$) or copilots ($j = 2$). Let G_b^j be the set of outgoing arcs from nodes in A_b in base $b \in B$ to midnight nodes I that could start vacations for pilots ($j = 1$) or copilots ($j = 2$). Let M_b^j be the set of outgoing arcs from arrival nodes to departure nodes D_b that could

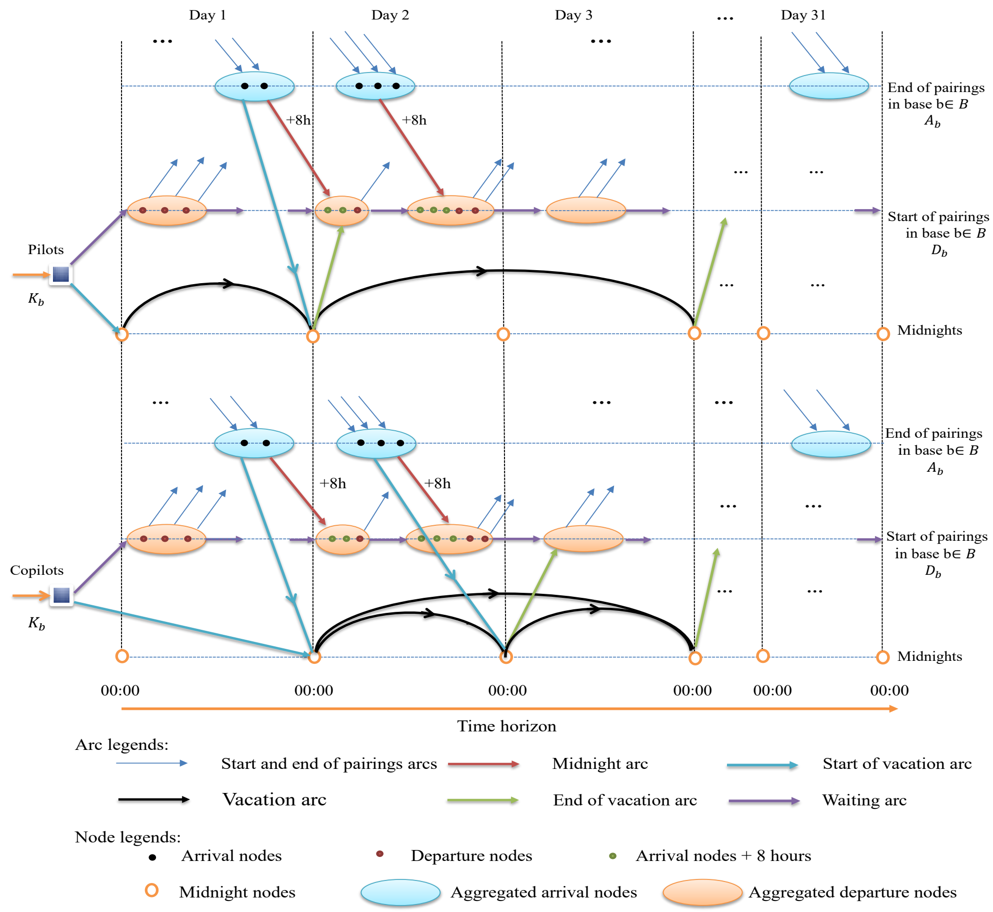

Figure 3: Network to incorporate pilot and copilot availability and VRs in crew pairing problem.

start a midnight for pilots ($j = 1$) or copilots ($j = 2$). The integer variables u_e indicate the flow on the outgoing arcs $e \in E_b^j$. Variable r_v is 1 if $v \in V_l$ is covered and 0 otherwise. Integer variable w_i^j indicates the flow on the waiting arc i between departure nodes i and $i+1$ in base $b \in B$. Integer variable h_g represents the flow on the outgoing arcs $g \in G_b^j$. We denote by t_m the flow on the outgoing arcs $m \in M_b^j$. The fixed value k_b is the number of available (co)pilots at base $b \in B$.

There are several binary parameters. Parameter α_i^p is 1 if end of pairing $p \in P_b$ is incoming at node $i \in A_b$. Parameter η_i^p is 1 if start of pairing $p \in P_b$ is outgoing at node $i \in D_b$. Parameter γ_i^g is 1 if arc $g \in G_b^j$ is outgoing from node $i \in A_b$. Parameter π_i^g is 1 if arc $g \in G_b^j$ is incoming at midnight node $i \in I$. Parameter τ_i^m is 1 if arc $m \in M_b^j$ is outgoing from node $i \in A_b$. Parameter ζ_i^m is 1 if arc $m \in M_b^j$ is incoming at node $i \in D_b$. Parameter ϑ_i^e is 1 if arc $e \in E_b^j$ is incoming at node $i \in D_b$. Parameter μ_i^e is 1 if arc $e \in E_b^j$ is outgoing from midnight node $i \in I$. Finally, parameter e_i^v is 1 if $v \in V_l$ starts on day $i \in I$ and -1 if it finishes on this day.

4.3 Pilot (Copilot) Assignment Notation

We define S_l to be the set of feasible schedules for (co)pilot $l \in L_b^j$. Variable x_s is 1 if schedule $s \in S_l$ is chosen and 0 otherwise. Variable e_p^j is 1 if $p \in P$ is not covered by pilots ($j = 1$) or copilots ($j = 2$) and 0 otherwise. Variable e_v is 1 if $v \in V_l$ is not covered and 0 otherwise. Let $\bar{c}_p$ be the penalty cost for uncovered pairing $p \in P$. Let $\bar{c}_v$ be the penalty cost for uncovered $v \in V_l$. The binary parameter e_p^s is 1 if $p \in P$ is covered by personalized schedule $s \in S_l$. The binary constant V_v^s is 1 if $v \in V_l$ is covered by schedule $s \in S_l$.

Using this notation, our integrated crew scheduling model is as follows:

$$\text{minimize} \left\{ \sum_{b\in B}\sum_{p\in P_b} c_p y_p + \sum_{f\in F} c_f e_f + \sum_{j=1}^{2}\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l} c_v r_v + \sum_{j=1}^{2}\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l} \bar{c}_v e_v + \sum_{j=1}^{2}\sum_{b\in B}\sum_{p\in P_b} \bar{c}_p e_p^j \right\} \quad (1)$$

subject to

Crew pairing constraints:

$$\sum_{b\in B}\sum_{p\in P_b} e_f^p y_p + e_f = 1 \qquad \forall f \in F \qquad (2)$$

$$\sum_{p\in P_b}\sum_{f\in F_p} \rho_f y_p \le E_b \qquad \forall b \in B \qquad (3)$$

Flow conservation constraints to extend pairing model:

$$\sum_{p\in P_b} \alpha_i^p y_p - \sum_{g\in G_b^j} \gamma_i^g h_g - \sum_{m\in M_b^j} \tau_i^m t_m = 0 \qquad \forall j \in \{1,2\}, \forall b \in B, \forall i \in A_b \qquad (4)$$

$$-\sum_{p\in P_b} \eta_i^p y_p + \sum_{e\in E_b^j} \vartheta_i^e u_e + \sum_{m\in M_b^j} \zeta_i^m t_m + w_{i-1}^j - w_i^j = 0 \qquad \forall j \in \{1,2\}, \forall b \in B, \forall i \in D_b \qquad (5)$$

$$-\sum_{e\in E_b^j} \mu_i^e u_e + \sum_{g\in G_b^j} \pi_i^g h_g + \sum_{l\in L_b^j}\sum_{v\in V_l} e_i^v r_v = 0 \qquad \forall j \in \{1,2\}, \forall b \in B, \forall i \in I \qquad (6)$$

$$w_{i_0'}^j + h_{i_0} = k_b \qquad \forall j \in \{1,2\}, \forall b \in B, i_0' \in D_b, i_0 \in G_b^j \qquad (7)$$

Pilot (j=1) and Copilot (j=2) assignment constraints:

$$\sum_{l\in L_b^j}\sum_{s\in S_l} e_p^s x_s + e_p^j \ge y_p \qquad \forall j \in \{1,2\}, \forall b \in B, \forall p \in P_b \qquad (8)$$

$$\sum_{s\in S_l} v_v^s x_s + e_v \ge r_v \qquad \forall j \in \{1,2\}, \forall b \in B, \forall l \in L_b^j, \forall v \in V_l \qquad (9)$$

$$\sum_{s\in S_l} x_s \le 1 \qquad \forall j \in \{1,2\}, \forall b \in B, \forall l \in L_b^j \qquad (10)$$

Variable declarations:

$$y_p \in \{0,1\} \qquad \forall b \in B, \forall p \in P_b \qquad (11)$$

$$e_f \in \{0,1\} \qquad \forall f \in F \qquad (12)$$

$$w_i^j \in \mathbb{Z} \qquad \forall j \in \{1,2\}, \forall b \in B, \forall i \in D_b \qquad (13)$$

$$u_e \in \mathbb{Z} \qquad \forall j \in \{1,2\}, \forall b \in B, \forall e \in E_b^j \qquad (14)$$

$$h_g \in \mathbb{Z} \qquad \forall j \in \{1,2\}, \forall b \in B, \forall g \in G_b^j \qquad (15)$$

$$t_m \in \mathbb{Z} \qquad \forall j \in \{1,2\}, \forall b \in B, \forall m \in M_b^j \qquad (16)$$

$$r_v \in \{0,1\} \qquad \forall j \in \{1,2\}, \forall b \in B, \forall l \in L_b^j, \forall v \in V_l \qquad (17)$$

$$e_v \in \{0,1\} \qquad \forall j \in \{1,2\}, \forall b \in B, \forall l \in L_b^j, \forall v \in V_l \qquad (18)$$

$$e_p^j \in \{0,1\} \qquad \forall j \in \{1,2\}, \forall b \in B, \forall p \in P_b \qquad (19)$$

$$x_s \in \{0,1\} \qquad \forall j \in \{1,2\}, \forall b \in B, \forall l \in L_b^j, \forall s \in S_l \qquad (20)$$

The objective function (1) minimizes the total cost of the pairings and maximizes the number of the satisfied VRs. Constraints (2) ensure that each scheduled flight is included in exactly one pairing. Constraints (3) specify an upper bound on the total credited flying time per base. Constraints (4) impose flow conservation for arrival nodes. Constraints (5) enforce flow conservation for departure nodes, and Constraints (6) impose flow conservation for midnight nodes. Constraints (7) ensure that the number of (co)pilots assigned does not exceed the number available at the base. In constraints (7), the indexes $i_0^{\prime}$ and i_0 indicate the first waiting arc and the first start-of-vacation arc at each base, respectively (see Figure 3). Constraints (8) are linking constraints ensuring that pilots and copilots stay together. These constraints should be equalities, but in Proposition 1 we show that the optimal objective value of both models is the same. The linking constraints (9) impose the VRs for each (co)pilot. Constraints (10) guarantee that at most one schedule is chosen for each (co)pilot. The integrality conditions are defined by constraints (11)–(20). Figure 4 shows the block diagonal structure of the integrated model. The block diagonal can be divided into three groups of rows and three groups of columns. The first group of rows relates to the pairing problem, the second relates to pilot assignment, and the third relates to copilot assignment. The pairing variables form the first group of columns, the flow conservation variables form the second, and the assignment variables form the third.

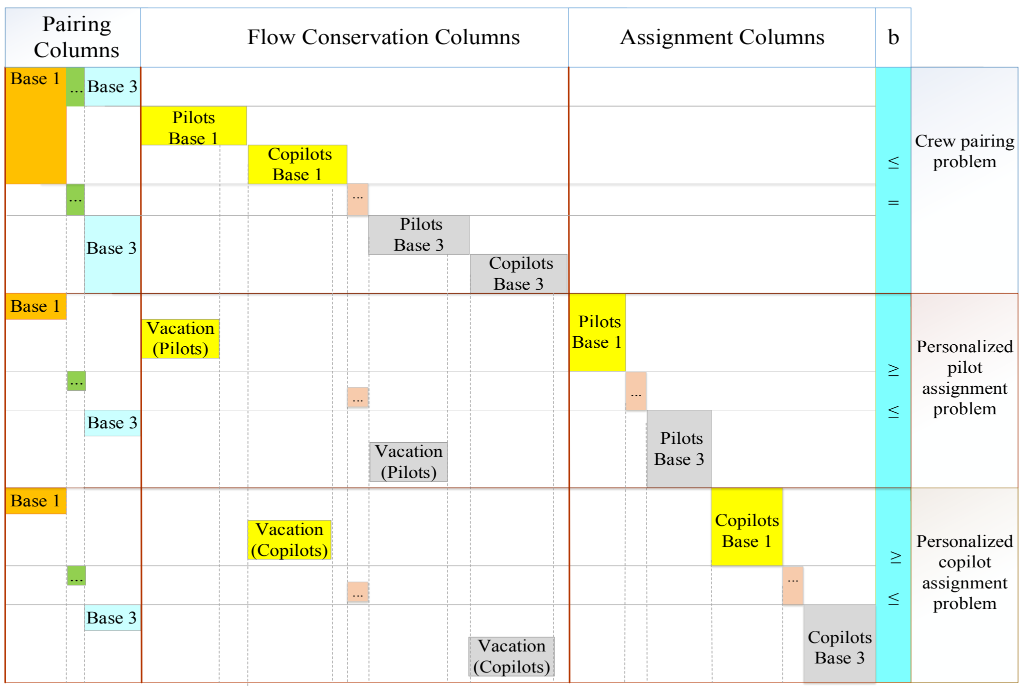

Figure 4: Block diagonal structure of integrated model.

Proposition 1 *Consider the following pilot (j=1) assignment models:*

$$\text{(P1) minimize} \sum_{b\in B}\sum_{l\in L_b^1}\sum_{v\in V_l}\bar{c}_v e_v + \sum_{b\in B}\sum_{p\in P_b}\bar{c}_p e_p^1 \tag{21}$$

subject to (9)–(10), (18)–(20), and

$$\sum_{l\in L_b^1}\sum_{s\in S_l} e_p^s x_s + e_p^1 \geq y_p \qquad \forall b\in B, \forall p\in P_b \tag{22}$$

$$\text{(P2) minimize} \sum_{b\in B}\sum_{l\in L_b^1}\sum_{v\in V_l}\bar{c}_v e_v + \sum_{b\in B}\sum_{p\in P_b}\bar{c}_p e_p^1 \tag{23}$$

subject to (9)–(10), (18)–(20), and

$$\sum_{l\in L_b^1}\sum_{s\in S_l} e_p^s x_s + e_p^1 = y_p \qquad \forall b\in B, \forall p\in P_b \tag{24}$$

The optimal objective values of models (P1) and (P2) are equal.

Proof. Let X_1 and X_2 be the feasible domains of models (P1) and (P2), respectively. Clearly, constraint (22) is a relaxation of constraint (24) and the other constraints in the two models are the same, so $X_2 \subseteq X_1$. On the other hand, the objective functions are the same, so the cost of an optimal solution of (P1) is less than or equal to the cost of an optimal solution of (P2), hence, (P1) is a relaxation of (P2). Let $x^* \in X_1$ be an optimal solution of (P1). If x^* covers some pairings more than once, define $x^{**} \in X_2$ by removing these pairings from the schedule of the supplementary pilot the cost of x^{**} is the same as x^*, because the cost of the uncovered VRs and pairings is not increased. Then $P1(x^*) = P2(x^{**})$, so the optimal objective values of (P1) and (P2) are the same. For the copilot ($j = 2$) model the argument is similar. This result justifies the use of "$\geq$" rather than "$=$" in constraint (8). □

5 Solution methodology

In this section, we describe three proposed approaches. SEQ first solves the pairing model (2)–(3) with consider two first terms of function (1) as objective function and then the (co)pilot assignment model (8)–(10) using CG. SEQEP first solves the extended pairing model (2)–(7) with consider three first terms of function (1) as objective function and then the (co)pilot assignment model (8)–(10) using CG. In both sequential approaches, the objective function of (co)pilot assignment model (8)–(10) is two last terms of function (1) and variables y_p and r_v are set equal to 1 in this model. The integrated model (1)–(20) has a block diagonal structure (see Figure 4) and is hence suitable for mathematical decomposition. INT solves this integrated model using Benders decomposition combined with CG. We now introduce the Benders decomposition with a master problem and two subproblems that are solved by CG method.

5.1 Benders decomposition for integrated model

Benders decomposition [6] is an iterative method for large-scale optimization problems where the coefficient matrix has a block diagonal structure. Model (1)–(20) can be decomposed into multiple smaller problems. The solution process iterates between a master problem that represents the pairing problem, and two subproblems that represent the pilot and copilot assignment problems as personalized assignment problems. At each iteration, the subproblems add two Benders cuts to the master problem to reflect the information obtained from the subproblem solutions. One cut is related to pilot assignment and the other to copilot assignment. The algorithm terminates when the upper bounds provided by the Benders master problem and the lower bounds provided by the subproblems are sufficiently close. See [31] for a comprehensive review.

5.2 Benders reformulation

For given non-negative values $\bar{y}_p$ ($b \in B; p \in P_b$) and $\bar{r}_v$ ($j \in \{1,2\}, b \in B, l \in L_b^j; v \in V_l$) satisfying constraints (2)–(7), the LP relaxation of model (1)–(20) reduces to problems involving assignment variables, called the *primal pilot* ($j = 1$) and *copilot* ($j = 2$) subproblems, are as follows:

$$P_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v) = \text{minimize} \sum_{j=1}^{2} \sum_{b \in B} \sum_{l \in L_b^j} \sum_{v \in V_l} \bar{c}_v e_v + \sum_{j=1}^{2} \sum_{b \in B} \sum_{p \in P_b} \bar{c}_p e_p^j \tag{25}$$

$$\text{subject to} \quad \sum_{l \in L_b^j} \sum_{s \in S_l} e_p^s x_s + e_p^j \geq \bar{y}_p \qquad \forall j \in \{1,2\}, \forall b \in B, \forall p \in \bar{P}_b \tag{26}$$

$$\sum_{s \in S_l} v_v^s x_s + e_v \geq \bar{r}_v \qquad \forall j \in \{1,2\}, \forall b \in B, \forall l \in L_b^j, \forall v \in V_l \tag{27}$$

$$\sum_{s \in S_l} x_s \leq 1 \qquad \forall j \in \{1,2\}, \forall b \in B, \forall l \in L_b^j \tag{28}$$

$$x_s \geq 0 \qquad \forall j \in \{1,2\}, \forall b \in B, \forall l \in L_b^j, \forall s \in S_l \tag{29}$$

$$e_v \geq 0 \qquad \forall j \in \{1,2\}, \forall b \in B, \forall l \in L_b^j, \forall v \in V_l \tag{30}$$

$$e_p^j \geq 0 \qquad \forall j \in \{1,2\}, \forall b \in B, \forall p \in \bar{P}_b \tag{31}$$

Let $\Omega = (\Omega_p^j \geq 0 | j \in \{1,2\}, b \in B, p \in \bar{P}_b)$, $\theta = (\theta_v \geq 0 | j \in \{1,2\}, b \in B, l \in L_b^j, v \in V_l)$ and $\lambda = (\lambda_l \leq 0 | j \in \{1,2\}, b \in B, l \in L_b^j)$ be the dual variables associated with constraints (26)–(28), respectively. The duals of the linear relaxations of the primal (co)pilot subproblems, called the *dual pilot* $(j = 1)$ and *copilot* $(j = 2)$ subproblems, are as follows:

$$D_{\bar{y}_p, \bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l) = \text{maximize} \sum_{j=1}^{2} \sum_{b \in B} \sum_{p \in \bar{P}_b} \bar{y}_p \Omega_p^j + \sum_{j=1}^{2} \sum_{b \in B} \sum_{l \in L_b^j} \sum_{v \in V_l} \bar{r}_v \theta_v + \sum_{j=1}^{2} \sum_{b \in B} \sum_{l \in L_b^j} \lambda_l \tag{32}$$

$$\begin{aligned}
\text{subject to} \quad & \sum_{p \in \bar{P}_b} e_p^s \Omega_p^j + \sum_{v \in V_l} V_v^s \theta_v + \lambda_l \leq 0 && \forall j \in \{1,2\}, \forall b \in B, \forall l \in L_b^j, \forall s \in S_l && (33)\\
& \Omega_p^j \leq \bar{c}_p && \forall j \in \{1,2\}, \forall b \in B, \forall p \in \bar{P}_b && (34)\\
& \theta_v \leq \bar{c}_v && \forall j \in \{1,2\}, \forall b \in B, \forall l \in L_b^j, \forall v \in V_l && (35)\\
& \Omega_p^j \geq 0 && \forall j \in \{1,2\}, \forall b \in B, \forall p \in \bar{P}_b && (36)\\
& \theta_v \geq 0 && \forall j \in \{1,2\}, \forall b \in B, \forall l \in L_b^j, \forall v \in V_l && (37)\\
& \lambda_l \leq 0 && \forall j \in \{1,2\}, \forall b \in B, \forall l \in L_b^j && (38)
\end{aligned}$$

Let Δ_j $(j \in \{1,2\})$ denote the polyhedra defined by constraints (33)–(38). Let P_{Δ_j} and R_{Δ_j} be the sets of extreme points and extreme rays of Δ_j, respectively.

Proposition 2 *The dual (co)pilot subproblems are always feasible and bounded.*

Proof. Assume that $\bar{\boldsymbol{C}}_{\boldsymbol{p}} = (\bar{c}_p \geq 0 | b \in B, p \in \bar{P}_b)$ and $\bar{\boldsymbol{C}}_{\boldsymbol{v}} = (\bar{c}_v \geq 0 | b \in B, l \in L_b^j, v \in V_l)$. Then $\Delta_j \neq \emptyset$ $(j \in \{1,2\})$ because the null vector 0 satisfies constraints (33)–(38), so the dual subproblems are always feasible. On the other hand, the solution $e_p^j = \bar{y}_p$ $(\forall j \in \{1,2\}, b \in B, p \in \bar{P}_b$ $)$, $e_v = \bar{r}_v$ $(\forall j \in \{1,2\}, b \in B, l \in L_b^j, v \in V_l)$, and $x_s = 0$ $(\forall j \in \{1,2\}, b \in B, l \in L_b^j, s \in S_l)$ satisfies constraints (26)–(31). Hence, the primal (co)pilot subproblems are feasible. By strong duality, the dual (co)pilot subproblems are always bounded. □

Consequently, for given non-negative values $\bar{y}_p$ $(b \in B, p \in P_b$ $)$ and $\bar{r}_v$ $(\forall j \in \{1,2\}, b \in B, l \in L_b^j, v \in V_l)$ and for all extreme rays $(\boldsymbol{\Omega}, \boldsymbol{\theta}, \boldsymbol{\lambda}) \in R_{\Delta_j}$ we have

$$\sum_{b \in B} \sum_{p \in \bar{P}_b} \Omega_p^j \bar{y}_p + \sum_{b \in B} \sum_{l \in L_b^j} \sum_{v \in V_l} \theta_v \bar{r}_v + \sum_{b \in B} \sum_{l \in L_b^j} \lambda_l \leq 0 \qquad \forall j \in \{1,2\}.$$

Hence, the primal and dual subproblems have finite optimal solutions and the optimal value is

$$\sum_{j=1}^{2} \Big(\max_{(\boldsymbol{\Omega}, \boldsymbol{\theta}, \boldsymbol{\lambda}) \in P_{\Delta_j}} \Big(\sum_{b \in B} \sum_{p \in \bar{P}_b} \Omega_p^j \bar{y}_p + \sum_{b \in B} \sum_{l \in L_b^j} \sum_{v \in V_l} \theta_v \bar{r}_v + \sum_{b \in B} \sum_{l \in L_b^j} \lambda_l \Big)\Big).$$

The LP relaxation of model (1)–(20) can thus be reformulated as

$$\begin{aligned}
\text{Minimize} \Big\{ & \sum_{b \in B} \sum_{p \in P_b} c_p y_p + \sum_{f \in F} c_f e_f + \sum_{j=1}^{2} \sum_{b \in B} \sum_{l \in L_b^j} \sum_{v \in V_l} c_v r_v + \\
& \sum_{j=1}^{2} \Big(\max_{(\boldsymbol{\Omega}, \boldsymbol{\theta}, \boldsymbol{\lambda}) \in P_{\Delta_j}} \Big(\sum_{b \in B} \sum_{p \in P_b} \Omega_p^j y_p + \sum_{b \in B} \sum_{l \in L_b^j} \sum_{v \in V_l} \theta_v r_v + \sum_{b \in B} \sum_{l \in L_b^j} \lambda_l \Big)\Big) \Big\}
\end{aligned} \tag{39}$$

subject to (2)–(7) and (11)–(17).

By introducing the additional free variables Z^j and the function $Z^j_{\Omega,\theta,\lambda}(y_p, r_v)$, we obtain the following reformulation, called the *Benders master problem*:

$$\text{minimize} \sum_{b\in B}\sum_{p\in P_b} c_p y_p + \sum_{f\in F} c_f e_f + \sum_{j=1}^{2}\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l} c_v r_v + \sum_{j=1}^{2} Z^j \tag{40}$$

subject to (2)–(7), (11)–(17), and

$$Z^j \geq Z^j_{\Omega,\theta,\lambda}(y_p, r_v) = \sum_{b\in B}\sum_{p\in \bar{P}_b} \Omega_p^j y_p + \sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l} \theta_v r_v + \sum_{b\in B}\sum_{l\in L_b^j} \lambda_l \quad \forall j\in\{1,2\}, (\boldsymbol{\Omega},\boldsymbol{\theta},\boldsymbol{\lambda}) \in P_{\Delta_j} \tag{41}$$

where each $Z^j_{\Omega,\theta,\lambda}(y_p, r_v)$ is a linear segment of the (co)pilot penalty functions. Constraints (41) are the *Benders (co)pilot optimality cuts.* Model (40)–(41) contains one such cut for each extreme point. However, most of these cuts are inactive at the optimal solution. To avoid enumerating all the extreme points, we use an iterative approach to generate subsets of the optimality cuts (41) as needed to recognize an optimal solution. We denote the subsets of extreme points available at iteration $t = 1, 2, \ldots$ by $P^t_{\Delta_1}$ and $P^t_{\Delta_2}$, for the primal pilot and copilot subproblems, respectively. Each iteration solves the relaxed crew pairing problem as a Benders master problem by considering these subsets. The optimal solution of the relaxed Benders master problem is used as input to the primal (co)pilot subproblems. The values of the dual variables associated with constraints (26)–(28) determine an extreme point of P_{Δ_1} and P_{Δ_2}. For each extreme point, one cut is added to the relaxed Benders master problem at each iteration.

[12] describe a heuristic called *three-phase Benders decomposition and CG* for the case where the Benders subproblem is an IP. We extend this, developing a two-phase approach. In the first phase (*phase I*), the integrality requirements are relaxed and at each iteration of the Benders decomposition, the LP relaxation of the Benders master problem and the primal (co)pilot subproblems are solved by CG. All the Benders cuts generated in the first phase are retained. The second phase has two steps. In the first step (*phase II_{LP}*), the integrality constraints are introduced only for the master problem variables, and the resulting mixed-integer problem is solved by generating additional Benders cuts. All these cuts are retained. In the second step (*phase II_{IP}*), the integrality constraints are introduced for the subproblem variables, and the resulting integer problem is solved by generating integer Benders cuts (see Theorem 2). Figure 5 gives a flowchart for our approach. In practice, we often stop each phase before the optimality conditions are met. As we approach optimality, new cuts have little or no effect on the optimal Benders decomposition solution. To avoid the well-known tailing-off effect, we generate new cuts until the relative difference between the lower and upper bounds is less than or equal to 0.1% for phase I and phase II_{LP} and 0.05% for phase II_{IP}.

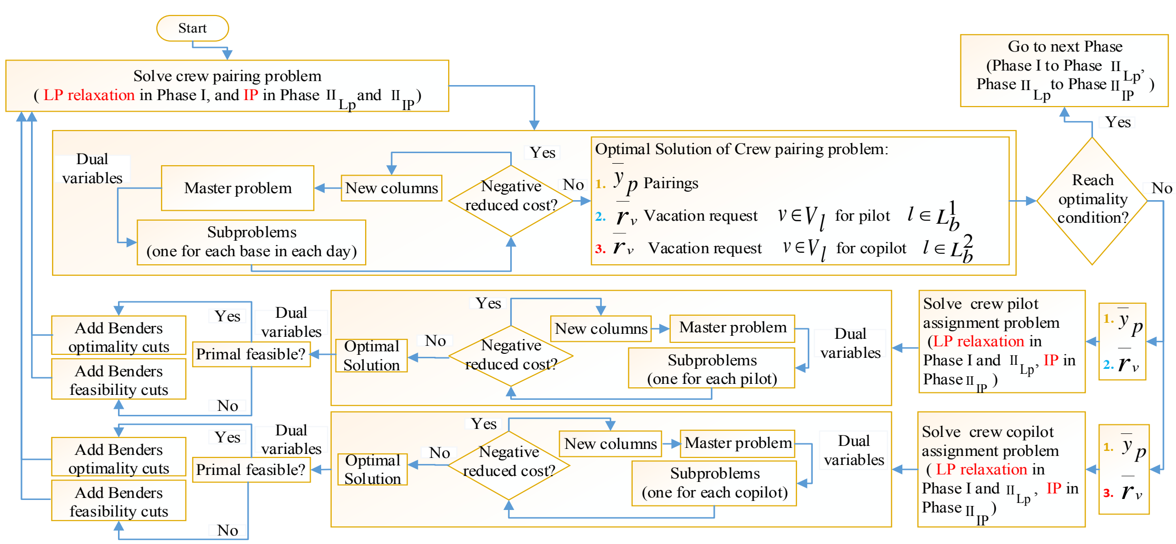

Figure 5: Flowchart of three-phase Benders decomposition.

5.3 Column generation

The crew pairing and (co)pilot assignment problems contain many variables: the number of pairing and schedule variables grows exponentially as the number of flights increases. We avoid this issue by using a CG method embedded in a branch-and-bound scheme. CG decomposes the main problem into a restricted master problem (RMP) and several subproblems. The RMP is obtained by replacing the sets y_p and x_s by the subsets $y_p^t \subseteq y_p$ and $x_s^t \subseteq x_s$ at iteration $t = 0, 1, \ldots$. At each iteration of the CG process, we solve the RMP using an LP solver over a subset of the variables (columns) to produce primal and dual solutions. Based on the dual solution, we solve the subproblems to find negative reduced cost variables. We then add these variables to the RMP. We iterate until no negative reduced cost variables are identified. In practice, as we approach optimality new pairings and schedules with negative reduced costs have little or no effect on the optimal value of the RMP. We therefore stop the CG when the optimal value of the RMP improves by less than 0.1% over five iterations. In Sections 5.3.1 and 5.3.2, we describe the subproblems for the crew pairing and (co)pilot assignment problems. To define these subproblems we use the network structures of [34] and [22], respectively. The subproblems are defined on a directed acyclic time-space network. The network corresponds to a resource-constrained shortest path problem that is solved using a label-setting algorithm [19].

5.3.1 Crew pairing subproblems

There is one subproblem for each crew base and each day, as illustrated in Figure 6. Having a separate subproblem for each day allows us to control the maximum pairing duration (four days) by limiting the subnetwork duration. There are five node types: *source, sink, departure, arrival,* and *waiting.* The source and sink nodes represent the start and end of the pairing. Each flight is defined by an arc from a departure node to an arrival node that must be covered exactly once. Finally, for each departing flight there is a waiting node, allowing that flight to start a duty. The nodes are grouped by crew base and sorted in chronological order.

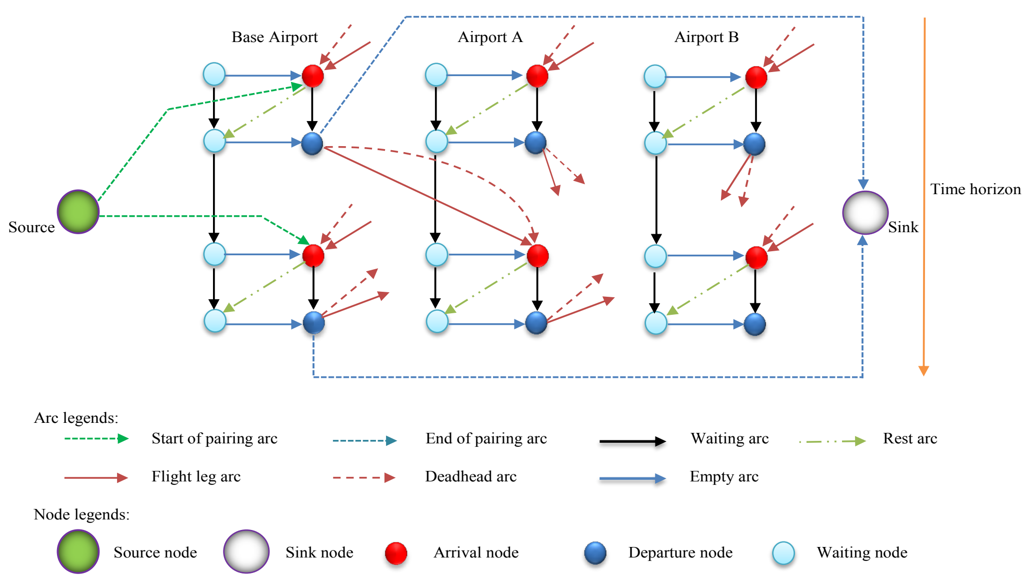

Figure 6: Network structure for crew pairing subproblems.

The network has seven arc types: *start of pairing, end of pairing, flight, deadhead, rest, waiting,* and *empty.* The start-of-pairing arcs connect the source node to each departure node at the base. The associated cost is $\varphi_i^1 + \varphi_i^2$, where φ_i^j $(j \in \{1, 2\})$ are the dual variables for constraints (5). The end-of-pairing arcs link each arrival node at the base to the sink node, and the associated cost is $-(\phi_i^1 + \phi_i^2)$, where ϕ_i^j $(j \in \{1, 2\})$ are the dual variables for constraints (4). Each flight arc and deadhead arc is defined by a departure/arrival node and fixed departure/arrival dates and times. The cost of a flight arc f is $-\beta_f - \rho_f \delta_b$, where ρ_f is the duration of flight f and β_f and δ_b are the dual variables for constraints (2) and (3). A deadhead arc has a fixed cost for each occurrence of deadheading in a pairing and a variable cost that depends on the length of the deadhead. A short-rest arc links the arrival node of a flight to the departure nodes of all the flights at the same base if the time interval is between the minimum and the ideal maximum rest time; its cost is equal to the fixed rest cost. If the waiting time exceeds the ideal maximum rest time, a long-rest arc connects the arrival node of the flight to the earliest waiting node. Waiting arcs either link two consecutive waiting nodes to extend the long rest duration or connect two consecutive flights in a duty. For the costs of long-rest and waiting arcs, we use the formulation of [30]. Finally, an empty arc links each waiting node to its corresponding departure node; its cost is zero.

Remark 1 *In airline crew scheduling there are usually enough crew members to cover all the tasks. The first iteration of Benders decomposition sequentially solves the crew pairing and personalized assignment problems. If there are uncovered pairings, reserve crew members are added. This ensures that the primal (co)pilot subproblems can cover all the pairings. The penalty cost for uncovered pairings is much greater than the cost associated with VRs ($\bar{c}_p \gg \bar{c}_v$). Consequently, the Benders solution covers all the pairings.*

Remark 2 *The Benders master problem (40)–(41) is difficult to solve with CG. Let K_1 and K_2 be the sets of Benders (co)pilot optimality cuts. Let γ_k be the dual variable associated with Benders optimality cut $k \in K_j$ ($j \in \{1,2\}$). The reduced cost of variable y_p is given by*

$$\bar{C}_p = C_p - (\sum_{f \in F} e_f^p \beta_f + \sum_{f \in F_p} \rho_f \delta_b + \sum_{j=1}^{2} \sum_{i \in A_b} \alpha_i^p \phi_i^j - \sum_{j=1}^{2} \sum_{i \in D_b} \eta_i^p \varphi_i^j - \sum_{j=1}^{2} \sum_{k \in K_j} \Omega_p^j \gamma_k)$$

To reflect the information from the restricted MMP (RMMP) in the MSP, the arc costs in MSP must be modified by the optimal dual solution for the current RMMP. The cost of a flight arc is $-\beta_f - \rho_f \delta_b$. The costs of the input and output arcs of each network in the pairing subproblems are set to $-(\phi_i^1 + \phi_i^2)$ and $\varphi_i^1 + \varphi_i^2$, respectively. However, there are no arcs corresponding to pairings in the pairing-generation subproblems (see Figure 6), so we cannot transfer the information on optimality cuts from the RMMP to the MSP.

To overcome this difficulty, we introduce the concept of weak Benders cuts in Proposition 3 that do not take into account the dual variable corresponding to the pairings. Theorem 2 proves that with these weak Benders cuts and the suggested dual solution of the primal (co)pilot subproblem, the Benders approach reach optimality in phase II_{IP}.

Proposition 3 *The following weak Benders cuts*

$$Z^j \geq \sum_{b \in B} \sum_{l \in L_b^j} \sum_{v \in V_l} \theta_v r_v + \sum_{b \in B} \sum_{l \in L_b^j} \lambda_l \qquad \forall j \in \{1,2\}, (\boldsymbol{\Omega}, \boldsymbol{\theta}, \boldsymbol{\lambda}) \in P_{\Delta_j} \tag{42}$$

are lower bounds on the (co)pilot satisfaction penalty function.

Proof. Let $\Omega = (\Omega_p^j \geq 0 | j \in \{1,2\}, b \in B, p \in \bar{P}_b)$ be the dual variables for constraints (26). Since Ω_p^j is positive for all $p \in \bar{P}_b$ we have

$$Z^j \geq Z^j_{\Omega,\theta,\lambda}(y_p, r_v) = \sum_{b \in B} \sum_{p \in \bar{P}_b} \Omega_p^j y_p + \sum_{b \in B} \sum_{l \in L_b^j} \sum_{v \in V_l} \theta_v r_v + \sum_{b \in B} \sum_{l \in L_b^j} \lambda_l \geq$$
$$\sum_{b \in B} \sum_{l \in L_b^j} \sum_{v \in V_l} \theta_v r_v + \sum_{b \in B} \sum_{l \in L_b^j} \lambda_l \qquad \forall j \in \{1,2\}, (\boldsymbol{\Omega}, \boldsymbol{\theta}, \boldsymbol{\lambda}) \in P_{\Delta_j}$$

and so

$$Z^j \geq Z^j_{\Omega,\theta,\lambda}(y_p, r_v) \geq \sum_{b \in B} \sum_{l \in L_b^j} \sum_{v \in V_l} \theta_v r_v + \sum_{b \in B} \sum_{l \in L_b^j} \lambda_l \qquad \forall j \in \{1,2\}, (\boldsymbol{\Omega}, \boldsymbol{\theta}, \boldsymbol{\lambda}) \in P_{\Delta_j}$$

This completes the proof. □

Lemma 1 *The integer optimal solution of the primal (co)pilot subproblem is an upper bound for the optimal LP relaxation of the dual (co)pilot subproblem.*

Proof. Let $P_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{LP}$ and $P_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{IP}$ be the objective functions of the LP relaxation and the IP primal (co)pilot subproblem, respectively, and let $D_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP}$ be the objective function of the LP relaxation dual (co)pilot subproblem. We denote the optimal objective values of these problems by $P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{LP}$, $P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{IP}$, and $D^*_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP}$. By strong duality, we have $D^*_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP} = P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{LP}$. Let X^P_{LP} and X^P_{IP} be the feasible domains of the LP relaxation and the IP primal (co)pilot subproblem, respectively. Clearly $X^P_{IP} \subseteq X^P_{LP}$, so $D^*_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP} = P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{LP} \le P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{IP}$. This completes the proof. □

Theorem 1 *Some optimal solutions of the linear relaxation primal (co)pilot subproblem are integer.*

Proof. It suffices to show that there exists an optimal integer primal solution such that $P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{IP} = D^*_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP}$ and therefore $P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{IP} = P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{LP}$.

Let $\bar{e}_p^j$ and $\bar{e}_v$ be the values of variables e_p^j and e_v after we solve the primal (co)pilot subproblem (25)–(31). Let $\bar{c}_p \ge 0$ and $\bar{c}_v \ge 0$ be the penalty costs for uncovered pairings and VRs. Consider the following dual solution: if pairing $p \in P_b$ is covered, i.e., $\bar{e}_p^j = 0$, then $\Omega_p^j = 0$, otherwise $\Omega_p^j = \bar{c}_p$. If vacation $v \in V_l$ is covered, i.e., $\bar{e}_v = 0$, then $\theta_v = 0$, otherwise $\theta_v = \bar{c}_v$, and $\lambda_l = -max_{s\in S_l}\{\sum_{p\in\bar{P}_b} e_p^s\Omega_p^j + \sum_{v\in V_l} V_v^s\theta_v\}$. Clearly, the first component of this solution satisfies constraints (34) and (36), the second satisfies constraints (35) and (37), and the third satisfies constraints (33) and (38). Hence, this is a feasible solution of the dual (co)pilot subproblem (32)–(38). The Benders solution covers all the pairings (Remark (1)), i.e., $\bar{e}_p^j = 0$, and thus $\Omega_p^j = 0$. Consequently, the IP (or LP) optimal value of the primal (co)pilot subproblem is

$$P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{IP} = \sum_{j=1}^{2}\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l} \bar{c}_v\bar{e}_v.$$

In the IP primal optimal solution, there are two possibilities for the coverage of each VR: (i) $v \in V_l$ is covered completely, i.e., $\bar{e}_v = 0$, and therefore $\theta_v = 0$. Hence, $V_v^s\theta_v = 0$ and $\bar{r}_v\theta_v = \bar{c}_v\bar{e}_v$. (ii) $v \in V_l$ is not covered, i.e., $\bar{e}_v = \bar{r}_v$, and therefore $\theta_v = \bar{c}_v$ and the value of parameter V_v^s must be 0. Hence, $V_v^s\theta_v = 0$ and $\bar{r}_v\theta_v = \bar{c}_v\bar{e}_v$.

The proposed dual solution gives $\Omega_p^j = 0$, and from conditions (i) and (ii) we have $V_v^s\theta_v = 0$, so $\lambda_l = -max_{s\in S_l}\{\sum_{p\in\bar{P}_b} e_p^s\Omega_p^j + \sum_{v\in V_l} V_v^s\theta_v\} = 0$. If we put this integer dual solution constructed from the IP optimal solution of the primal (co)pilot subproblem into the objective function of the dual (co)pilot subproblem (32)–(38) we obtain

$$\begin{aligned} D_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP} &= \sum_{j=1}^{2}\sum_{b\in B}\sum_{p\in\bar{P}_b} \bar{y}_p\Omega_p^j + \sum_{j=1}^{2}\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l} \bar{r}_v\theta_v + \sum_{j=1}^{2}\sum_{b\in B}\sum_{l\in L_b^j} \lambda_l = \\ &\sum_{j=1}^{2}\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l} \bar{r}_v\theta_v = \sum_{j=1}^{2}\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l} \bar{c}_v\bar{e}_v = P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{IP}. \end{aligned}$$

Hence, $D_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP} = P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{IP}$. On the other hand, since the dual (co)pilot subproblem is a maximization, clearly $D_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP} \le D^*_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP}$. Also, by lemma 1, we have $D^*_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP} \le P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{IP}$, so $D^*_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP} \le D_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP} = P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{IP} \le D^*_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP}$. Thus, $D^*_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP} = P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{IP}$ and this integer dual solution is an optimal dual solution. By strong duality, we have $D^*_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP} = P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{LP}$, so $P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{LP} = P^*_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v)_{IP}$. This means that this optimal integer solution is an optimal solution for the LP relaxation of the primal (co)pilot subproblems. This completes the proof. □

From this theorem, we can easily deduce the following corollary.

Corollary 1 *Let $D^*_{\bar{y}_p,\bar{r}_v}(\Omega^j_p, \theta_v, \lambda_l)_{IP}$ be the optimal objective value of the dual of the IP primal (co)pilot subproblem. From Theorem 1, we have $D^*_{\bar{y}_p,\bar{r}_v}(\Omega^j_p, \theta_v, \lambda_l)_{IP} = P^*_{\bar{y}_p,\bar{r}_v}(e^j_p, e_v)_{IP}$. Thus, the integer-programming duality gap in the Benders subproblems is zero, so the integer Benders cuts derived from the proposed dual solutions in Theorem 1 are valid cuts.*

Theorem 2 *The Benders master problem*

$$minimize \sum_{b\in B}\sum_{p\in P_b} c_p y_p + \sum_{f\in F} c_f e_f + \sum_{j=1}^{2}\sum_{b\in B}\sum_{l\in L^j_b}\sum_{v\in V_l} c_v r_v + \sum_{j=1}^{2} Z^j \tag{43}$$

subject to (2)–(7), (11)–(17) with Benders cuts (42) can reach an optimal solution in phase II_{IP}.

Proof. It suffices to show that when all the pairings are covered the optimal dual variables of all the pairings can be 0. We use the optimal dual solutions from Theorem 1. Figure 7 shows that the $Z^j_{\Omega,\theta,\lambda}(y_p, r_v)$ functions can be seen as supporting hyperplanes of the (co)pilot penalty function. Points A and C correspond to extreme points of MMP where $\Omega^j_p > 0$, and point B is an extreme point of MMP where $\Omega^j_p = 0$. At each iteration of the

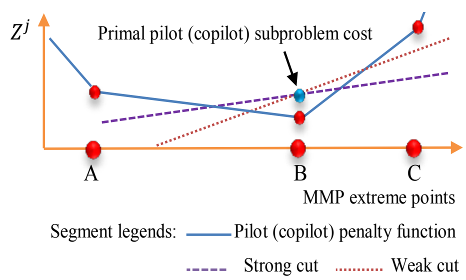

Figure 7: Graphical representation of penalty function in Benders master problem.

Benders process, the strong Benders cut (41) adds a segment $Z^j_{\Omega,\theta,\lambda}(y_p, r_v)$ to the (co)pilot penalty function, with the exact cost of the primal (co)pilot subproblem at the point B = ($\bar{y}_p$, $\bar{r}_v$) (the current solution). The weak Benders cut (42) has the same value at this point, so it adds a segment to the (co)pilot penalty function, with the exact cost at this point. The (co)pilot penalty function becomes exact at the extreme points of the MMP where all the generated pairings can be covered by the primal (co)pilot subproblem. Furthermore, since the number of these extreme points is finite, finite termination of the algorithm at an optimal solution is guaranteed. □

5.3.2 Personalized assignment subproblems

In the personalized assignment problem, there is one subproblem for each (co)pilot. There are five node types: *source, sink, pairing start, pairing end,* and *midnight.* The source and sink nodes represent the start and end of the schedules. Each pairing is defined by a pairing-start node and a pairing-end node. The start and end of each midnight are defined by midnight nodes. There are nine arc types: *start of schedule, end of schedule, start of pairing, pairing, short VR, long VR, rest, start of day-off,* and *day-off.* The start-of-schedule arc connects the source node to the first midnight node of the horizon. Its cost is $-\lambda_l$, where λ_l is the dual variable for constraint (28). The end-of-schedule arc connects the last midnight node to the sink node; its cost is 0. The start-of-pairing arcs connect each midnight node to a start-of-pairing node; the cost is 0. A pairing arc connects a start-of-pairing node to an end-of-pairing node. The cost of these arcs is $-\Omega_p$, where Ω_p is the dual variable for constraints (26). A rest arc links an end-of-pairing node at the base to the start node of a pairing whose departure time is greater than or equal to the arrival time plus the minimum rest time (8 hours in our tests). A start-of-day-off arc links an end-of-pairing node to the first midnight node at the base to start a day off. A day-off arc connects a pair of consecutive midnight nodes at the base. The cost of all these arcs is 0. A short VR arc links two midnight nodes covering between one and three days, and a long VR arc links two midnight nodes covering between four and ten days; its cost is $-\theta_v$, where θ_v are the dual variables for constraints (27). The reduced cost C_s of a variable x_s is $\bar{C}_s = C_s - (\sum_{p \in \bar{P}} e_p^s \Omega_p + \sum_{v \in V_l} V_v^s \theta_v + \lambda_l)$.

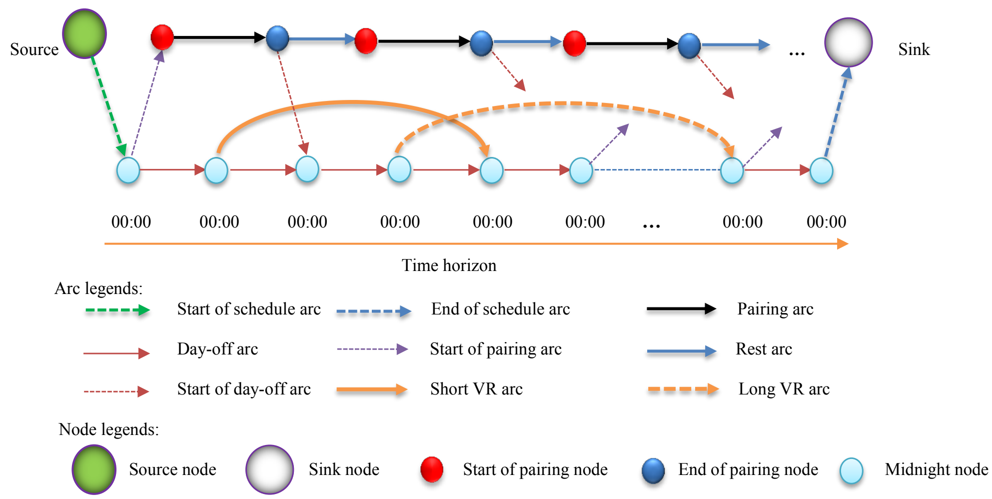

Figure 8: Network structure for personalized crew assignment subproblems.

5.4 Integer solution

The linear relaxation solutions of the crew pairing and (co)pilot assignment problems are fractional, so we embed the CG algorithm in a branch-and-bound framework. However, it is impossible to explore the whole search tree. Hence, we use two branching heuristics, column fixing and inter-task fixing, to derive integer solutions in reasonable computational times. The column fixing strategy branches directly on the variables y_p and x_s. We fix to 1 all the variables with a fractional value greater than a predetermined threshold (0.85 for our tests). If no suitable variable exists, we apply inter-task fixing. Let P be the set of all feasible pairings and S the set of all feasible schedules. For each ordered pair of flights f_1 and f_2 and pairings p_1 and p_2, we define two subsets as follows: $P(f_1, f_2) = \{p \in P$: flight f_2 is covered immediately after flight $f_1\}$ and $S(p_1, p_2) = \{s \in S$: pairing p_2 is covered immediately after pairing $p_1\}$. Let $r_{f_1 f_2} = \sum_{p \in P(f_1, f_2)} y_p$ and $r_{p_1 p_2} = \sum_{s \in S(p_1, p_2)} x_s$ be the total flow between these two flights and pairings. We select the maximum fractional values, $r_{f_1 f_2}$ and $r_{p_1 p_2}$. We then require these two flights to be covered consecutively in the same pairing (by setting $r_{f_1 f_2} = 1$) and these two pairings to be covered consecutively in the same schedule (by setting $r_{p_1 p_2} = 1$). These branching strategies produce good integer solutions with small optimality gaps (see Table 5).

5.5 Reducing computational time

We use two techniques to reduce the computational time.

5.5.1 Personalized Crew Assignment Problem

We solve the pilot and copilot assignment problems in parallel. This is possible since these problems are independent of each other.

5.5.2 Benders master problem

We warm start the Benders master problem after adding Benders cuts and after branching. We start with the feasible columns generated to date and with the current base. After solving the pairing problem at each iteration of the Benders decomposition we use the current optimal solution as the initial solution for the next iteration. The results show that this idea works well.

6 Computational experiments

This section presents our experiments with the three proposed approaches: SEQ, SEQEP and INT. We applied the algorithms to three datasets derived from a one-month flight schedule of a major North American airline. The characteristics of each instance are given in Table 1. In our test, each (co)pilot requests two vacations in each month. These are divided into two categories based on their length: a long vacation is between 4 and 10 days and a short vacation is between 1 and 3 days. The last two columns of Table 1 indicate the total number of vacation days for pilots and copilots. We conducted our tests on a Linux computer with an Intel Core i7-1770 CPU clocked at 3.40 GHz, using a single processor. Our implementation is coded in C++ using the commercial GENCOL column generation library, version 4.5. The RMPs are solved by CPLEX 12.4.

Table 1: Instance characteristics.

		#				VRs					
						Number of vacations				Vacation days	
Instance	Flights	Airports	Bases	Pilots	Copilots	Pilots		Copilots		Pilots	Copilots
						Short	Long	Short	Long		
1	1013	26	3	33	33	47	19	38	28	193	234
2	1500	35	3	34	34	43	25	43	25	200	173
3	1854	41	3	47	47	52	42	47	47	315	334

6.1 Analysis of computational refinements

This section discusses the effect of the aggregation and the impact of the warm start strategy, as described in Sections 3 and 5.5.2. Table 2 gives the total number of constraints in the extended pairing problem for each instance before and after the aggregation. The aggregation reduces the number of constraints by an average of 37.36%; the reduction increases with the size of the instance.

Table 2: Number of constraints for extended crew pairing model.

Instance	Before aggregation	After aggregation	Size reduction (%)
1	3104	2098	32.50
2	4565	2987	34.56
3	5627	3093	45.03
Average			37.36

We solved each aggregated instance with both the basic algorithm and the *refined* version that includes warm starting; see Table 3. Here, the number of columns generated and the CPU time indicate the effort needed to solve the Benders master problem using INT. The last two columns of Table 3 give the ratios for the CPU times and the number of columns generated for the basic and refined algorithms. The computational times and numbers of generated columns are reduced considerably for all the instances. The refined algorithm is an average of 1.93 times faster than the basic algorithm, and the number of columns generated increases by a factor of 1.73.

Table 3: Impact of warm start strategy on extended crew pairing problem.

	Basic algorithm		Refined algorithm			
Instance	CPU (min)	No. Columns generated	CPU (min)	No. Columns generated	CPU Basic/Refined	No. Columns Basic/Refined
1	24.30	66757	16.46	30527	1.47	2.18
2	108.35	136797	48.68	88755	2.22	1.54
3	119.21	184134	56.53	123479	2.11	1.49
Average					1.93	1.73

6.2 Results for integrated approach

In this section, for each phase of INT we indicate the time spent, the number of cuts generated, and the total cost at the end; see Table 4. The results show that the percentage gaps between the phases are small. The gap between phases II_{LP} and II_{IP}, which is between 0.00% and 0.03%, justifies the use of the three-phase approach. The performance of INT is related to the number of cuts: in each instance eight to fourteen cuts suffice to find a good solution. For the first two instances, most of the CPU time is spent on phase II_{LP}. In this phase the CG is embedded in a branch-and-bound framework to obtain an integer solution for crew pairing variables, and therefore more time is needed. For the largest instance, phase I is the most time-consuming, and this can be explained as follows. This instance has between 1300 and 1400 fractional pairings in the solution of the relaxation and between 300 and 350 pairings when the integrality constraints are added in phase II_{LP}. Therefore, size of (co)pilot assignment problems in phase I is much larger than phase II_{LP} and requires more time. For all three instances, phase II_{IP} has the lowest CPU time, because only a few cuts are generated in this phase. Adding Benders optimality cuts in the first two phases helps us to find an optimal solution more quickly in phase II_{IP}.

Table 4: Computational results for integrated approach in each phase.

	CPU time (min)			Number of cuts			Cost			Gap (%)		
Instance	I	II_{LP}	II_{IP}	I	II_{LP}	II_{IP}	I	II_{LP}	II_{IP}	I & II_{LP}	I & II_{IP}	II_{LP} & II_{IP}
1	6.66	12.76	1.50	4	4	0	163495	163847	163847	0.21	0.21	0.00
2	24.88	41.66	5.50	6	6	2	239023	239537	239628	0.21	0.25	0.03
3	77.76	58.22	19.31	4	6	2	290160	290880	290977	0.24	0.28	0.03

6.3 Computational gap in integrated and sequential approaches

In this experiment, we study the optimality gap of each problem for each approach. We report in Table 5 three integrality gaps: the crew pairing gap, the pilot assignment gap, and the copilot assignment gap. The gap is the percentage difference between the LP and integer solutions of the problem. The relatively small gaps indicate that the three approaches produce good solutions. The larger integrality gaps for the pairing problems confirm that it is difficult to find integer solutions; this is because the branching strategies are heuristic and the pairing problem is much larger than the assignment problem. On average, the integrality gaps obtained by INT are smaller than those for SEQEP and SEQ for the assignment problem.

Table 5: Optimality gap (%).

	Crew pairing	Pilot assignment	Copilot assignment
SEQ			
Instance 1	0.07	0.10	0.00
Instance 2	1.86	0.00	0.22
Instance 3	0.88	1.55	0.21
Average	0.93	0.75	0.55
SEQEP			
Instance 1	0.03	0.09	0.00
Instance 2	2.47	0.34	0.00
Instance 3	0.80	0.81	0.00
Average	1.10	0.41	0.00
INT			
Instance 1	0.29	0.00	0.00
Instance 2	1.69	0.12	0.00
Instance 3	1.07	0.00	0.00
Average	1.01	0.04	0.00

6.4 Number of branching nodes, iterations, and columns

Table 6 presents for each approach the total number of branch-and-bound nodes explored, CG iterations performed, and columns generated during the solution process. INT has more branching nodes, iterations, and generated columns than SEQEP and SEQ. This is because INT performs an average of six iterations, and for SEQEP and SEQ is one. We conclude that the pairing problem is more difficult than the assignment problem. In the former problem there are more generated columns and branching nodes.

Table 6: Results.

	Crew pairing			Pilot assignment			Copilot assignment		
	Number of			Number of			Number of		
	Branching nodes	Generation iterations	Columns generated	Branching nodes	Generation iterations	Columns generated	Branching nodes	Generation iterations	Columns generated
SEQ									
Instance 1	30	310	18690	21	459	3157	23	398	2845
Instance 2	138	1000	37343	27	1315	8369	29	1282	8509
Instance 3	169	1171	51987	38	609	7217	39	599	7855
SEQEP									
Instance 1	11	207	14313	17	315	2527	23	465	2806
Instance 2	161	1129	35444	26	1397	8080	27	1329	8224
Instance 3	170	1179	51378	34	543	7250	36	544	7387
Phase I									
Instance 1	-	152	11830	-	158	4863	-	165	4880
Instance 2	-	169	16537	-	1878	20099	-	1817	19684
Instance 3	-	159	20744	-	185	7353	-	191	7311
Phase II$_{\mathbf{LP}}$									
Instance 1	83	516	17153	-	155	3826	-	141	3639
Instance 2	404	2360	43663	-	666	5057	-	653	4924
Instance 3	501	2800	66721	-	159	6492	-	168	6684
Phase II$_{\mathbf{IP}}$									
Instance 1	34	212	1544	20	443	3324	19	374	2683
Instance 2	274	1585	28555	54	2634	14954	24	2429	15074
Instance 3	310	1949	36014	72	1037	14311	80	1240	14824
Total- INT									
Instance 1	117	880	30527	20	756	12013	19	680	11202
Instance 2	678	4104	88755	54	5178	40110	48	4899	39682
Instance 3	1381	4908	123479	72	881	28156	80	1599	28819

6.5 Coverage of VRs

We now analyze the performance of the approaches in terms of the percentage of satisfied VRs; see Table 7. The results clearly show that the percentage of satisfied VRs decreases as the complexity of the instance increases. INT has significantly better performance than SEQEP, and SEQEP has significantly better performance than SEQ. The last three rows of Table 7 provide a pairwise comparison of the approaches. For example, on average, INT increases the coverage of long and short VRs by 15.66% and 10.66% for pilots and 15.00% and 12.00% for copilots compared to SEQEP. On average, INT increases the coverage of VRs by 16% for pilots and 18.33% for copilots in terms of the total number of days requested when compared with SEQEP.

Table 7: Satisfaction of VRs (%).

	Pilot assignment		Copilot assignment		Pilot assignment	Copilot assignment
	Long	Short	Long	Short	Days	Days
SEQ						
Instance 1	57.00	73.00	50.00	78.00	62.00	56.00
Instance 2	12.00	58.00	28.00	53.00	32.00	38.00
Instance 3	23.00	67.00	23.00	55.00	34.00	30.00
Average	30.66	66.00	33.66	62.00	42.66	41.33
SEQEP						
Instance 1	72.00	80.00	53.00	86.00	74.00	61.00
Instance 2	43.00	73.00	43.00	79.00	60.00	58.00
Instance 3	30.00	70.00	34.00	61.00	40.00	38.00
Average	48.33	74.33	43.33	75.33	58.00	52.33
INT						
Instance 1	77.00	97.00	82.00	86.00	90.00	82.00
Instance 2	50.00	81.00	53.00	89.00	65.00	71.00
Instance 3	65.00	77.00	40.00	87.00	67.00	59.00
Average	64.00	85.00	58.33	87.33	74.00	70.66
Improvements						
SEQEP vs. SEQ	17.67	8.33	9.67	13.13	15.34	11.00
INT vs. SEQEP	15.67	10.67	15.00	12.00	16.00	18.33
INT vs. SEQ	34.66	19.00	24.67	25.33	31.34	29.33

6.6 Coverage of flights and pairings

We define large penalty costs to ensure that the percentages of uncovered flights and pairings are small. INT and SEQEP cover all the flights and pairings for all the instances. With SEQ 0.32% and 0.60% of the (co)pilot pairings in the second and third instances are uncovered, respectively. This is because with SEQ the pairing problem does not take into account how many pilots and copilots are available.

6.7 Computational time and cost for integrated and sequential approaches

We now compare the total CPU time (in minutes) and the solution cost for the pairing and assignment problems; see Table 8. For SEQEP and SEQ, the total CPU time is the time to solve the pairing and assignment problems. For INT the total CPU time is the time to solve all three phases. The pilot and copilot problems were solved in parallel, as explained in Section 5.5.1.

On average, INT decreases the pairing cost by 0.83% compared with SEQEP and 0.10% compared with SEQ. These results confirm the effectiveness of the warm start strategy, especially since the INT model has more constraints. Clearly, if an exact branching strategy were used, the results might be different. On average, SEQEP increases the pairing cost by 0.72% compared with SEQ. This increase might be necessary to increase the number of satisfied VRs.

For the assignment costs, INT yields significant savings compared with SEQEP, and SEQEP yields significant savings compared with SEQ. For example, INT decreases the pilot (copilot) costs by 54.92% (53.24%), in comparison with SEQEP. The computational time increases by a factor of 2.55.

Table 8: Comparisons of CPU time and cost of solution.

	Total CPU (min)	Pairing cost	Pilot assignment cost	Copilot assignment cost
SEQ				
Instance 1	3.66	172666	2400	3300
Instance 2	14.90	247343	6000	5700
Instance 3	50.03	296608	7350	8550
Average	22.86	238872	5250	5850
SEQEP				
Instance 1	7.42	172534	2100	2550
Instance 2	32.70	248675	2850	3000
Instance 3	55.81	300617	5700	6000
Average	31.97	240608	3550	3850
INT				
Instance 1	20.92	172547	600	1200
Instance 2	69.04	246128	1950	1650
Instance 3	155.29	297177	2250	2550
Average	81.75	238617	1600	1800
Cost saving (%) & CPU ratio				
SEQEP vs. SEQ	1.39	-0.72	32.38	34.18
INT vs. SEQEP	2.55	0.83	54.92	53.24
INT vs. SEQ	3.57	0.10	69.52	69.23

7 Conclusion

We have proposed two novel mathematical models: an extended model for the crew pairing problem, and a completely integrated model for the crew pairing and personalized assignment problems. We use aggregation in the extended pairing problem to decrease the number of constraints and reduce the CPU time. We have presented two sequential solution approaches (SEQ and SEQEP), and an integrated approach (INT). The sequential approaches are based on CG. The integrated approach is based on combined Benders decomposition and CG, and it uses a warm start strategy. The theoretical results for the integrated approach show that optimality can be achieved.

We studied real-world instances from a major US carrier. The results show that the novel extended crew pairing model leads to monthly schedules that are significantly better in terms of satisfied VRs. The integrated approach yields significant improvements compared with the sequential approaches. However, the SEQEP can solve large problems in reasonable computational times.

Future work could consider other crew preferences, such as preferences for specific flights.

References

[1] R. Anbil, E. Gelman, B. Patty, and R. Tanga. Recent advances in crew-pairing optimization at American Airlines. Interfaces, 21(1):62–74, 1991.

[2] R. Anbil, R. Tanga, and E. L. Johnson. A global approach to crew-pairing optimization. IBM Systems Journal, 31(1):71–78, 1992.

[3] C. Barnhart and R. G. Shenoi. An approximate model and solution approach for the long-haul crew pairing problem. Transportation Science, 32(3):221–231, 1998.

[4] C. Barnhart, A. M. Cohn, E. L. Johnson, D. Klabjan, G. L. Nemhauser, and P. H. Vance. Airline crew scheduling. In Handbook of Transportation Science, pages 517–560. Springer, 2003.

[5] J. E. Beasley and B. Cao. A tree search algorithm for the crew scheduling problem. European Journal of Operational Research, 94(3):517–526, 1996.

[6] J. F. Benders. Partitioning procedures for solving mixed-variables programming problems. Numerische Mathematik, 4(1):238–252, 1962.

[7] R. E. Bixby, J. W. Gregory, I. J. Lustig, R. E. Marsten, and D. F. Shanno. Very large-scale linear programming: A case study in combining interior point and simplex methods. Operations Research, 40(5):885–897, 1992.

[8] K. Boubaker, G. Desaulniers, and I. Elhallaoui. Bidline scheduling with equity by heuristic dynamic constraint aggregation. Transportation Research Part B: Methodological, 44(1):50–61, 2010.

[9] K. W. Campbell, R. B. Durfee, and G. S. Hines. Fedex generates bid lines using simulated annealing. Interfaces, 27(2):1–16, 1997.

[10] C.-H. Chen, T.-K. Liu, J.-H. Chou, and C.-C. Wang. Multiobjective airline scheduling: An integrated approach. In SICE Annual Conference (SICE), 2012 Proceedings of, pages 1266–1270. IEEE, 2012.

[11] I. T. Christou, A. Zakarian, J.-M. Liu, and H. Carter. A two-phase genetic algorithm for large-scale bidline-generation problems at Delta Air Lines. Interfaces, 29(5):51–65, 1999.

[12] J.-F. Cordeau, G. Stojković, F. Soumis, and J. Desrosiers. Benders decomposition for simultaneous aircraft routing and crew scheduling. Transportation Science, 35(4):375–388, 2001.

[13] G. Desaulniers, J. Desrosiers, Y. Dumas, S. Marc, B. Rioux, M. M. Solomon, and F. Soumis. Crew pairing at Air France. European Journal of Operational Research, 97(2):245–259, 1997.

[14] M. Gamache, F. Soumis, G. Marquis, and J. Desrosiers. A column generation approach for large-scale aircrew rostering problems. Operations Research, 47(2):247–263, 1999.

[15] C. Gao, E. Johnson, and B. Smith. Integrated airline fleet and crew robust planning. Transportation Science, 43(1):2–16, 2009.

[16] I. Gershkoff. Optimizing flight crew schedules. Interfaces, 19(4):29–43, 1989.

[17] B. Gopalakrishnan and E. L. Johnson. Airline crew scheduling: State-of-the-art. Annals of Operations Research, 140(1):305–337, 2005.

[18] K. L. Hoffman and M. Padberg. Solving airline crew scheduling problems by branch-and-cut. Management Science, 39(6):657–682, 1993.

[19] S. Irnich and G. Desaulniers. Shortest path problems with resource constraints. In Column Generation, pages 33–65. Springer, 2005.

[20] A. I. Jarrah and J. T. Diamond. The problem of generating crew bidlines. Interfaces, 27(4):49–64, 1997.

[21] A. Kasirzadeh. Optimisation intégrée des rotations et des blocs mensuels personnalisés des équipages en transport aérien. Ph.D. thesis, Department of Mathematics and Industrial Engineering, Polytechnique Montréal, 2015.

[22] A. Kasirzadeh, M. Saddoune, and F. Soumis. Airline crew scheduling: Models, algorithms, and data sets. EURO Journal on Transportation and Logistics, pages 1–27, 2014.

[23] D. Klabjan, E. L. Johnson, G. L. Nemhauser, E. Gelman, and S. Ramaswamy. Solving large airline crew scheduling problems: Random pairing generation and strong branching. Computational Optimization and Applications, 20(1):73–91, 2001.

[24] D. Klabjan, E. L. Johnson, G. L. Nemhauser, E. Gelman, and S. Ramaswamy. Airline crew scheduling with time windows and plane-count constraints. Transportation Science, 36(3):337–348, 2002.

[25] B. Maenhout and M. Vanhoucke. A hybrid scatter search heuristic for personalized crew rostering in the airline industry. European Journal of Operational Research, 206(1):155–167, 2010.

[26] R. E. Marsten and F. Shepardson. Exact solution of crew scheduling problems using the set partitioning model: Recent successful applications. Networks, 11(2):165–177, 1981.

[27] A. Mercier and F. Soumis. An integrated aircraft routing, crew scheduling and flight retiming model. Computers & Operations Research, 34(8):2251–2265, 2007.

[28] A. Mercier, J.-F. Cordeau, and F. Soumis. A computational study of Benders decomposition for the integrated aircraft routing and crew scheduling problem. Computers & Operations Research, 32(6):1451–1476, 2005.

[29] N. Papadakos. Integrated airline scheduling. Computers & Operations Research, 36(1):176–195, 2009.

[30] F. Quesnel, G. Desaulniers, and F. Soumis. A new heuristic branching scheme for the crew pairing problem with base constraints. Computers & Operations Research, 80:159–172, 2016.

[31] R. Rahmaniani, T. G. Crainic, M. Gendreau, and W. Rei. The Benders decomposition algorithm: A literature review. European Journal of Operational Research, 2016.

[32] M. Saddoune, G. Desaulniers, I. Elhallaoui, and F. Soumis. Integrated airline crew scheduling: A bi-dynamic constraint aggregation method using neighborhoods. European Journal of Operational Research, 212(3):445–454, 2011.

[33] M. Saddoune, G. Desaulniers, I. Elhallaoui, and F. Soumis. Integrated airline crew pairing and crew assignment by dynamic constraint aggregation. Transportation Science, 46(1):39–55, 2012.

[34] M. Saddoune, G. Desaulniers, and F. Soumis. Aircrew pairings with possible repetitions of the same flight number. Computers & Operations Research, 40(3):805–814, 2013.

[35] R. Sandhu and D. Klabjan. Integrated airline fleeting and crew-pairing decisions. Operations Research, 55(3): 439–456, 2007.

[36] S. Shao, H. D. Sherali, and M. Haouari. A novel model and decomposition approach for the integrated airline fleet assignment, aircraft routing, and crew pairing problem. Transportation Science, 51:233–249, 2015.

[37] N. Souai and J. Teghem. Genetic algorithm based approach for the integrated airline crew-pairing and rostering problem. European Journal of Operational Research, 199(3):674–683, 2009.

[38] S. Subramanian and H. D. Sherali. An effective deflected subgradient optimization scheme for implementing column generation for large-scale airline crew scheduling problems. INFORMS Journal on Computing, 20(4): 565–578, 2008.

[39] P. H. Vance, C. Barnhart, E. L. Johnson, and G. L. Nemhauser. Airline crew scheduling: A new formulation and decomposition algorithm. Operations Research, 45(2):188–200, 1997.

[40] J. D. Weir and E. L. Johnson. A three-phase approach to solving the bidline problem. Annals of Operations Research, 127(1-4):283–308, 2004.

[41] F. Zeghal and M. Minoux. Modeling and solving a crew assignment problem in air transportation. European Journal of Operational Research, 175(1):187–209, 2006.