UNIVERSITÉ DE MONTRÉAL

OPTIMISATION INTÉGRÉE DES ROTATIONS ET DES BLOCS MENSUELS PERSONNALISÉS DES PILOTES ET DES COPILOTES SIMULTANÉMENT

VAHID ZEIGHAMI
DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIÆ DOCTOR
(MATHÉMATIQUES DE L'INGÉNIEUR)
JANVIER 2019

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée :

OPTIMISATION INTÉGRÉE DES ROTATIONS ET DES BLOCS MENSUELS PERSONNALISÉS DES PILOTES ET DES COPILOTES SIMULTANÉMENT

présentée par : ZEIGHAMI Vahid
en vue de l'obtention du diplôme de : Philosophiæ Doctor
a été dûment acceptée par le jury d'examen constitué de :

M. ROUSSEAU Louis-Martin, Ph. D., président.
M. SOUMIS François, Ph. D., membre et directeur de recherche.
M. GENDRON Bernard, Ph. D., membre.
M. FRANGIONI Antonio, Ph. D., membre externe.

DÉDICACE

À ma femme, Sheida,
À mes parents et ma famille,
de tout mon cœur…
A tous ceux qui me sont chers,

REMERCIEMENTS

je remercie sincèrement mon directeur de recherche, M. François Soumis, pour ses excellents conseils, son immense savoir, son soutien continu et son engagement qui a eu un impact considérable sur mon travail.

Je remercie M. François Lessard pour ses conseils tout au long de ce long processus.

Je voudrais remercier M. Antonio Frangioni, M. Bernard Gendron et M. Louis-Martin Rousseau pour avoir accepté de siéger à mon comité de thèse. J'apprécie vraiment leur attention et des évaluations.

Je voudrais remercier le personnel de AD-OPT pour fournir les ensembles de données et la bibliothèque de logiciels GENCOL.

je remercie fortement tous mes amis du centre de recherche GERAD et CIRRELT pour leur aide et leur soutien sans fin. De plus, je voudrais remercier le personnel du GERAD. Finalement, je tiens à remercier mes parents et ma femme pour leur soutien interminable.

RÉSUMÉ

Le problème de construction des blocs mensuels pour les membres d'équipage consiste à déterminer des horaires mensuels pour les membres d'équipage des compagnies aériennes tels que tous les vols planifiés sur un horizon de planification donné (généralement un mois) sont couverts tout en satisfaisant un certain nombre de contraintes. En raison de sa taille et de sa complexité, ce problème est généralement résolu séquentiellement en deux étapes : la construction des rotations suivie par la construction des blocs mensuels. Une rotation est une séquence de vols, de connexions et de pauses effectuée par un équipage partant et revenant à la même base. Le problème de construction des rotations consiste à déterminer un ensemble de rotations réalisables à un coût minimal, de telle sorte que chaque vol soit couvert exactement une seule fois. Dans le problème d'affectation des membres d'équipage, l'objectif est de construire des horaires mensuels à partir de ces rotations pour un ensemble donné de pilotes et de copilotes. La construction des rotations et des blocs mensuels doit respecter les règles de la sécurité aérienne, les règles d'opération de la compagnie et les règles contenues dans les conventions collectives entre les employés et la compagnie aérienne. Cependant, il peut s'avérer impossible que l'approche séquentielle obtienne une solution globale optimale car le domaine de décision du problème d'affectation des membres d'équipage est réduit par les décisions précédemment prises dans le problème de construction des rotations des membres d'équipage. L'objectif principal de cette thèse est de proposer des modèles intégrés et de nouvelles approches qui permettent de résoudre le problème de planification des membres d'équipage pour un ensemble donné de pilotes et de copilotes simultanément. Tous les tests réalisés dans cette thèse se basent sur des instances réelles fournies par une compagnie aérienne américaine. À part l'introduction, la revue de littérature et la conclusion, cette thèse comprend trois chapitres principaux dont chacun présente les travaux réalisés pour un objectif de recherche bien précis.

Dans le premier objectif, nous proposons en premier temps une extension du problème de construction des rotations des membres d'équipage qui intègre les demandes de vacances du pilote et du copilote au stade du couplage des équipages. Deuxièmement, nous présentons un modèle qui intègre complètement les problèmes de construction des rotations et le problème d'affectation des équipes simultanément pour les pilotes et les copilotes. Pour résoudre ce modèle intégré, nous développons une méthode qui combine la décomposition de Benders et la génération de colonnes.

Dans un cas plus général concernant le problème de planification des équipages de ligne

aérienne, chaque pilote/copilote a la possibilité de choisir chaque mois un ensemble de vols préférés parmi les vols réguliers. Le deuxième objectif de la thèse consiste à étudier la difficulté d'utiliser la méthode proposée dans le premier objectif lorsque nous considérons un ensemble de vols préférés et de demandes de vacances pour chaque pilote et copilote.

Quant au troisième objectif de la thèse, nous considérons le problème de planification d'équipage (pilotes et copilotes) dans un contexte personnalisé où chaque pilote/copilote demande un ensemble de préférences pour des vols spécifiques et des vacances par mois. En effet, nous proposons un modèle intégré qui permet de générer des blocs mensuels personnalisés pour les pilotes et les copilotes simultanément en une seule étape où nous gardons les rotations dans les deux problèmes aussi similaires que possible afin de réduire la propagation des perturbations pendant l'opération. Pour résoudre ce modèle intégré, nous développons une méthode qui combine la relaxation lagrangienne, la génération de colonnes et l'agrégation dynamique des contraintes. Le processus de résolution itère entre le modèle intégré des pilotes et le modèle intégré des copilotes en estimant les effets des décisions du premier problème sur le second problème.

ABSTRACT

The airline crew scheduling problem consists of determining crew schedules for airline crew members such that all the scheduled flights over a planning horizon (usually a month) are covered and the constraints are satisfied. Due to its complexity, this problem is usually solved in two phases: the crew pairing followed by the crew assignment. 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 airline crew members, while respecting all the safety and collective agreement rules. However, finding an optimal global solution via sequential approach may become impossible because the decision domain of the crew assignment problem is reduced by previously made decisions in the crew pairing problem. The main goal of this dissertation is to propose integrated models and approaches to solving the crew scheduling problem for a given set of pilots and copilots simultaneously. We conduct computational experiments on a set of real instances from a major US carrier.

In the first essay of this dissertation, 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 model, we develop a method that combines Benders decomposition and column generation.

In a more general case in the airline crew scheduling problem, each pilot and copilot have the option of choosing a set of preferred flights from the scheduled flights per month. In chapter 5, we study the difficulty of using the proposed method in the first essay when we consider a set of preferred flights and vacation requests for each pilot and copilot.

In the third essay of this dissertation, we consider the pilot and copilot crew scheduling problems in a personalized context where each pilot and copilot requests a set of preferences flights and vacations per month. We propose a model that completely integrates the crew pairing and personalized assignment problems to generate personalized monthly schedules for a given set of pilots and copilots simultaneously. The proposed model keeps the pairings in the two problems as similar as possible so the propagation of the perturbations during the operation is reduced. To solve this integrated model, we develop an integrated approach that combines alternating Lagrangian decomposition, column generation, and dynamic constraint aggregation.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

LISTE DES SIGLES ET ABBREVIATIONS

CG	Column Generation
DCA	Dynamic Constraint Aggregation
VRs	Vacation Requests
BD	Benders' Decomposition
SEQ	sequential approach
SEQEP	Sequential approach with the Extended crew pairing problem
INT	Integrated approach
RMP	Restricted Master Problem
PFs	Preferred Flights
ALD	Alternating Lagrangian Decomposition

CHAPITRE 1 INTRODUCTION

Le processus de planification du transport aérien se décompose généralement de quatre étapes principales : problème de planification des vols (flight scheduling problem), problème d'affectation des flottes (fleet assignment problem), problème de routage d'avion (aircraft routing problem) et problème de construction des horaires des membres d'équipage (crew scheduling problem). En pratique, les quatre étapes sont résolues séquentiellement et la sortie d'une étape est l'entrée de l'étape suivante. Le problème de planification des vols détermine un ensemble de vols avec des heures de départ et d'arrivée spécifiques, dans le but est de maximiser les bénéfices attendus. Le problème d'affectation des avions aux vols détermine le type d'avion à affecter à chaque vol régulier afin de maximiser les profits. Le problème de routage des avions détermine la séquence des segments de vol à effectuer par chaque avion de telle sorte à couvrir chaque segment de vol exactement une fois tout en assurant l'entretien des avions. Finalement, les compagnies aériennes résolvent le problème de construction des horaires des membres d'équipage en considérant les itinéraires des avions et un ensemble de règles de travail définies par la convention collective et les autorités du transport aérien.

Pour l'industrie aérienne, les coûts d'équipage sont les plus élevés après les coûts de carburant (Barnhart et al., 2003a). En raison de sa complexité, ce problème est généralement résolu en deux étapes : la construction des rotations suivie par la construction des blocs mensuels. Le problème de construction des rotations consiste à déterminer un ensemble des rotations réalisables compte tenu des vols, de telle sorte que le coût des rotations soit minimisé et chaque vol soit couvert exactement une seule fois. Une rotation (pairing) est une séquence d'un ou plusieurs services de vols séparés par des périodes de repos (arrêts de nuit). Un service de vol (duty) est une journée de travail composée de vols consécutifs séparés par des périodes de connexion. Un service de vols peut contenir des vols de repositionnement (deadheads) qui sont des vols où l'équipage est considéré comme passagers. L'objectif du problème d'affectation des membres d'équipage est de construire des horaires mensuels à partir de ces rotations pour chaque membre de l'équipage, de telle sorte que chaque rotation soit couverte exactement une fois tout en respectant les règles de sécurité, ainsi que les règles définies par la convention collective. Les membres de l'équipage sont formés pour un type d'avion spécifique et sont associés à une base. Tous les rotations assignées à un membre d'équipage doivent commencer et se terminer à la même base. Une base est un grand aéroport où les membres de l'équipage sont associés. La construction des horaires de l'équipage peut différer d'une compagnie aérienne à l'autre. Généralement, il existe trois approches différentes utilisées pour ce problème : construction des blocs anonymes (bidline), construction personnalisée avec ancienneté stricte et construction personnalisée avec un objectif global. La première approche (bidline) produit des horaires mensuels anonymes et après les membres de l'équipage choisissent leurs horaires en fonction de leur ordre d'ancienneté. Cette approche a commencé à être utilisée par certaines compagnies aériennes nord-américaines. L'approche personnalisée avec une ancienneté stricte, qui devient de plus en plus populaire en Amérique du Nord consiste à maximiser séquentiellement la satisfaction des employés par ordre décroissant d'ancienneté. Quant à la dernière approche, elle permet de construire des horaires mensuels dont l'objectif est de maximiser la somme de la satisfaction personnelle sans aucun avantage pour le plus d'ancienneté. Cette approche est privilégiée par les compagnies aériennes européennes.

Dans cette thèse, nous nous concentrons sur les problèmes d'affectation personnalisés des équipes dans un contexte personnalisé avec un objectif global, où chaque vol doit être couvert par un pilote et un copilote, et chaque pilote et copilote demandent un ensemble de vacances (VRs) et de vols préférés (PFs) par mois. L'utilisation d'une approche séquentielle pour résoudre le problème de planification des équipes réduit considérablement la complexité du processus, mais peut mener à des solutions sous-optimales car les contraintes des horaires ne sont pas prises en compte lors de la construction des rotations. En effet, la phase de construction des rotations ne peut pas déterminer les meilleures rotations pour la phase d'affectation des équipages. Par conséquent, il est difficile de maximiser la satisfaction des préférences de l'équipage. Cette thèse concerne le développement de plusieurs modèles intégrés et des approches pour résoudre le problème de la planification des horaires d'équipage pour un ensemble donné de pilotes et copilotes simultanément. Pour analyser les performances de nos développements, nous considérons quatre ensembles de données provenant d'une grande compagnie aérienne nord-américaine. Nous terminons cette section avec un aperçu des contributions et de la structure de cette thèse, qui fournit diverses contributions numériques et théoriques.

1.1 Contribution à la thèse

Les contributions les plus remarquables de cette thèse sont les suivantes :

- Première contribution,
  - Nous proposons un nouveau modèle extension pour le problème de rotations des membres d'équipage qui prend en compte les demandes de vacances des pilotes et copilotes pour obtenir de meilleures rotations pour l'étape d'affectation de l'équipage;
  - Nous proposons une règle d'agrégation dynamique pour réduire le nombre de contraintes

dans le modèle extension de construction des rotations ;

- Nous proposons un nouveau modèle intégré pour le problème de rotations d'équipage et l'affectation personnalisée;
- Nous proposons une approche intégrée basée sur combinaison des approches de décomposition de Benders et de génération de colonnes;
- Nous proposons une stratégie efficace pour accélérer la convergence de l'algorithme de résolution du problème;
- Nous développons les résultats théoriques pour montrer que l'approche proposée atteint l'optimalité;
- Les résultats montrent que l'approche intégrée produit des améliorations significatives par rapport aux approches séquentielles ;

- Deuxième contribution,
  - Nous étudions la difficulté d'utiliser l'approche proposée dans le premier objectif lorsque l'on considère un ensemble de PFs et VRs pour chaque pilote et copilote;
- Troisième contribution,
  - Nous proposons un nouveau modèle intégré pour le problème de rotations d'équipage et d'affectation personnalisée pour générer des plannings personnalisés mensuels pour un ensemble donné de pilotes et copilotes simultanément;
  - Garder les rotations dans les deux problèmes aussi semblables que possible afin de réduire la propagation des perturbations pendant l'opération;
  - Nous proposons une approche intégrée basée sur l'alternance de la décomposition Lagrangienne, génération de colonnes et l'agrégation dynamique de contraintes;
  - L'approche développée utilise la décomposition lagrangienne de façon alternée proposant une nouvelle façon de mettre à jour les multiplicateurs de Lagrange;
  - La solution calculée par l'approche intégrée proposée est une solution optimale pour le modèle intégré si la similarité des rotations entre pilotes et copilotes est de 100 %;
  - Quelques itérations suffisent pour trouver une solution (presque) optimale pour le problème grand et complexe;

- Les résultats montrent que le temps de calcul requis par l'approche proposée est proche de l'approche séquentielle. En moyenne, c'est seulement 1,65 fois plus longtemps que ceux obtenus par approche séquentielle;
- Les résultats montrent aussi que l'approche intégrée apporte des améliorations significatives par rapport à l'approche séquentielle;

CHAPITRE 2 ORGANISATION DE LA THÈSE

Après l'introduction dans Chapitre 1, Chapitre 3 présente la revue de la littérature. Chapitre 4 propose deux nouveaux modèles mathématiques : un nouveau modèle extension pour le problème de rotations des membres d'équipage qui prend en considération les demandes des vacances des pilotes et des copilotes, et un deuxième modèle qui intègre complétement le problème de construction des rotations et le problème de construction des blocs mensuels pour les pilotes et les copilotes simultanément. Ce dernier a été résolu en développant une méthode basée sur la décomposition de Benders et la génération de colonnes. Nous comparons par la suite les résultats obtenus par cette méthode avec ceux obtenus par les deux méthodes séquentielles.

Le deuxième objectif de cette thèse est décrit dans chapitre 5. Sa contribution étudie la difficulté d'utiliser la méthode proposée dans le premier objectif si nous voulons considérer un ensemble de vols préférés pour chaque pilote et copilote par mois.

Chapitre 6, quant à lui, constitue le troisième objectif de cette thèse où nous considérons les pilotes et les copilotes dans un contexte personnalisé où chaque pilote et copilote demande un ensemble de vols et vacances préférés par mois. Le modèle proposé intègre complètement les problèmes des rotations d'équipage et l'affectation personnalisée des pilotes/copilotes afin de générer simultanément des horaires mensuels personnalisés en une seule étape d'optimisation. Le modèle proposé maintient les rotations dans les deux problèmes aussi semblables que possible afin de réduire la propagation des perturbations survenant au cours de l'opération. Ce chapitre se termine par la présentation des comparaisons numériques par rapport à l'approche séquentielle traditionnelle. Chapitre 7 présente une synthèse du travail. Enfin, la discussion de cette thèse et les conclusions sont données dans Chapitre 8 et Chapitre 9, respectivement.

CHAPITRE 3 REVUE DE LITTÉRATURE

Au cours des dernières décennies, divers modèles et méthodes ont été introduits pour résoudre le problème de la planification des horaires d'équipage (voir Kasirzadeh et al. (2017)). La littérature relative à ce problème peut être divisée en trois catégories. La première concerne le problème de rotations de l'équipage, la seconde concerne le problème d'affectation de l'équipage, quant à la troisième, elle concerne le modèle intégré pour construit simultanée les rotations et les blocs.

3.1 Problème de rotation d'équipage

Marsten and Shepardson (1981) ont proposé une approche pour résoudre le problème de construction des rotations basée sur la relaxation lagrangienne et la méthode du sous-gradient. Ils ont appliqué leur algorithme à un ensemble de données provenant de Flying Tiger Line, Pacific Southwest Airlines, Continental Airlines et Helsinki City Transport. Gershkoff (1989) a introduit un algorithme heuristique itératif pour résoudre le problème journalier de rotation. À chaque itération, des rotations possibles sont construites pour un sous-ensemble de vols jusqu'à ce qu'aucune amélioration supplémentaire ne soit possible ou une restriction de temps est satisfaite. Anbil et al. (1992) ont développé un algorithme heuristique. où des millions des rotations possibles sont énumérées à priori et plusieurs milliers sont utilisées pour le solveur LP. À chaque itération, la plupart des rotations non reliées à une base sont ignorées et de nouvelles rotations sont ajoutées. Le processus continue jusqu'à ce que toutes les rotations sont prises en compte. Beasley and Cao (1996) ont proposé une approche basée sur la combinaison de la méthode du point intérieur et la méthode simplex pour trouver une solution LP pour les très grands problèmes. Hoffman and Padberg (1993) ont proposé une approche de branchement. Ils ont généré des rotations de manière heuristique et des coupes sont utilisées pour trouver une solution entière. Des méthodes de résolution basées sur la génération de colonnes (CG) ont été développées par Desaulniers et al. (1997), Barnhart and Shenoi (1998), Vance et al. (1997), Klabjan et al. (2001), et Subramanian and Sherali (2008). Dück et al. (2011) ont présenté un algorithme de génération de colonnes et de découpes (column- and cut-generation). L'objectif est de minimiser le cout des rotations pour couvrir un ensemble de vols. Cet algorithme s'applique aux petites et moyennes instances d'une compagnie aérienne européenne. Saddoune et al. (2013) ont utilisé une approche d'horizon fuyant basée sur la génération de colonnes pour résoudre le problème de rotations.

3.2 Problème d'affectation d'équipage (Blocs mensuels d'équipage)

Plusieurs approches ont été proposées pour le problème de l'affectation des équipages. Pour le problème d'affectation de type anonyme (bidline), Beasley and Cao (1996) ont présenté un algorithme basé sur la relaxation lagrangienne et l'optimisation de sous-gradients, intégré dans une recherche arborescente afin de trouver une solution optimale. Jarrah and Diamond (1997) ont proposé d'utiliser la génération de colonnes à priori pour le problème d'affectation anonyme dans le but de minimiser le nombre de blocs anonymes et de maximiser le temps de crédit couvert. Campbell et al. (1997) ont développé un système de génération de blocs basé sur le recuit simulé, l'objectif est de minimiser le nombre de blocs et de minimiser le temps de vol non effectué dans ces blocs. Pour résoudre le problème de génération de blocs chez Delta Air Lines, Christou et al. (1999) ont proposé une approche en deux phases basée sur des algorithmes génétiques pour la génération des blocs anonymes. L'objectif est de maximiser le nombre moyen et la qualité des blocs. Les résultats, sur des problèmes ayant jusqu'à 320 membres d'équipage, a montré que l'algorithme fournit d'importantes économies par rapport à l'approche semi-automatisée. Weir and Johnson (2004) ont présenté une méthode en trois phases utilisant la programmation en nombres entiers mixtes pour la génération des blocs. Dans la première phase, un problème en nombres entiers mixte est résolu pour fournir des blocs. L'objectif de la deuxième phase est de construire des blocs mensuels à partir de ces blocs anonymes qui couvrent toutes les rotations. Si la deuxième phase échoue, une troisième phase intègre alors les rotations non couvertes dans les blocs mensuels. De bons résultats pour un maximum de 150 membres d'équipage ont été présentés. En utilisant génération de colonnes, Gamache et al. (1998) ont développé une approche utilisant une séquence de problèmes de recouvrement résolue par génération de colonnes pour la fabrication de blocs mensuels personnalisés avec séniorité pour créer des horaires mensuels personnalisés pour les pilotes et les officiers. Les résultats pour les instances moyennes chez Air Canada ont confirmé la qualité des solutions, en termes de coûts et de temps de calcul. Gamache et al. (1999) ont aussi résolu par génération de colonnes le problème de fabrication des blocs personnalisés. Un objectif global pour des problèmes de près de 1000 agents de bord d'Air France. Ils ont produit des économies de 6% comparé aux logiciels en usage. Ces deux derniers systèmes sont maintenant utilisés dans une vingtaines de compagnies aériennes. El Moudani et al. (2001) ont proposé une approche heuristique basée sur un algorithme génétique. Cette approche produit des blocs mensuels moins coûteux qui permettent d'atteindre un niveau spécifique de la satisfaction de l'équipage. Les résultats pour les données d'une compagnie aérienne moyen-courriers sont donnés. Maenhout and Vanhoucke (2010) and Boubaker et al. (2010) ont décrit deux algorithmes heuristiques pour le problème de planification des blocs anonymes basé sur SPP. Des résultats sont donnés pour des instances avec un maximum de 150 pilotes et 800 rotations. Kasirzadeh et al. (2017) ont formulé le problème d'horaire personnalisé de l'équipage via une couverture d'ensemble et ils ont utilisé la méthode génération de colonnes comme approche de résolution.

3.3 Intégration des problèmes de rotations et de blocs mensuels

Les approches séquentielles peuvent conduire à de mauvaises solutions. Seulement quelques chercheurs ont étudié les modèles intégrés. Zeghal and Minoux (2006) ont proposé un modèle de programmation linéaire en nombres entier utilisant des contraintes de cliques pour l'intégration des rotations et l'affectation des blocs mensuels d'équipage. Ils ont résolu de petits problèmes avec 59 à 210 vols et quelques vols par rotation. Guo et al. (2006) ont décrit une approche heuristique pour intégrer partiellement les problèmes de rotations d'équipage et de blocs mensuels. Ils ont construit une série des rotations séparées par des repos hebdomadaires, puis ils ont ajusté ces rotations pour prendre en compte les demandes de l'équipage. Les résultats pour une compagnie aérienne européenne a indiqué une économie significative de coût pour les horaires. Saddoune et al. (2012) ont développé un modèle et une algorithme basé sur la génération de colonnes et l'agrégation dynamique de contraintes pour l'intégration du problème des rotations et des blocs mensuels dans le cas non personnalisé, où l'objectif est de minimiser le coût total et le nombre de pilotes. Ils ont rapporté de bons résultats pour sept ensembles de données d'une compagnie aérienne nord-américaine contenant jusqu'à 7000 vols par mois. Azadeh et al. (2013) ont introduit une métaheuristique hybride pour le problème de planification d'équipage dans laquelle l'objectif est de minimiser le coût total d'équipage tout en respectant les réglementations. Ils ont proposé deux algorithmes hybrides basés sur l'algorithme génétique et l'optimisation de colonies de fourmis. Kasirzadeh (2015) ont présenté un algorithme heuristique basé sur génération de colonnes et DCA pour l'intégration des rotations et l'affectation de blocs mensuels d'équipage.

CHAPITRE 4 ARTICLE 1: COMBINING BENDERS DECOMPOSITION AND COLUMN GENERATION FOR INTEGRATED CREW PAIRING AND PERSONALIZED CREW ASSIGNMENT PROBLEMS

Accepté pour publication dans *Transportation Science* (09 January 2019) et écrit par :

VAHID ZEIGHAMI

Polytechnique Montréal

FRANÇOIS SOUMIS

Polytechnique Montréal

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 a given set of 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

4.1 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 order of seniority. 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 Kasirzadeh et al. (2014) and Gopalakrishnan and Johnson (2005). The literature mostly focuses on either crew pairing or assignment; there are a few studies of integrated models and solution methodologies. Marsten and Shepardson (1981) and Beasley and Cao (1996) proposed a solution technique for crew pairing based on Lagrangian relaxation and subgradient optimization. Solution methodologies based on column generation (CG) were developed by Desaulniers et al. (1997), Barnhart and Shenoi (1998), Vance et al. (1997), Klabjan et al. (2001), and Subramanian and Sherali (2008). Heuristic algorithms were developed by Gershkoff (1989), Anbil et al. (1991), Anbil et al. (1992), Bixby et al. (1992), and Hoffman and Padberg (1993). Saddoune et al. (2013) used a rolling-horizon approach based on CG to solve the pairing problem.

Several approaches have been proposed for the assignment problem. Jarrah and Diamond (1997) 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. Campbell et al. (1997) developed a bidline generator system based on simulated annealing. To solve the bidline-generation problem at Delta Air Lines, Christou et al. (1999) proposed a two-phase approach based on genetic algorithms. Weir and Johnson (2004) presented a three-phase method using mixed integer programming for bidline generation. Using CG, Gamache et al. (1999) solved the preferential bidding problem with seniority to construct personalized monthly schedules for pilots and officers. Maenhout and Vanhoucke (2010) and Boubaker et al. (2010) described two heuristic algorithms for the SPP-based bidline scheduling problem. Kasirzadeh et al. (2014) 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 Papadakos (2009). Integrated aircraft routing and crew pairing was proposed by Cordeau et al. (2001), Barnhart et al. (2003b), Mercier et al. (2005), and Chen et al. (2012). Integrated flight scheduling, aircraft routing, and crew pairing was considered by Klabjan et al. (2002). Sandhu and Klabjan (2007) and Gao et al. (2009) considered the integration of fleet assignment and crew pairing. Integrated aircraft routing, crew scheduling, and flight retiming was studied by Mercier and Soumis (2007). Shao et al. (2015) considered the integration of the fleet assignment, aircraft routing, and crew pairing problems. Integrated crew pairing and assignment was studied by Zeghal and Minoux (2006); 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. Souai and Teghem (2009) proposed three heuristic approaches based on the genetic algorithm to integrate crew pairing and assignment. Saddoune et al. (2011, 2012) 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. Kasirzadeh (2015) 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 (BD) 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 pairing variables are dropped from the Benders cuts (thus giving weaker cuts) to alleviate pricing issues in the MSP. We prove that with these weak Benders cuts and the suggested dual solution of the SMP, combined BD and CG reach optimality. We also use a warm-start strategy that relies on the solution process of BD to speed up the convergence.

The remainder of this paper is organized as follows. Section 4.2 gives the problem statement, and Section 4.3 describes the extension of the crew pairing problem. Section 4.4 presents the integrated mathematical formulation, and Section 4.5 describes the solution methodology. Section 4.6 presents the datasets, experiments, and computational results. Finally, Section 4.7 provides concluding remarks and discusses future research.

4.2 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 Saddoune et al. (2013) and Kasirzadeh et al. (2014).

4.2.1 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. The maximum number of landings per duty is 5. 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 Quesnel et al. (2016).

4.2.2 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.

4.3 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 4.3. 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.

4.3.1 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 4.1 shows 11 arriving flights (terminating at black nodes) and 12 departing flights (starting at red nodes) for two consecutive days at a base. Figure 4.2 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).

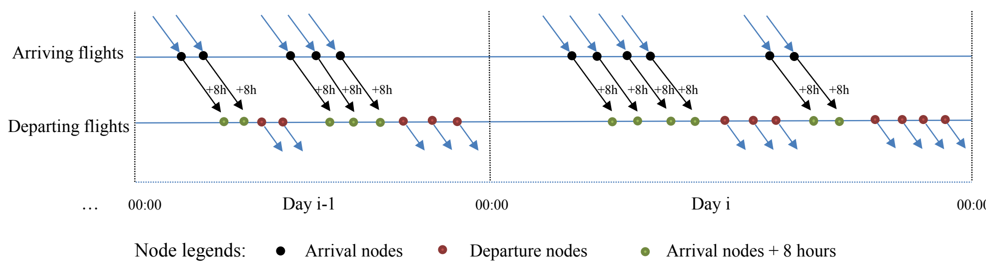

Figure 4.1 Example of arriving and departing flights at a base.

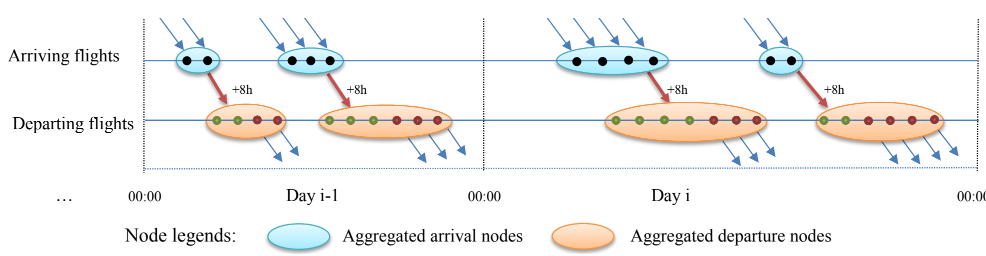

Figure 4.2 Example of node aggregation at a base.

4.3.2 Network Representation of Flow Conservation

Figure 4.3 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 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.

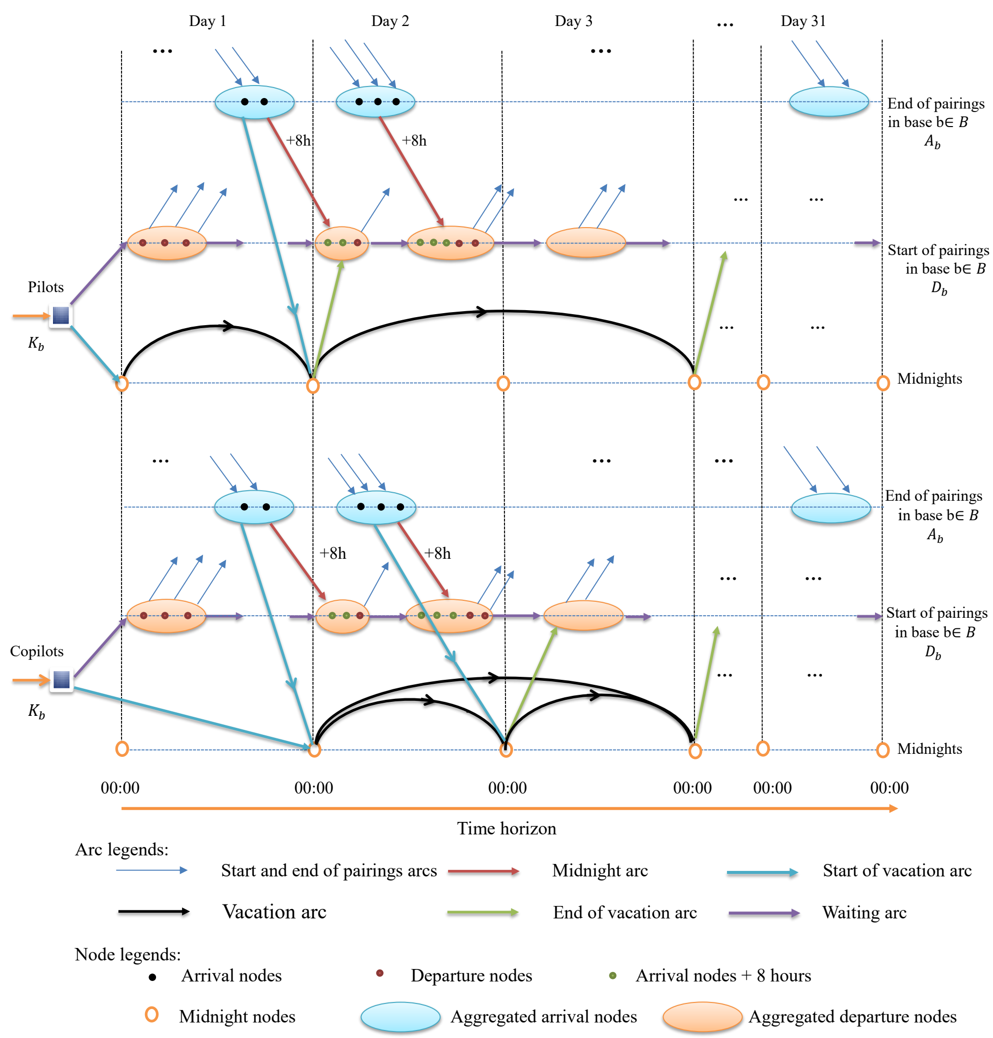

Figure 4.3 Network to incorporate pilot and copilot availability and VRs in crew pairing problem.

4.4 Mathematical Formulation

This section introduces the notation and describes the integrated crew scheduling model that includes the extended crew pairing and personalized assignment models. Different sets are introduced for (co)pilots, bases, flights, vacations, etc. These sets include persons or objects, not indices. To solve these problems we propose three approaches, described in Section 4.5. They are the sequential (SEQ) approach, the sequential approach with the extended crew pairing problem (SEQEP), and the integrated (INT) approach.

4.4.1 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 $j \in J = \{1, 2\}$, where $j = 1$ refers to the pilot assignment problem and $j = 2$ refers to the copilot assignment problem. Let L_b^j be the set of pilots (for $j = 1$) and copilots (for $j = 2$) 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 binary variable y_p is 1 if pairing $p \in P_b$ is chosen and 0 otherwise. The binary 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 binary parameter ψ_f^p is 1 if flight f is covered by pairing p and 0 otherwise. The parameter Q_b is the maximum available 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.

4.4.2 Flow Conservation Notation to Extend Pairing Model

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 start a midnight for pilots ($j = 1$) or copilots ($j = 2$). The integer variables u_e indicate the flow on the outgoing arcs $e \in E_b^j$. The binary 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 σ_i^v is 1 if $v \in V_l$ starts on day $i \in I$ and -1 if it finishes on this day.

4.4.3 Pilot (Copilot) Assignment Notation

We define S_l to be the set of feasible schedules for (co)pilot $l \in L_b^j$. The binary variable x_s is 1 if schedule $s \in S_l$ is chosen and 0 otherwise. The binary variable e_p^j is 1 if $p \in P$ is not covered by pilots ($j = 1$) or copilots ($j = 2$) and 0 otherwise. The binary 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 a_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, the proposed integrated crew scheduling model is as follows:

$$\min\Big\{ \sum_{b\in B}\sum_{p\in P_b} c_p y_p + \sum_{f\in F} c_f e_f + \sum_{j\in J}\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l} c_v r_v + \sum_{j\in J}\Big(\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l} \bar{c}_v e_v + \sum_{b\in B}\sum_{p\in P_b} \bar{c}_p e_p^j \Big)\Big\} \tag{4.1}$$

subject to

Crew pairing constraints:

$$\sum_{b\in B}\sum_{p\in P_b} \psi_f^p y_p + e_f = 1 \qquad \forall f \in F \tag{4.2}$$

$$\sum_{p\in P_b}\sum_{f\in F_p} \rho_f y_p \le Q_b \qquad \forall b \in B \tag{4.3}$$

Flow conservation constraints to extend pairing model:

$$\sum_{p\in P_b} \alpha_i^p y_p - \sum_{g\in G_b^j} \gamma_i^g h_g - \sum_{m\in M_b^j} \tau_i^m t_m = 0 \qquad \forall j \in J, \forall b \in B, \forall i \in A_b \tag{4.4}$$

$$-\sum_{p\in P_b} \eta_i^p y_p + \sum_{e\in E_b^j} \vartheta_i^e u_e + \sum_{m\in M_b^j} \zeta_i^m t_m^j + w_{i-1}^j - w_{i,b}^j = 0 \quad \forall j \in J, \forall b \in B, \forall i \in D_b \tag{4.5}$$

$$-\sum_{e\in E_b^j} \mu_i^e u_e + \sum_{g\in G_b^j} \pi_i^g h_g + \sum_{l\in L_b^j} \sum_{v\in V_l} \sigma_i^v r_v = 0 \qquad \forall j \in J, \forall b \in B, \forall i \in I \tag{4.6}$$

$$w_{\acute{i_0}}^j + h_{i_0} = k_b \qquad \forall j \in J, \forall b \in B, \acute{i_0} \in D_b, i_0 \in G_b^j \tag{4.7}$$

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

$$\sum_{l\in L_b^j} \sum_{s\in S_l} a_p^s x_s + e_p^j \geq y_p \qquad \forall j \in J, \forall b \in B, \forall p \in P_b \tag{4.8}$$

$$\sum_{s\in S_l} v_v^s x_s + e_v \geq r_v \qquad \forall j \in J, \forall b \in B, \forall l \in L_b^j, \forall v \in V_l \tag{4.9}$$

$$\sum_{s\in S_l} x_s \leq 1 \qquad \forall j \in J, \forall b \in B, \forall l \in L_b^j \tag{4.10}$$

Variable declarations:

$$y_p \in \{0,1\} \qquad \forall b \in B, \forall p \in P_b \tag{4.11}$$

$$e_f \in \{0,1\} \qquad \forall f \in F \tag{4.12}$$

$$w_i^j \in \mathbb{Z} \qquad \forall j \in J, \forall b \in B, \forall i \in D_b \tag{4.13}$$

$$u_e \in \mathbb{Z} \qquad \forall j \in J, \forall b \in B, \forall e \in E_b^j \tag{4.14}$$

$$h_g \in \mathbb{Z} \qquad \forall j \in J, \forall b \in B, \forall g \in G_b^j \tag{4.15}$$

$$t_m \in \mathbb{Z} \qquad \forall j \in J, \forall b \in B, \forall m \in M_b^j \tag{4.16}$$

$$r_v \in \{0,1\} \qquad \forall j \in J, \forall b \in B, \forall l \in L_b^j, \forall v \in V_l \tag{4.17}$$

$$e_v \in \{0,1\} \qquad \forall j \in J, \forall b \in B, \forall l \in L_b^j, \forall v \in V_l \tag{4.18}$$

$$e_p^j \in \{0,1\} \qquad \forall j \in J, \forall b \in B, \forall p \in P_b \tag{4.19}$$

$$x_s \in \{0,1\} \qquad \forall j \in J, \forall b \in B, \forall l \in L_b^j, \forall s \in S_l \tag{4.20}$$

The objective function (4.1) finds a trade-off between maximizing the number of satisfied VRs and minimizing the total cost of the pairings. Constraints (4.2) ensure that each scheduled flight is included in at most one pairing. However, the penalty cost c_f encourages the model to cover each flight. Constraints (4.3) specify an upper bound on the total available flying time per base. Constraints (4.4) impose flow conservation for arrival nodes. Constraints (4.5) enforce flow conservation for departure nodes, and Constraints (4.6) impose flow conservation for midnight nodes. Constraints (4.7) ensure that the number of (co)pilots assigned does not exceed the number available at the base. In constraints (4.7), the indexes $\acute{i_0}$ and i_0 indicate the first waiting arc and the first start-of-vacation arc at each base, respectively (see Figure 4.3). Constraints (4.8) 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 (4.9) impose the VRs for each (co)pilot. Constraints (4.10) guarantee that at most one schedule is chosen for each (co)pilot. The integrality conditions are defined by constraints (4.11)–(4.20). Figure 4.4 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.

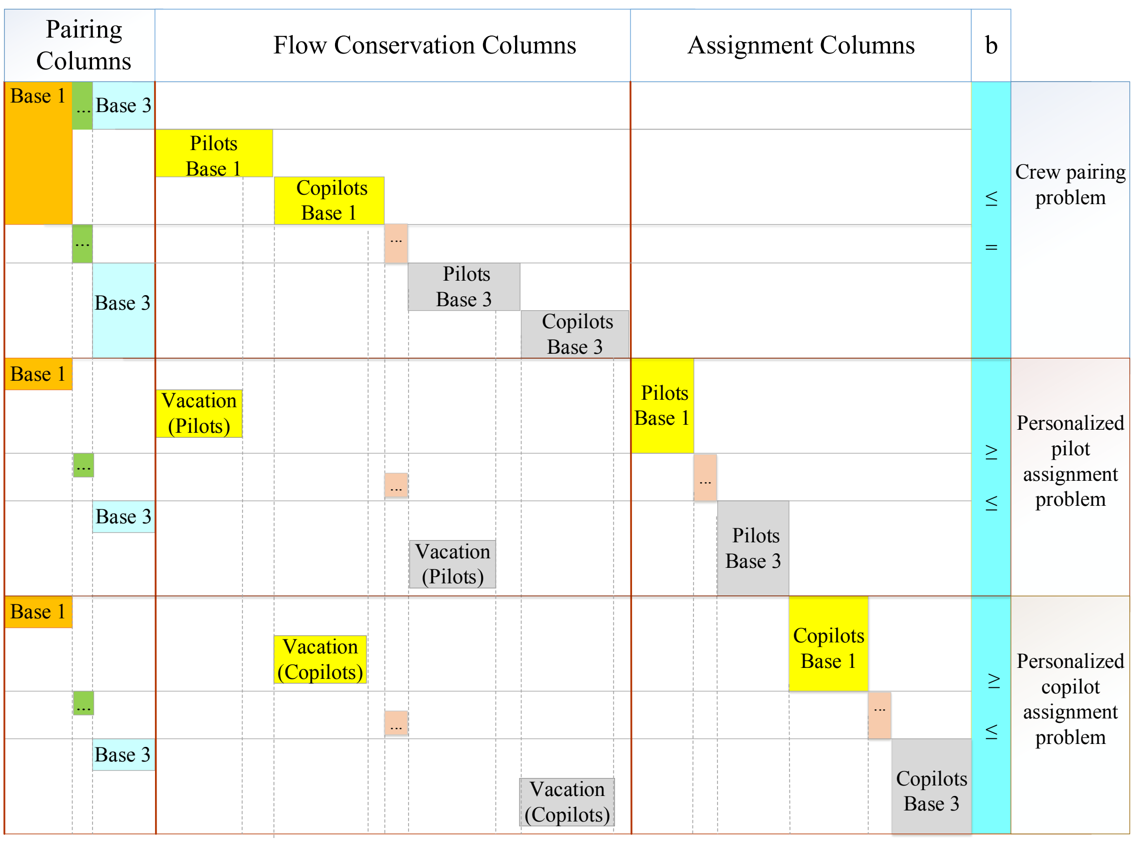

Figure 4.4 Block diagonal structure of integrated model.

Proposition 1 *Consider the following integrated crew scheduling models:*

$$(P1) \quad min\Big\{\sum_{b\in B}\sum_{p\in P_b} c_p y_p + \sum_{f\in F} c_f e_f + \sum_{j\in J}\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l} c_v r_v + \sum_{j\in J}\Big(\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l} \bar{c}_v e_v + \sum_{b\in B}\sum_{p\in P_b} \bar{c}_p e_p^j\Big)\Big\} \tag{4.21}$$

subject to (4.2)–(4.7), (4.9)–(4.20), *and*

$$\sum_{l\in L_b^j}\sum_{s\in S_l} a_p^s x_s + e_p^j \geq y_p \qquad \forall j \in J, \forall b \in B, \forall p \in P_b \tag{4.22}$$

$$(P2) \quad min\Big\{\sum_{b\in B}\sum_{p\in P_b} c_p y_p + \sum_{f\in F} c_f e_f + \sum_{j\in J}\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l} c_v r_v + \sum_{j\in J}\Big(\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l} \bar{c}_v e_v + \sum_{b\in B}\sum_{p\in P_b} \bar{c}_p e_p^j\Big)\Big\} \tag{4.23}$$

subject to (4.2)–(4.7), (4.9)–(4.20), *and*

$$\sum_{l\in L_b^j}\sum_{s\in S_l} a_p^s x_s + e_p^j = y_p \qquad \forall j \in J, \forall b \in B, \forall p \in P_b \tag{4.24}$$

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 (4.22) is a relaxation of constraint (4.24) 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 (co)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. This result justifies the use of "≥" rather than "=" in constraint (4.8). □

4.5 Solution Methodology

In this section, we describe three proposed approaches. SEQ first solves the pairing model (4.2)–(4.3) consists of the two first terms of function (4.1) as objective function and then the (co)pilot assignment model (4.8)–(4.10) using CG. SEQEP first solves the extended pairing model (4.2)–(4.7) consists of the three first terms of function (4.1) as objective function and then the (co)pilot assignment model (4.8)–(4.10) using CG. In both sequential approaches, the objective function of (co)pilot assignment model (4.8)–(4.10) consists of the last two terms of function (4.1) and variables y_p and r_v are set equal to 1 in this model. The integrated model (4.1)–(4.20) has a block diagonal structure (see Figure 4.4) and is hence suitable for mathematical decomposition. INT solves this integrated model using BD combined with CG. We now introduce the BD with a master problem and two subproblems that are solved by CG method.

4.5.1 Benders' Decomposition for Integrated Model

BD (Benders, 1962) is an iterative method for large-scale optimization problems where the coefficient matrix has a block diagonal structure. Model (4.1)–(4.20) 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 at most 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 lower bounds provided by the Benders master problem and the upper bounds provided by the subproblems are sufficiently close. See Rahmaniani et al. (2017) for a comprehensive review of BD.

4.5.2 Benders Reformulation

For given non-negative values $\bar{y}_p$ $(b \in B; p \in P_b)$ and $\bar{r}_v$ $(j \in J, b \in B, l \in L_b^j; v \in V_l)$ satisfying constraints (4.2)–(4.7), the LP relaxation of model (4.1)–(4.20) reduces to problems involving assignment variables, called the *primal pilot* $(j = 1)$ and *primal copilot* $(j = 2)$ subproblems, are as follows:

$$P_{\bar{y}_p,\bar{r}_v}(e_p^j, e_v) = \min \sum_{j\in J} \left(\sum_{b\in B} \sum_{l\in L_b^j} \sum_{v\in V_l} \bar{c}_v e_v + \sum_{b\in B} \sum_{p\in P_b} \bar{c}_p e_p^j \right) \tag{4.25}$$

$$\text{subject to} \quad \sum_{l\in L_b^j} \sum_{s\in S_l} a_p^s x_s + e_p^j \geq \bar{y}_p \qquad \forall j \in J, \forall b \in B, \forall p \in \bar{P}_b \tag{4.26}$$

$$\sum_{s\in S_l} v_v^s x_s + e_v \geq \bar{r}_v \qquad \forall j \in J, \forall b \in B, \forall l \in L_b^j, \forall v \in V_l \tag{4.27}$$

$$\sum_{s\in S_l} x_s \leq 1 \qquad \forall j \in J, \forall b \in B, \forall l \in L_b^j \tag{4.28}$$

$$x_s \geq 0 \qquad \forall j \in J, \forall b \in B, \forall l \in L_b^j, \forall s \in S_l \tag{4.29}$$

$$e_v \geq 0 \qquad \forall j \in J, \forall b \in B, \forall l \in L_b^j, \forall v \in V_l \tag{4.30}$$

$$e_p^j \geq 0 \qquad \forall j \in J, \forall b \in B, \forall p \in \bar{P}_b \tag{4.31}$$

Let $\Omega = (\Omega_p^j \geq 0 | j \in J, b \in B, p \in \bar{P}_b)$, $\theta = (\theta_v \geq 0 | j \in J, b \in B, l \in L_b^j, v \in V_l)$ and $\lambda = (\lambda_l \leq 0 | j \in J, b \in B, l \in L_b^j)$ be the dual variables associated with constraints (4.26)–(4.28), respectively. The duals of the linear relaxations of the primal (co)pilot subproblems, called the *dual pilot ($j = 1$)* and *dual copilot ($j = 2$)* subproblems, are as follows:

$$D_{\bar{y}_p,\bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l) = \max \sum_{j\in J} \left(\sum_{b\in B} \sum_{p\in \bar{P}_b} \bar{y}_p \Omega_p^j + \sum_{b\in B} \sum_{l\in L_b^j} \sum_{v\in V_l} \bar{r}_v \theta_v + \sum_{b\in B} \sum_{l\in L_b^j} \lambda_l \right) \tag{4.32}$$

$$\text{subject to} \quad \sum_{p\in \bar{P}_b} a_p^s \Omega_p^j + \sum_{v\in V_l} V_v^s \theta_v + \lambda_l \leq 0 \qquad \forall j \in J, \forall b \in B, \forall l \in L_b^j, \forall s \in S_l \tag{4.33}$$

$$\Omega_p^j \leq \bar{c}_p \qquad \forall j \in J, \forall b \in B, \forall p \in \bar{P}_b \tag{4.34}$$

$$\theta_v \leq \bar{c}_v \qquad \forall j \in J, \forall b \in B, \forall l \in L_b^j, \forall v \in V_l \tag{4.35}$$

$$\Omega_p^j \geq 0 \qquad \forall j \in J, \forall b \in B, \forall p \in \bar{P}_b \tag{4.36}$$

$$\theta_v \geq 0 \qquad \forall j \in J, \forall b \in B, \forall l \in L_b^j, \forall v \in V_l \tag{4.37}$$

$$\lambda_l \leq 0 \qquad \forall j \in J, \forall b \in B, \forall l \in L_b^j \tag{4.38}$$

Remark 1 *We note that e_v and e_p^j are not bounded from above in the primal subproblem defined by inequalities* (4.25)–(4.31)*, the primal subproblem has complete recourse, and the null vector 0 satisfies constraints* (4.33)–(4.38)*. Therefore, the dual is always feasible and bounded.*

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

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

Where P_{Δ_j} denote the sets of extreme points constraints (4.33)–(4.38). The LP relaxation of model (4.1)–(4.20) can thus be reformulated as

$$\begin{aligned}\min\Big\{&\sum_{b\in B}\sum_{p\in P_b}c_py_p+\sum_{f\in F}c_fe_f+\sum_{j\in J}\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l}c_vr_v+\\&\sum_{j\in J}\big(\max_{(\boldsymbol{\Omega},\boldsymbol{\theta},\boldsymbol{\lambda})\in P_{\Delta_j}}\big(\sum_{b\in B}\sum_{p\in P_b}\Omega_p^jy_p+\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l}\theta_vr_v+\sum_{b\in B}\sum_{l\in L_b^j}\lambda_l\big)\big)\Big\}\end{aligned}\tag{4.39}$$

subject to (4.2)–(4.7)) and (4.11)–(4.17).

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*:

$$\min\sum_{b\in B}\sum_{p\in P_b}c_py_p+\sum_{f\in F}c_fe_f+\sum_{j\in J}\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l}c_vr_v+\sum_{j\in J}Z^j\tag{4.40}$$

subject to (4.2)–(4.7), (4.11)–(4.17), and

$$Z^j\geq Z^j_{\Omega,\theta,\lambda}(y_p,r_v)=\sum_{b\in B}\sum_{p\in\bar{P}_b}\Omega_p^jy_p+\sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l}\theta_vr_v+\sum_{b\in B}\sum_{l\in L_b^j}\lambda_l\quad\forall j\in J,(\boldsymbol{\Omega},\boldsymbol{\theta},\boldsymbol{\lambda})\in P_{\Delta_j}\tag{4.41}$$

where each $Z^j_{\Omega,\theta,\lambda}(y_p,r_v)$ is a linear segment of the (co)pilot penalty functions. Constraints (4.41) are the *Benders (co)pilot optimality cuts*. Model (4.40)–(4.41) 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 (4.41) as needed to recognize an optimal solution. We denote the subsets of extreme points available at iteration $t=1,2,\dots$ by $P^t_{\Delta_1}$ and $P^t_{\Delta_2}$, for the primal pilot and copilot subproblems, respectively. Each iteration solves the relaxed extended 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 Benders subproblems (primal (co)pilot subproblems (4.25)–(4.31)). The values of the dual variables associated with constraints (4.26)–(4.28) 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.

Cordeau et al. (2001) 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. Figure 4.5 gives a flowchart for our approach. In the first phase (*phase I*), the integrality requirements are relaxed and at each iteration of the BD, the LP relaxation of the Benders master problem and the Benders 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 Benders 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 Benders subproblem variables, and the resulting integer problem is solved by generating integer Benders cuts (see Theorem 2). In this flowchart, $\bar{y}_p$ and $\bar{r}_v$ denote the optimal solution of a Benders master problem that is used as input to the Benders subproblems. 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 BD 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}. In the three-phase method proposed by Cordeau et al. (2001), the last phase is solved just once and their method is heuristic. However, in our approach, the last phase is solved by generating integer Benders cuts (see Theorem 2) until the optimality condition is reached. In other words, the approach presented in this paper achieves an exact solution.

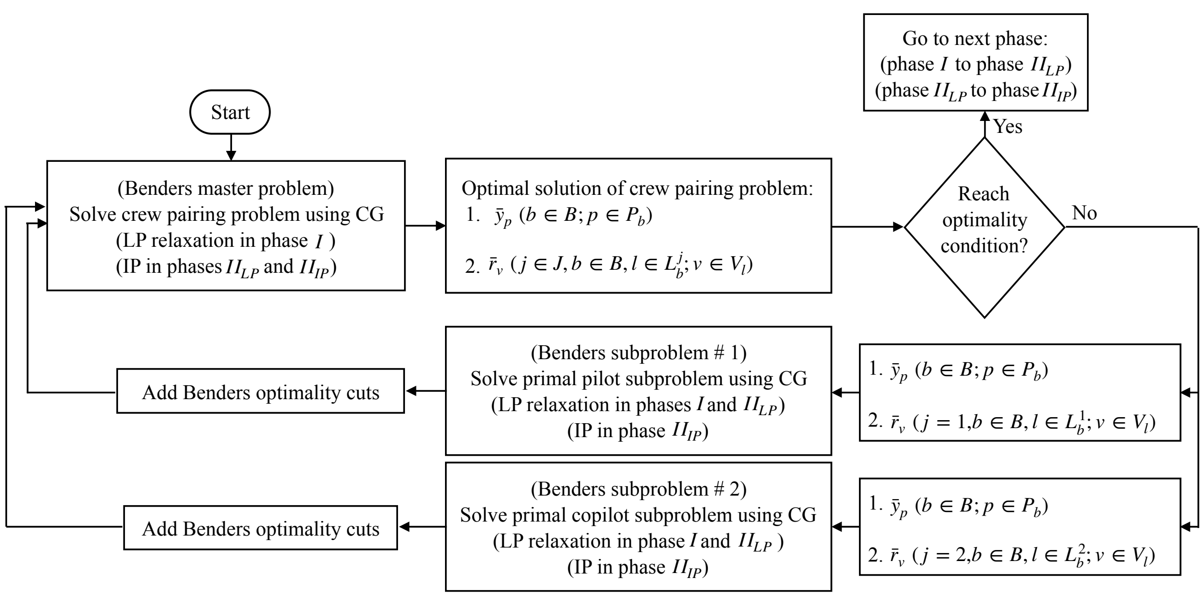

Figure 4.5 Flowchart of three-phase Benders' decomposition.

Remark 2 *Clearly, for given non-negative values* $\bar{y}_p$ *(*$b \in B; p \in P_b$*) satisfying constraints* (4.2)*, the maximum value of* $\bar{y}_p$ *is 1. Since the LP relaxation of the Benders master problem (the crew pairing problem) is solved in phase I, we add the constraints* $r_v \leq 1$ *to our model to ensure that the upper bounds on* r_v *are satisfied. In phases* II_{LP} *and* II_{IP}*, the integrality constraints are enforced for the Benders master problem, so* $r_v \in \{0, 1\}$*. Hence, during the solution process, for a given non-negative value* $\bar{r}_v$ *(*$j \in J, b \in B, l \in L_b^j; v \in V_l$*) satisfying constraints* (4.7)*, the maximum value of* $\bar{r}_v$ *is 1. Thus, the maximum value of the right-hand side of constraints* (4.26) *and* (4.27) *is 1. Also,* $\bar{c}_p \geq 0$ *and* $\bar{c}_v \geq 0$ *are the penalty costs for uncovered pairings* $p \in P$ *and uncovered vacations* $v \in V_l$*, respectively. The model* (4.25)–(4.31) *is a minimization problem, so* e_v *and* e_p^j *must take the smallest possible values. Clearly, in the worst case, these variables are equal to 1. The constraints* $e_v \leq 1$ *and* $e_p^j \leq 1$ *are redundant in the Benders subproblems* (4.25)–(4.31) *because the optimal values of* e_v *and* e_p^j *are never greater than 1. In other words, a feasible solution with* $e_v > 1$ *and* $e_p^j > 1$ *cannot be an extreme point.*

4.5.3 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 (or pricing problems). The RMP is obtained by replacing the sets P_b and S_l by the subsets $P_b^t \subseteq P_b$ and $S_l^t \subseteq S_l$ 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; see Figure 4.6. 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 4.5.3 and 4.5.3, we describe the subproblems for the crew pairing and (co)pilot assignment problems. To define these subproblems we use the network structures of Saddoune et al. (2013) and Kasirzadeh et al. (2014), 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 (Irnich and Desaulniers, 2005) as described in Section 4.5.4.

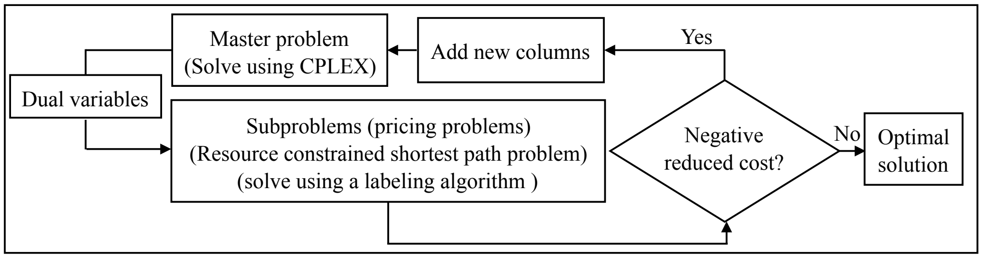

Figure 4.6 Flowchart of CG.

Crew Pairing Subproblems.

There is one subproblem (or pricing problem) for each crew base and each day, as illustrated in Figure 4.7. 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. 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 $(j \in J)$ are the dual variables for constraints (4.5). 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 $(j \in J)$ are the dual variables for constraints (4.4). 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 (4.2) and (4.3). 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 Quesnel et al. (2016). Finally, an empty arc links each waiting node to its corresponding departure node; its cost is zero. 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 $k \in K_j$ $(j \in J)$.

The crew pairing subproblems for the extended crew pairing problem are exactly the same as the standard crew pairing problem. But since the constraints in these two problems are different, the formulas for calculating reduced cost of variable y_p are different. The reduced cost of variable y_p in INT, SEQEP and SEQ is given respectively by

$$\text{INT: } \bar{C}_p = C_p - (\sum_{f \in F} e_f^p \beta_f + \sum_{f \in F_p} \rho_f \delta_b + \sum_{j \in J} \sum_{i \in A_b} \alpha_i^p \phi_i^j - \sum_{j \in J} \sum_{i \in D_b} \eta_i^p \varphi_i^j - \sum_{j \in J} \sum_{k \in K_j} \Omega_p^j \gamma_k) \quad (4.42)$$

$$\text{SEQEP: } \bar{C}_p = C_p - (\sum_{f \in F} e_f^p \beta_f + \sum_{f \in F_p} \rho_f \delta_b + \sum_{j \in J} \sum_{i \in A_b} \alpha_i^p \phi_i^j - \sum_{j \in J} \sum_{i \in D_b} \eta_i^p \varphi_i^j) \quad (4.43)$$

$$\text{SEQ: } \bar{C}_p = C_p - (\sum_{f \in F} e_f^p \beta_f + \sum_{f \in F_p} \rho_f \delta_b) \quad (4.44)$$

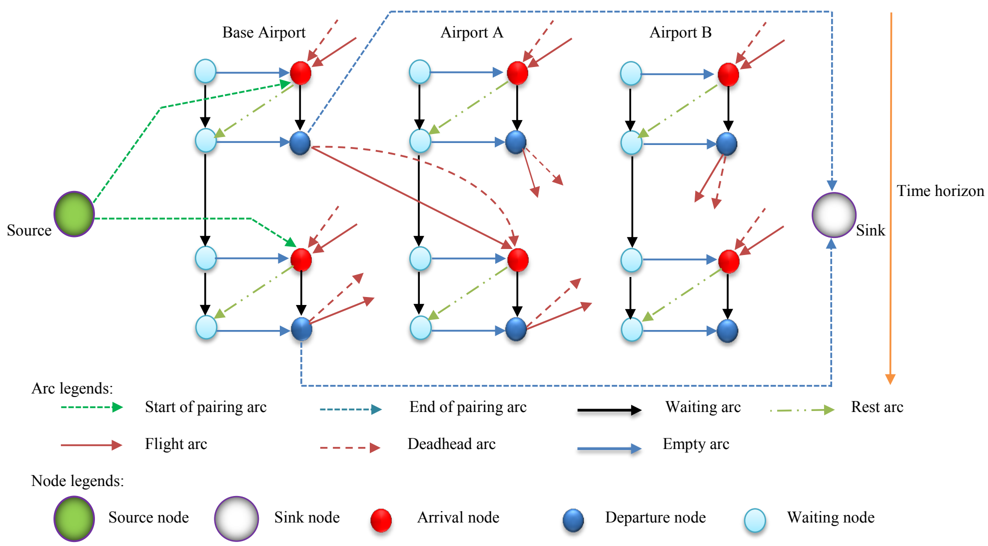

Figure 4.7 Network structure for crew pairing subproblems.

Remark 3 *The Benders master problem* (4.40)–(4.41) *is difficult to solve with CG. 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 4.7), so we cannot transfer the information on optimality cuts from the RMMP to the MSP.*

To overcome this difficulty, we introduce in Proposition 2 the concept of weak Benders cuts 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 reaches optimality in phase II_{IP}.

Proposition 2 *Since $\Omega_p^j \geq 0$, for all $p \in \bar{P}_b$ the following weak Benders cuts*

$$Z^j \geq \sum_{b \in B} \sum_{l \in L_b^j} \sum_{v \in V_l} \theta_v r_v + \sum_{b \in B} \sum_{l \in L_b^j} \lambda_l \qquad \forall j \in J, (\boldsymbol{\Omega}, \boldsymbol{\theta}, \boldsymbol{\lambda}) \in P_{\Delta_j} \tag{4.45}$$

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

Remark 4 *To use the standard crew pairing problem instead of the extended crew pairing problem for the Benders approach, it is enough to remove the flow conservation constraints* (4.4)–(4.7) *from the integrated model* (4.1)–(4.20) *and set the variable r_v equal to 1 in the constraint* (4.9)*, as we explained in Section 4.5. Hence, in the primal Benders subproblem* (4.25)–(4.31)*, instead of $\bar{r}_v$ we will have 1. Thus, the weak Benders cuts are as follows:*

$$Z^j \geq \sum_{b \in B} \sum_{l \in L_b^j} \sum_{v \in V_l} \theta_v + \sum_{b \in B} \sum_{l \in L_b^j} \lambda_l \qquad \forall j \in J, (\boldsymbol{\Omega}, \boldsymbol{\theta}, \boldsymbol{\lambda}) \in P_{\Delta_j} \tag{4.46}$$

Clearly, the right-hand side of Benders cut (4.46) is a constant value. Therefore, Benders decomposition cannot solve the proposed integrate model.

Remark 5 *In airline crew scheduling there are usually enough crew members to cover all the tasks. The first iteration of BD 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.*

Definition 1 *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 duality theory, the optimal integer solution of the primal (co)pilot subproblem is an upper bound on the dual (co)pilot subproblem, i.e., $D^*_{\bar{y}_p, \bar{r}_v}(\Omega_p^j, \theta_v, \lambda_l)_{LP} = P^*_{\bar{y}_p, \bar{r}_v}(e_p^j, e_v)_{LP} \leq P^*_{\bar{y}_p, \bar{r}_v}(e_p^j, e_v)_{IP}$.

Theorem 1 *There exist at least one integer optimal solution for the linear relaxation primal (co)pilot subproblem.*

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

Let $\bar{e}^j_p$ and $\bar{e}_v$ be the values of variables e^j_p and e_v after we solve the primal (co)pilot subproblem (4.25)–(4.31). Let $\bar{c}_p \geq 0$ and $\bar{c}_v \geq 0$ be the penalty costs for uncovered pairings and VRs. Consider the following dual solution: if pairing $p \in P_b$ is covered, i.e., $\bar{e}^j_p = 0$, then $\Omega^j_p = 0$, otherwise $\Omega^j_p = \bar{c}_p$. 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 $\lambda_l = -max_{s\in S_l}\{\sum_{p\in\bar{P}_b} a^s_p\Omega^j_p + \sum_{v\in V_l} V^s_v\theta_v\}$. Clearly, the first component of this solution satisfies constraints (4.34) and (4.36), the second satisfies constraints (4.35) and (4.37), and the third satisfies constraints (4.33) and (4.38). Hence, this is a feasible solution of the dual (co)pilot subproblem (4.32)–(4.38). The Benders solution covers all the pairings (Remark (5)), i.e., $\bar{e}^j_p = 0$, and thus $\Omega^j_p = 0$. Consequently, the IP (or LP) optimal value of the primal (co)pilot subproblem is

$$P^*_{\bar{y}_p,\bar{r}_v}(e^j_p, e_v)_{IP} = \sum_{j\in J}\sum_{b\in B}\sum_{l\in L^j_b}\sum_{v\in V_l}\bar{c}_v\bar{e}_v.$$

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, $V^s_v\theta_v = 0$ 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^s_v must be 0. Hence, $V^s_v\theta_v = 0$ and $\bar{r}_v\theta_v = \bar{c}_v\bar{e}_v$. The proposed dual solution gives $\Omega^j_p = 0$, and from conditions (i) and (ii) we have $V^s_v\theta_v = 0$, so $\lambda_l = -max_{s\in S_l}\{\sum_{p\in\bar{P}_b} a^s_p\Omega^j_p + \sum_{v\in V_l} V^s_v\theta_v\} = 0$. 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 (4.32)–(4.38) we obtain

$$D_{\bar{y}_p,\bar{r}_v}(\Omega^j_p, \theta_v, \lambda_l)_{LP} = \sum_{j\in J}\left(\sum_{b\in B}\sum_{p\in\bar{P}_b}\bar{y}_p\Omega^j_p + \sum_{b\in B}\sum_{l\in L^j_b}\sum_{v\in V_l}\bar{r}_v\theta_v + \sum_{b\in B}\sum_{l\in L^j_b}\lambda_l\right) =$$
$$\sum_{j\in J}\sum_{b\in B}\sum_{l\in L^j_b}\sum_{v\in V_l}\bar{r}_v\theta_v = \sum_{j\in J}\sum_{b\in B}\sum_{l\in L^j_b}\sum_{v\in V_l}\bar{c}_v\bar{e}_v = P^*_{\bar{y}_p,\bar{r}_v}(e^j_p, e_v)_{IP}.$$

Hence, $D_{\bar{y}_p,\bar{r}_v}(\Omega^j_p, \theta_v, \lambda_l)_{LP} = P^*_{\bar{y}_p,\bar{r}_v}(e^j_p, e_v)_{IP}$. On the other hand, since the dual (co)pilot subproblem is a maximization, clearly $D_{\bar{y}_p,\bar{r}_v}(\Omega^j_p, \theta_v, \lambda_l)_{LP} \leq D^*_{\bar{y}_p,\bar{r}_v}(\Omega^j_p, \theta_v, \lambda_l)_{LP}$. Also, we have $D^*_{\bar{y}_p,\bar{r}_v}(\Omega^j_p, \theta_v, \lambda_l)_{LP} \leq P^*_{\bar{y}_p,\bar{r}_v}(e^j_p, e_v)_{IP}$, so $D^*_{\bar{y}_p,\bar{r}_v}(\Omega^j_p, \theta_v, \lambda_l)_{LP} \leq D_{\bar{y}_p,\bar{r}_v}(\Omega^j_p, \theta_v, \lambda_l)_{LP} = P^*_{\bar{y}_p,\bar{r}_v}(e^j_p, e_v)_{IP} \leq D^*_{\bar{y}_p,\bar{r}_v}(\Omega^j_p, \theta_v, \lambda_l)_{LP}$. Thus, $D^*_{\bar{y}_p,\bar{r}_v}(\Omega^j_p, \theta_v, \lambda_l)_{LP} = P^*_{\bar{y}_p,\bar{r}_v}(e^j_p, e_v)_{IP}$ and this integer dual solution is an optimal dual solution. By strong duality, we have $D^*_{\bar{y}_p,\bar{r}_v}(\Omega^j_p, \theta_v, \lambda_l)_{LP}$

$= P^*_{\bar{y}_p,\bar{r}_v}(e^j_p, e_v)_{LP}$, so $P^*_{\bar{y}_p,\bar{r}_v}(e^j_p, e_v)_{LP} = P^*_{\bar{y}_p,\bar{r}_v}(e^j_p, e_v)_{IP}$. 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 IP dual (co)pilot subproblem* (4.32)–(4.38). *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*

$$min \sum_{b \in B} \sum_{p \in P_b} c_p y_p + \sum_{f \in F} c_f e_f + \sum_{j \in J} \sum_{b \in B} \sum_{l \in L^j_b} \sum_{v \in V_l} c_v r_v + \sum_{j \in J} Z^j \tag{4.47}$$

subject to (4.2)–(4.7), (4.11)–(4.17) *with Benders cuts* (4.45) *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 4.8 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 Benders process, the strong Benders cut (4.41) 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 (4.45) 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. Furthermore, since the number of these extreme points is finite, finite termination of the algorithm at an optimal solution is guaranteed. □

Remark 6 *Theorem 1 proves that the given dual solution is optimal when the IP primal (co)pilot subproblems* (4.25)–(4.31) *are solved. In the first two phases (*I *and* $\mathrm{II_{LP}}$*), LP relaxations of the primal (co)pilot subproblems* (4.25)–(4.31) *are solved. Therefore, the dual solution in Theorem 1 is not necessarily optimal. Hence, in these two phases, the Benders cuts of Proposition 2 are weaker than the optimal Benders cuts* (4.41). *In phase* $\mathrm{II_{IP}}$*, the IP primal (co)pilot subproblems* (4.25)–(4.31) *are solved. Thus, the dual solution of Theorem 1 is optimal. Theorem 2 proves that in this phase the Benders cuts of Proposition 2 become strong optimality cuts with the dual solution of Theorem 1.*

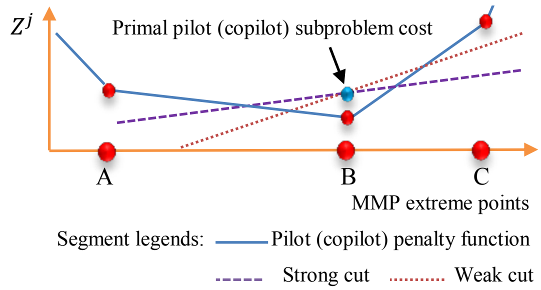

Figure 4.8 Graphical representation of penalty function in Benders master problem in phase II_{IP}

.

Personalized (Co)Pilot Assignment Subproblems.

In the personalized assignment problem, there is one subproblem (or pricing problem) for each (co)pilot, as illustrated in Figure 4.9. 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 day is 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 (4.28). 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 (4.26). 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 (4.27). The reduced cost C_s of a variable x_s is given by

$$\bar{C}_s = C_s - (\sum_{p \in \bar{P}} e_p^s \Omega_p + \sum_{v \in V_l} V_v^s \theta_v + \lambda_l) \tag{4.48}$$

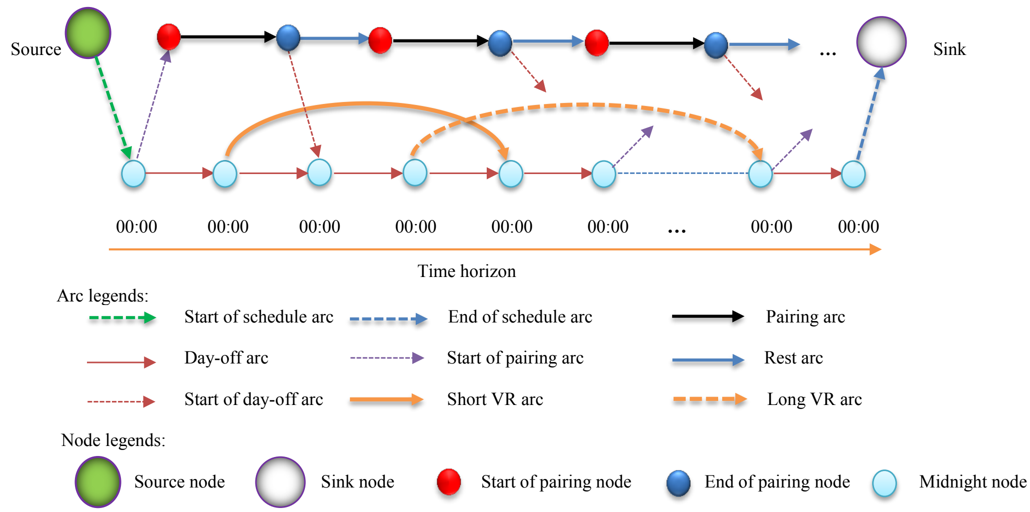

Figure 4.9 Network structure for personalized crew assignment subproblems.

4.5.4 Column Generation Subproblem Solver

The subproblems (or pricing problems) are solved using a labeling algorithm. Every feasible pairing or schedule corresponds to a path from the source node to the sink node in the above network. The feasibility constraints on the pairings and schedules are enforced via resource constraints in the network. A resource is a quantity that varies along a path and contains information about a partial path in the network. For each feasibility constraint there is a resource window at each node, and the value of every constrained resource must be within this interval (e.g., a pairing cannot exceed 4 days). Resources are consumed at arcs of the network. For the crew assignment model, we use three resources for the schedule feasibility constraints: a lower bound on the number of days off in a schedule, a maximum number of consecutive working days, and a maximum available flying time. For the crew pairing model, we use five resources: the maximum pairing duration, the maximum number of duties in a pairing, the maximum number of landings per duty, the maximum working time per duty, and the maximum duty duration. The cost of a feasible path corresponding to a pairing or schedule from the source node to the sink node is the sum of the arc costs on the path. In CG, the subproblem finds paths with a negative reduced cost. Hence, the arc costs must be updated during the solution process based on the dual variables for the master problem constraints. A label is associated with each partial path starting from the source node. Each label has two resource components and a cost component. The resource components determine the value of the resources at the last node of the partial path, and the cost component calculates the reduced cost of this path. At the source node, the components are initialized to 0. We add a new arc to the partial path if the resource consumption of the new path is within the bounds on the resource windows of the new node. For large networks, enumerating all the feasible paths can lead to long computational times at each CG iteration. To avoid this, we eliminate labels using a dominance rule: label L_1 is dominated by label L_2 if the resource components and cost component of L_1 are less than or equal to the corresponding components of L_2. We use a heuristic and an exact version of the labeling algorithm. The heuristic version considers a subset of the resource components with the reduced-cost component in the dominance rule. If no negative reduced cost paths are found, the exact version uses all the resource components.

4.5.5 Integer Solution

The linear relaxation solutions of the crew pairing and (co)pilot assignment problems may be 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: $P(f_1, f_2) = \{p \in P : \text{flight } f_2 \text{ is covered immediately after flight } f_1\}$ and $S(p_1, p_2) = \{s \in S : \text{pairing } p_2 \text{ is covered immediately after pairing } p_1\}$. 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 4.6).

4.5.6 Reducing Computational Time

We use two techniques to reduce the computational time.

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.

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 BD we use the current optimal solution as the initial solution for the next iteration. The results show that this idea works well.

4.6 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 4.1. They all involve 3 crew bases and between 1,011 and 1854 flights. The number of airports varies between 26 and 41. The numbers of pilots and copilots per base for each instance is given in Table 4.2. The pairing and schedule feasibility is restricted by the rules stated in Section 4.2 using the parameter values of Saddoune et al. (2013) for the pairings and of Kasirzadeh et al. (2014) for the schedules. 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 4.1 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. The datasets are available at www.gerad.ca/en/papers/G-2017-41.

Table 4.1 Instance Characteristics

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

Table 4.2 Number of Pilots and Copilots per Base

	Base 1		Base 2		Base 3		Total	
Instance	No. Pilots	No. Copilots	No. Pilots	No. Copilots	No. Pilots	No. Copilots	No. Pilots	No. Copilots
1	7	7	20	20	6	6	33	33
2	10	10	9	9	15	15	34	34
3	10	10	30	30	7	7	47	47

4.6.1 Analysis of Computational Refinements

This section discusses the effect of the aggregation and the impact of the warm start strategy, as described in Sections 4.3 and 4.5.6. Table 4.3 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. We solved each aggregated instance with both the basic algorithm and the *refined* version that includes warm starting; see Table 4.4. 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 4.4 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 4.3 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

Table 4.4 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

4.6.2 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 4.5. 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, the size of the (co)pilot assignment problems in phase I is much larger than that in 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 4.5 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

4.6.3 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 4.6 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, and it is calculated when the algorithm terminates. To find a lower bound, we stop the CG at each branching node if the optimal value of the RMP improves by less than 0.1% over five iterations; see Section 4.5.3. The lower bound will improve during the branch and bound process since we explore many nodes. The number of branching nodes for each instance is given in Table 4.7. To obtain integer solutions, we use two branching heuristics, column fixing and inter-task fixing; see Section 4.5.5. 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.

The crew pairing models used in the three approaches are not exactly the same, because the number of constraints is different (Benders cuts and flow conservation constraints), and there are different terms in the objective function, as described in Sections 4.5. On the other hand, since the values of $\bar{y}_p$ and $\bar{r}_v$ in the proposed approaches may be different, the structure of the (co)pilot assignment model used in the three approaches is not exactly the same but very similar. Therefore, the integrality gaps for INT, SEQEP, and SEQ for the crew pairing and (co)pilot assignment problems are not comparable. We observe that, on average, the integrality gaps of INT are smaller than those of SEQEP and SEQ for the assignment problems.

Table 4.6 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

4.6.4 Number of Branching Nodes, Iterations, and Columns

Table 4.7 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 it performs one iteration. 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 4.7 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

4.6.5 Coverage of VRs

We now analyze the performance of the approaches in terms of the percentage of satisfied VRs; see Table 4.8. 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 4.8 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 4.8 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

4.6.6 Coverage of Flights and Pairings

We define large penalty costs to ensure that the percentages of uncovered flights and pairings are small. The results of this experiment are given in Table 4.9. 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.

Table 4.9 Uncovered Flights and Pairings (%)

	Uncovered flights	Uncovered pairings	
	Crew pairing	Pilot assignment	Copilot assignment
SEQ			
Instance 1	0.00	0.00	0.00
Instance 2	0.00	0.32	0.32
Instance 3	0.00	0.60	0.60
SEQEP			
Instance 1	0.00	0.00	0.00
Instance 2	0.00	0.00	0.00
Instance 3	0.00	0.00	0.00
INT			
Instance 1	0.00	0.00	0.00
Instance 2	0.00	0.00	0.00
Instance 3	0.00	0.00	0.00

4.6.7 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 4.10. 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 4.5.6. 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. These improvements can be explained by the fact that SEQ does not take into account the crew assignment constraints and objective during the building of the pairings. Hence, the pairings generated by the crew pairing problem may not be suitable for the objective of the crew assignment problem. Therefore, the number of uncovered VRs and pairings in SEQ is more than INT and SEQEP. The objective function of the crew assignment problem is to minimize the cost of the uncovered pairings and unsatisfied VRs. Hence, the cost of crew assignment problem that obtained by SEQ is more than those obtained by INT and SEQEP. For example, INT decreases the pilot (copilot) costs by 69.52% (69.23%), in comparison with SEQ. The computational time increases by a factor of 3.57.

Table 4.10 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

4.7 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 BD 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.

CHAPITRE 5 DIFFICULTÉ DE L'UTILISATION DE DÉCOMPOSITION DE BENDERS ET DE GÉNÉRATION DE COLONNES CONSIDÉRANT DES VOLS PRÉFÉRÉS POUR LES PILOTES ET LES COPILOTS

Résumé

Dans le chapitre 4, nous avons proposé un modèle intégré pour les problèmes de rotation d'équipage et d'affectation d'équipage personnalisée et une approche intégrée basée sur combinaison des approches de décomposition de Benders et de génération de colonnes. Nous considérons un ensemble de vacances préférés pour chaque pilote et copilote par mois. Dans un cas plus général de problème d'horaire des équipages des compagnies aériennes, chaque pilote et copilote ont la possibilité de choisir un ensemble de vols préférés par mois. Dans ce chapitre, nous étudions la difficulté d'utiliser la méthode proposée dans le chapitre précédent si nous voulons considérer un ensemble de vols préférés pour chaque pilote et copilote. En outre, nous étudions la difficulté d'utiliser la relaxation lagrangienne pour résoudre ce problème.

5.1 La difficulté d'utiliser des coupes Benders fortes définies par les variables duales des rotations

Dans la méthode proposée dans le chapitre précédent, les rotations sont générées par le problème maître Benders et les blocs mensuels pour les pilotes et copilotes sont générés par les sous-problèmes Benders. Les coupes Benders sont définies comme suit:

$$Z^j \geq \sum_{b\in B}\sum_{p\in \bar{P}_b} \Omega_p^j y_p + \sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l} \theta_v r_v + \sum_{b\in B}\sum_{l\in L_b^j} \lambda_l \quad \forall j \in J, (\boldsymbol{\Omega}, \boldsymbol{\theta}, \boldsymbol{\lambda}) \in P_{\Delta_j} \tag{5.1}$$

Ces coupes doivent être placées dans le MMP, et leurs variables duales doivent être transférées au MSP pour évaluer les valeurs de coût de la nouvelle colonne. Pour refléter les informations du MMP (RMMP) dans MSP, les coûts d'arc dans MSP doivent être modifiés par la solution duale optimale pour le RMMP. Le coût réduit de la variable y_p est donné par:

$$\bar{C}_p = C_p - (\sum_{f\in F} e_f^p \beta_f + \sum_{f\in F_p} \rho_f \delta_b + \sum_{j\in J}\sum_{i\in A_b} \alpha_i^p \phi_i^j - \sum_{j\in J}\sum_{i\in D_b} \eta_i^p \varphi_i^j - \sum_{j\in J}\sum_{k\in K_j} \Omega_p^j \gamma_k) \tag{5.2}$$

Le coût d'un arc de vol est $-\beta_f - \rho_f \delta_b$. Les coûts des arcs d'entrée et de sortie de chaque réseau dans les sous-problèmes de rotation sont fixés à $-(\phi_i^1 + \phi_i^2)$ et $\varphi_i^1 + \varphi_i^2$, respectivement.

Cependant, il n'existe pas d'arcs correspondant à des rotations dans les sous-problèmes de génération de rotation, donc nous ne pouvons pas transférer les informations sur les optimalité coupes du RMMP au MSP.

5.2 La difficulté d'utiliser des coupes Benders faibles

Pour surmonter la difficulté d'utiliser des coupes Benders (5.1), nous avons introduit le concept de coupes Benders faibles (4.45) qui ne prennent pas en compte la variable duale correspondant aux rotations ($\Omega_p^j = 0$) et défini comme suit

$$Z^j \geq \sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l}\theta_v r_v + \sum_{b\in B}\sum_{l\in L_b^j}\lambda_l \quad \forall j\in J, (\boldsymbol{\Omega},\boldsymbol{\theta},\boldsymbol{\lambda})\in P_{\Delta_j} \tag{5.3}$$

Le théorème 2 du chapitre 4 prouve qu'avec ces coupes Benders faibles et la solution duale suggérée des sous-problèmes pilote et copilote, l'approche de Benders atteint l'optimalité. Pour considérer les PFs, un coût négatif est associé à chaque PF dans les sous-problèmes pilote et copilote pour inciter ces modèles à satisfaire les PF autant que possible. n_s est le nombre de vols préférés dans $s \in S_l$ et M représentent un coût négatif pour couvrir chaque vol préféré. c_s est le coût $s \in S_l$, où $c_s = Mn_s$. Le modèle d'affectation pilote (copilote) proposé est le suivant:

$$P_{\bar{y}_p,\bar{r}_v}(x_s, e_p^j, e_v) = \min \sum_{j\in J}\left(\sum_{b\in B}\sum_{l\in L_b^j}\sum_{s\in S_l} c_s x_s + \sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l}\bar{c}_v e_v + \sum_{b\in B}\sum_{p\in P_b}\bar{c}_p e_p^j\right) \tag{5.4}$$

$$\text{sujet à} \quad \sum_{l\in L_b^j}\sum_{s\in S_l} a_p^s x_s + e_p^j = \bar{y}_p \qquad \forall j\in J, \forall b\in B, \forall p\in \bar{P}_b \tag{5.5}$$

$$\sum_{s\in S_l} v_v^s x_s + e_v \geq \bar{r}_v \qquad \forall j\in J, \forall b\in B, \forall l\in L_b^j, \forall v\in V_l \tag{5.6}$$

$$\sum_{s\in S_l} x_s \leq 1 \qquad \forall j\in J, \forall b\in B, \forall l\in L_b^j \tag{5.7}$$

$$x_s \geq 0 \qquad \forall j\in J, \forall b\in B, \forall l\in L_b^j, \forall s\in S_l \tag{5.8}$$

$$e_v \geq 0 \qquad \forall j\in J, \forall b\in B, \forall l\in L_b^j, \forall v\in V_l \tag{5.9}$$

$$e_p^j \geq 0 \qquad \forall j\in J, \forall b\in B, \forall p\in \bar{P}_b \tag{5.10}$$

Soit $\Omega = (\Omega_p^j | j\in J, b\in B, p\in \bar{P}_b)$, $\theta = (\theta_v \geq 0 | j\in J, b\in B, l\in L_b^j, v\in V_l)$ et $\lambda = (\lambda_l \leq 0 | j\in J, b\in B, l\in L_b^j)$ sont les duales variables associées aux contraintes (5.5) - (5.7), respectivement. Les duales des relaxations linéaires des sous-problèmes (co)pilotes sont les suivants:

$$D_{\bar{y}_p,\bar{r}_v}(\Omega_p^j,\theta_v,\lambda_l) = \max\sum_{j\in J}\Bigg(\sum_{b\in B}\sum_{p\in\bar{P}_b}\bar{y}_p\Omega_p^j + \sum_{b\in B}\sum_{l\in L_b^j}\sum_{v\in V_l}\bar{r}_v\theta_v + \sum_{b\in B}\sum_{l\in L_b^j}\lambda_l\Bigg) \tag{5.11}$$

$$\text{sujet à}\quad \sum_{p\in\bar{P}_b} a_p^s\Omega_p^j + \sum_{v\in V_l} V_v^s\theta_v + \lambda_l \le c_s \qquad \forall j\in J,\forall b\in B,\forall l\in L_b^j,\forall s\in S_l \tag{5.12}$$

$$\Omega_p^j \le \bar{c}_p \qquad \forall j\in J,\forall b\in B,\forall p\in\bar{P}_b \tag{5.13}$$

$$\theta_v \le \bar{c}_v \qquad \forall j\in J,\forall b\in B,\forall l\in L_b^j,\forall v\in V_l \tag{5.14}$$

$$\Omega_p^j \quad \text{libre} \qquad \forall j\in J,\forall b\in B,\forall p\in\bar{P}_b \tag{5.15}$$

$$\theta_v \ge 0 \qquad \forall j\in J,\forall b\in B,\forall l\in L_b^j,\forall v\in V_l \tag{5.16}$$

$$\lambda_l \le 0 \qquad \forall j\in J,\forall b\in B,\forall l\in L_b^j \tag{5.17}$$

Remarque 7 *Si nous considérons un ensemble de PFs pour chaque pilote et copilote, la duale valeur de certaines rotations peut être négative. Voir l'exemple suivant: supposons que nous avons un pilote et un rotation, et la rotation inclut les vols i et j, vols i être de vol préféré pour ce pilote. Supposons que dans la solution optimale du sous-problème pilote, cette rotation et toutes les vacances préférées sont couverts. Ainsi, la valeur des variables* e_v *et* e_p^j *sont égaux à zéro. Par conséquent, la valeur optimale du sous-problème pilote est la suivante:* $P_{\bar{y}_p,\bar{r}_v}(x_s^*,e_p^j,e_v) = c_s = Mn_s = M \le 0$*. Considérons la solution duale suivante:* $\theta_v = 0$ *et* $\lambda_l = 0$*, et* $\Omega_p^j = c_s = M \le 0$*. Clairement, le premier composant de cette solution satisfait les contraintes* (5.14) *et* (5.16)*, la deuxième satisfait la contrainte* (5.17)*, la troisième satisfait les contraintes* (5.12), (5.15) *and* (5.17)*. Donc, c'est une solution réalisable du duale sous-problème (co)pilote* (5.11)–(5.17)*. Si nous plaçons cette solution dans la fonction objectif du duale sous-problème (co)pilote* (5.11) - (5.17)*, la valeur optimale du duale sous-problème pilote est* $D_{\bar{y}_p,\bar{r}_v}(\Omega_p^j,\theta_v,\lambda_l) = \sum_{b\in B}\sum_{p\in\bar{P}_b}\bar{y}_p\Omega_p^j = \Omega_p^j = M$*. Donc, nous avons:* $D_{\bar{y}_p,\bar{r}_v}(\Omega_p^j,\theta_v,\lambda_l) = P_{\bar{y}_p,\bar{r}_v}(x_s^*,e_p^j,e_v) = M \le 0$*. Selon la théorie de la dualité, la valeur optimale du problème primal est égale au problème dual. Par conséquent, cette solution est une solution optimale pour duale sous-problème pilote* (5.11)–(5.17)*. Puisque* $\Omega_p^j = M \le 0$*, donc, la duale valeur de cet rotation est négative.*

Proposition 3 *Si nous considérons un ensemble de PFs pour chaque pilote et copilote, les coupes Benders faibles* (5.3) *ne sont pas valides.*

Preuve. Soit $\Omega = (\Omega_p^j | j\in J, b\in B, p\in\bar{P}_b)$ sont les doubles variables des contraintes (4.26). Puisque Ω_p^j peut prendre une valeur positive ou négative, les coupes Benders faibles ne sont pas valides.

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

En outre, dans les coupes Benders faibles la duale valeur de chaque rotation est égale à 0 ($\Omega_p^j = 0$), donc, nous ne pouvons pas transférer des informations de vols préférés des sous-problèmes Benders au problème maître Benders. □

5.3 La difficulté d'utiliser des coupes Benders définies par les variables duales des vols

La difficulté d'utiliser des coupes Benders fortes est qu'il n'y a pas d'arcs correspondant aux rotations dans les sous-problèmes de génération de rotation, nous ne pouvons pas transférer les informations sur les coupes optimalité du RMMP au MSP. Mais, à chaque vol, il existe un arc dans les sous-problèmes de génération de rotation (MSP). Donc, il semble que si nous trouvons la duale valeur de chaque vol, nous pouvons envoyer cette information des sous-problèmes de Benders aux problèmes maître Benders. Nous prouvons que même nous trouvons la duale valeur de chaque vol, la combinaison des approches de décomposition de Benders et de génération de colonnes ne fonctionnent pas quand on considère les PFs. Une difficulté est qu'il ne fournit pas une solution complète duale pour chaque vol. Il fournit une solution duale agrégée pour chaque rotation qui doit être désagrégée. Pour ce faire, le système linéaire suivant, doit être résolu:

$$\sum_{f\in F} e_f^p \alpha_f^j = \Omega_p^j \quad \forall j \in J, \forall p \in \bar{P} \tag{5.18}$$

où α_f^j est la dual valeur du vol f. En remplaçant Ω_p^j avec $\sum_{f\in F} e_f^p \alpha_f^j$ dans les coupes Benders (5.1), les coupes Benders sont

$$Z^j \geq \sum_{p\in P}\sum_{f\in F} e_f^p \alpha_f^j y_p + \sum_{l\in L}\sum_{v\in V_l} \theta_v r_v + \sum_{l\in L} \lambda_l \quad \forall j \in J, (\boldsymbol{\Omega}, \boldsymbol{\theta}, \boldsymbol{\lambda}) \in P_{\Delta_j} \tag{5.19}$$

Proposition 4 *Les coupes Benders* (5.19) *obtenu avec la dual valeur de vols ne peut pas traiter avec des PFs.*

Preuve. Pour simplifier, nous supposons que nous avons juste PFs. Contrainte (4.2) assurer que chaque vol doit être couvert par exactement une rotation et il définit comme suit

$$\sum_{p \in P} e_f^p y_p = 1 \quad \forall f \in F \tag{5.20}$$

donc nous avons

$$\sum_{p \in P} \sum_{f \in F} e_f^p \alpha_f^j y_p = \sum_{f \in F} \sum_{p \in P} e_f^p \alpha_f^j y_p = \sum_{f \in F} \alpha_f^j (\sum_{p \in P} e_f^p y_p) = \sum_{f \in F} \alpha_f^j = A^j \text{ (valeur fixe et indépendante de } y_p)$$

Cela signifie que le premier terme dans les coupes Benders (5.19) (i.e., $\sum_{p \in P} \sum_{f \in F} e_f^p \alpha_f^j y_p$) est égal à une valeur fixe et indépendant des rotations sélectionnés. Par conséquent, nous avons

$$\sum_{p \in P} \sum_{f \in F} e_f^p \alpha_f^j y_p = A^j \Longrightarrow Z^j \geq A^j + \sum_{b \in B} \sum_{l \in L_b^j} \lambda_l \quad \forall j \in J, (\boldsymbol{\Omega}, \boldsymbol{\theta}, \boldsymbol{\lambda}) \in P_{\Delta_j} \tag{5.21}$$

Clairement, le côté droit de coupé Benders (5.21) est une valeur constante. Par conséquent, les coupes Benders ne peuvent pas traiter avec PFs. En fait, dans ce cas, l'approche INT est exactement l'approche SEQ et la solution de la variable y_p en deux itérations consécutives ne change pas. □

Remarque 8 *L'approche SEQ ne peut pas trouver la solution optimale globale. Voir l'exemple suivant:*
Supposons que nous avons deux pilotes: A et B. Dans la solution optimale du problème de rotation de l'équipage, laisser les vols i et j sont couverts par la rotation P_1 (i.e., $P_1 = \{i, j\}$); et vols m et n sont couverts par le rotation P_2 (i.e., $P_2 = \{m, n\}$). Soit les vols i et n être les vols préférés pour pilote A et les vols j et m sont les vols préférés pour pilote B. Supposons que, dans la solution optimale du problème d'affectation du pilote, la rotation P_1 est couvert par le pilote A; et le rotation P_2 est couvert par le pilote B. Donc, dans la solution optimale du problème d'affectation de pilote, deux vols préférés sont couverts. Considérons deux rotations possibles: $P_3 = \{i, n\}$ et $P_4 = \{m, j\}$ tels que:

$$C_1 P_1 + C_2 P_2 = C_3 P_3 + C_4 P_4 \tag{5.22}$$

Où C_i est le coût de rotation i. Nous utilisons ces deux rotations dans le problème d'affectation du pilote, et supposer que la rotation P_3 est couvert par le pilote A; et le rotaion P_4 est couvert par le pilote B. Par conséquent, dans la solution du problème d'affectation de pilote, quatre vols préférés sont couverts.

5.4 La difficulté d'utiliser la méthode de relaxation lagrangienne

Pour surmonter la difficulté évoquée, nous proposons un nouveau modèle intégré. La figure 5.1 présente la structure générale du nouveau modèle qui inclut les quatre problèmes d'optimisation: rotations pilote, rotations copilote, bloc mensuels pilotes et bloc mensuels copilotes. Pour réduire la propagation des perturbations pendant l'opération, les rotations dans les deux problèmes doivent être semblables autant que possible.

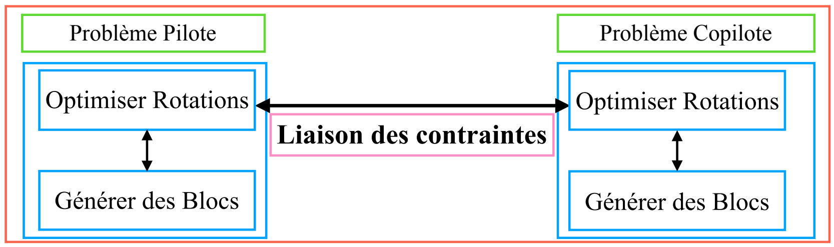

Figure 5.1 Modèle intégré pour pilotes et copilotes

Remarque 9 *La combinaison de décomposition de Benders et de génération de colonnes et les approches séquentielles pour résoudre ces problèmes produisent les mêmes rotations pour les pilotes et les copilotes.*

Ce modèle intégré a une structure diagonale de bloc et est donc adapté à la décomposition mathématique, en particulier à la relaxation lagrangienne. La relaxation lagrangienne classique utilisant des méthodes de subgradient ou de bundle nécessite généralement de nombreuses itérations pour atteindre une solution optimale; voir Crainic et al. (2001) et Akhavan Kazemzadeh et al. (2018). Par exemple, Akhavan Kazemzadeh et al. (2018) a proposé une méthode pour le problème de conception de réseau basée sur la relaxation lagrangienne et l'optimisation de sous-gradient. Ils ont considéré deux ensembles d'instances différents avec deux et trois couches. Le nombre de contraintes pour le premier ensemble d'instances varie entre 20612-30664 et pour le deuxième ensemble d'instances varie entre 42768-64162. Ils fixent le nombre maximal d'itérations de sous-programmes sur 3000.

Le problème que nous considérons dans cette thèse a la même structure que Akhavan Kazemzadeh et al. (2018). Le nombre de contraintes pour nos problèmes varie entre 8738-164289. La résolution de chaque problème (pilote et copilote) prend beaucoup de temps. Par exemple, la plus grande instance, le temps processeur de chaque problème est proche de 10 heures. Si nous avons besoin de 100 itérations de sous-gradient pour trouver une bonne solution, nous avons besoin de $100 * 10 = 1000$ heures. Par conséquent, l'utilisation de la relaxation lagrangienne avec des méthodes de subgradient ou de bundle ne peut en pratique résoudre notre problème.

5.5 Conclusion

Dans ce chapitre, nous avons étudié la difficulté d'utiliser la combinaison de décomposition de Benders et de génération de colonnes que nous avons développée dans le chapitre précédent quand on considère un ensemble de PFs et VRs pour chaque pilote et copilote. Nous avons étudié trois types de coupes Benders: coupes Benders fortes définies par les variables duales des rotations, coupes Benders faibles, et cupes Benders définies par les variables duales des vols. Nous avons montré qu'aucun d'entre eux ne pouvait s'occuper des PFs. De plus, la décomposition lagrangienne en utilisant des méthodes ou sous-gradient ou de bundle nécessite généralement de nombreuses itérations pour aboutir à une solution optimale, donc, en pratique ne peut pas résoudre notre problème. Pour surmonter ces difficultés, Dans le chapitre suivant, nous présentons un nouveau modèle intégré où chaque pilote et copilote demande un ensemble de VRs et de PFs par mois. Pour résoudre ce modèle, nous développons une approche intégrée basée sur l'alternance de la décomposition Lagrangienne, génération de colonnes et l'agrégation dynamique de contraintes. L'approche développée utilise la décomposition lagrangienne de façon alternée proposant une nouvelle façon de mettre à jour les multiplicateurs de Lagrange. Nous montrerons qu'avec cette nouvelle méthode quelques itérations suffisent pour trouver une solution (presque) optimale pour le problème grand et complexe.

CHAPITRE 6 ARTICLE 2: COMBINING ALTERNATING LAGRANGIAN DECOMPOSITION, COLUMN GENERATION AND DYNAMIC CONSTRAINT AGGREGATION FOR INTEGRATED CREW PAIRING AND PERSONALIZED ASSIGNMENT PROBLEMS FOR PILOTS AND COPILOTS SIMULTANEOUSLY

Article soumis a *Transportation Science* (NOVEMBER 2018) et écrit par :

VAHID ZEIGHAMI

Polytechnique Montréal

FRANÇOIS SOUMIS

Polytechnique Montréal

Abstract

The airline crew scheduling problem involves determining schedules for airline crew members such that all the scheduled flights over a planning horizon (usually a month) are covered and the constraints are satisfied. Because of its complexity, this problem is usually solved sequentially in two main steps: the crew pairing followed by the crew assignment. However, finding a globally optimal solution via the sequential approach may be impossible because the decision domain of the crew assignment problem is reduced by decisions made in the pairing problem. This study considers the crew scheduling problem in a personalized context where each pilot and copilot requests a set of preferred flights and vacations each month. We propose a model that completely integrates the crew pairing and personalized assignment problems to generate personalized monthly schedules for a given set of pilots and copilots simultaneously in a single optimization step. The model keeps the pairings in the two problems as similar as possible so that the propagation of perturbations arising during the operation is reduced. We develop an integrated algorithm that combines alternating Lagrangian decomposition, column generation, and dynamic constraint aggregation. We conduct computational experiments on a set of real instances from a major US carrier. Our integrated approach produces significant cost savings and better satisfaction of crew preferences compared with the traditional sequential approach.

Keywords: Integrated crew scheduling problem; column generation; dynamic constraint aggregation; alternating Lagrangian decomposition

6.1 Introduction

The airline crew scheduling problem is complex. For the airline industry, the crew costs are the largest after the fuel costs (Barnhart et al., 2003a). Hence, even slight improvements in the quality of the schedule can have a significant financial benefit. Given a set of flights to be operated by the same aircraft fleet, defined by departure/arrival stations and fixed departure/arrival dates and times, the crew scheduling problem constructs individual schedules for a set of available crew members. Because of its complexity, this problem is usually solved in two steps: crew pairing followed by crew assignment.

A pairing is a sequence of flights, connections, and rests starting and ending at the same crew base. Given the scheduled flights, the crew pairing problem involves determining a set of feasible pairings such that the cost of the pairings is minimized and each flight is covered exactly once.

In the crew assignment problem, the goal is to build monthly schedules from these pairings for each individual crew member such that every pairing is covered exactly once while respecting all the safety and collective agreement rules. There are three different approaches for this problem. The bidline approach, used in some North American airlines, produces anonymous monthly schedules, which the crew members then select in seniority order. The personalized approach with strict seniority, which is becoming more popular in North America, sequentially maximizes the satisfaction of the employees in decreasing order of seniority. The personalized approach with a global objective, which is often employed by European airlines and has recently started to be used by American airlines, produces the monthly schedules simultaneously to maximize the sum of the personal satisfactions without any advantages for seniority.

The sequential approach to crew scheduling considerably reduces the complexity of the process but may lead to significantly suboptimal solutions since the schedule constraints and objectives are not taken into account during the construction of the pairings. In fact, the pairing construction stage cannot determine the best pairings for the crew assignment stage. Hence, it is difficult to maximize the satisfaction of the preferences at the assignment stage, and we may not find the globally optimal solution of the crew scheduling problem.

The integration of two or more other airline planning problems involving crew scheduling has been studied; see the recent survey by Kasirzadeh et al. (2017). Cordeau et al. (2001) and Mercier et al. (2005) proposed an algorithm based on column generation (CG) and Benders decomposition (BD) for integrated aircraft routing and crew pairing. Integrated flight scheduling, aircraft routing, and crew pairing was considered by Klabjan et al. (2002). Cohn and Barnhart (2003) developed an extended crew pairing model that integrates crew scheduling and maintenance routing decisions. Sandhu and Klabjan (2007) introduced a model that completely integrates the fleeting and crew pairing stages. They proposed two approaches, the first based on a combination of Lagrangian relaxation and CG and the second a BD approach. Integrated aircraft routing, crew scheduling, and flight retiming was studied by Mercier and Soumis (2007). Papadakos (2009) introduced a set covering model for integrated fleet assignment, maintenance routing, and crew pairing and a method based on BD combined with CG. Gao et al. (2009) studied the integrated fleet and crew robust planning problem, providing fleet assignment solutions. Shao et al. (2015) considered integrated fleet assignment, aircraft routing, and crew pairing. Cacchiani and Salazar-González (2016) proposed two mixed integer linear programming models for integrated fleet assignment, aircraft routing, and crew pairing.

In recent decades, various models and methods have been introduced for the crew scheduling problem. Sequential approaches may lead to poor solutions, but only a few researchers have investigated integrated models. Zeghal and Minoux (2006) proposed an integer linear programming model using clique constraints for integrated pairing and bidline assignment. They consider small problems with 59 to 210 flights and a few flights per pairing. Guo et al. (2006) described a partially integrated crew scheduling approach based on pairing-chain generation. They construct a series of pairing chains containing weekly rests and then adjust these pairings to take into account the crew requests and prescheduled activities. Saddoune et al. (2012) developed a model and algorithm based on CG and dynamic constraint aggregation (DCA) for integrated crew pairing and assignment, where the objective is to minimize the total cost and the number of pilots. Kasirzadeh (2015) presented a heuristic algorithm based on CG and DCA for integrated crew pairing and personalized assignment.

Zeighami and Soumis (2017) proposed an integrated model for the crew pairing and personalized assignment problems and a method based on BD and CG. They considered a set of vacation requests (VRs) for each pilot and copilot each month. 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). The monthly schedules for pilots and copilots are generated by the Benders subproblems, which consist of a CG master problem (SMP) and a set of CG subproblems (SSP). The Benders cuts contain pairing variables. The challenge is that the dual values of the unknown pairings from SMP do not exist, and for known pairings there are no arcs corresponding to pairings in the MSP to transfer the information on the Benders cuts from the restricted MMP to the MSP. To overcome this difficulty, they proposed weak Benders cuts that do not take into account the dual variables corresponding to the pairings. They proved that with these weak Benders cuts and the suggested dual solution of the Benders subproblems, the Benders approach reaches optimality. In a more general airline crew scheduling problem, each pilot and copilot can choose a set of preferred flights (PFs) from the monthly schedule. Since the information on PFs is included in the pairings and in the proposed weak Benders cuts of Zeighami and Soumis (2017), the dual value of each pairing is zero, and it is not possible to transfer information on the PFs from the Benders subproblems to the master problem. Thus, Zeighami and Soumis (2017) cannot deal with PFs. The aim of this study is to fill this gap.

We consider the integrated pilot and copilot crew pairing and personalized crew assignment problems in a personalized context with a global objective. Each flight must be covered by one pilot and one copilot, and they request a set of VRs and PFs each month. The pilots and copilots have different schedules to maximize the satisfaction of their preferences. However, to reduce the propagation of perturbations arising during the operation, their pairings must be as similar as possible. We therefore optimize the schedules for the pilots and copilots simultaneously, taking their preferences into account; see Figure 6.1. For simplicity, we use the term *integrated pilot and copilot problem* to refer to the integrated pilot and copilot crew pairing and personalized crew assignment problem.

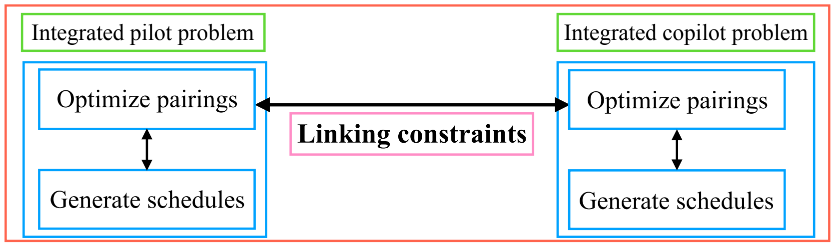

Figure 6.1 Integrated pilot and copilot problems optimized simultaneously within a comprehensively integrated framework

The contributions of this study are as follows: we introduce a novel integrated crew pairing and personalized assignment model to generate personalized monthly schedules for a given set of pilots and copilots simultaneously. To decrease the propagation of perturbations, we keep the pairings in the two problems as similar as possible. To solve this integrated model, we develop an algorithm based on alternating Lagrangian decomposition (ALD), CG, and DCA. The approach uses Lagrangian decomposition in an alternating fashion by proposing a new way to update the Lagrangian multipliers. The solution process iterates between pilots and copilots by estimating the effects of decisions made in one problem on the other. Also, we study the conditions for convergence. The computational results reveal that our approach outperforms the sequential approach in terms of cost savings and the satisfaction of crew preferences. Moreover, the computational time required is similar.

The remainder of this paper is organized as follows. Section 6.2 gives the problem statement, and Section 6.3 gives the mathematical formulation. The algorithm is outlined in Section 6.4, and Section 6.5 gives our experimental results and observations. Finally, Section 6.6 provides concluding remarks.

6.2 Problem Description: The Crew Pairing and Personalized Assignment Problems

In this section, we give the statement of first the crew pairing problem and then the personalized crew assignment problem. To define a pairing and a schedule we use the subset of feasibility rules given by Zeighami and Soumis (2017).

Each flight is defined by a flight number, departure/arrival airports, and fixed departure/arrival dates and times. Consider a set of flights and crew bases that must be operated by the same aircraft fleet over a planning horizon (usually one month). The crew pairing problem finds a set of minimum-cost pairings such that each flight is covered by exactly one pairing. 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 connection periods. A deadhead is a flight where the crew travels as passengers to be relocated. A pairing is feasible if it satisfies the following rules. 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 4 to 8 hours in a duty, a minimum of 4 hours is paid even if they are not worked, 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 Saddoune et al. (2012).

Pilots and copilots are trained for a specific aircraft fleet, and they are associated with a crew base. All pairings assigned to a pilot or copilot must start and end at his/her base. A base is a large airport where pilots and copilots are stationed (with an equal number of each rank); we assume that the number at each base is fixed. In our test, each pilot and copilot requests a set of PFs and VRs per month. A schedule is a sequence of pairings separated by rest periods. We build monthly schedules that cover all the pairings previously built, and each flight must be covered by one pilot and one copilot. A schedule is feasible if it satisfies the following safety and collective agreement rules. 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 rest time between any two consecutive pairings is 8 hours. A penalty cost is associated with each unsatisfied VR, and a negative cost is associated with each satisfied PF. The objective function maximizes the number of satisfied VRs and PFs.

6.3 Mathematical Formulation

In this section, we present our integrated crew pairing and personalized crew assignment model for a given set of pilots and copilots within a completely integrated framework. The model constructs the pairings and monthly schedules for each pilot and copilot in a single step directly from flights instead of pairings. We keep the pairings in the two problems as similar as possible, while satisfying the feasibility rules impacting the pairings and monthly schedules. We use the following notation. Let F be the set of scheduled flights to be covered. Let the set of pilots be L and the set of copilots O. We denote by V_l the set of VRs for pilot $l \in L$ and by V_o the set of VRs for copilot $o \in O$. Let S_l be the set of feasible schedules for pilot $l \in L$ and S_o the set of feasible schedules for copilot $o \in O$. The binary variable x_s is equal to 1 if schedule $s \in S_l$ is allocated to pilot $l \in L$ and 0 otherwise, and y_s is equal to 1 if schedule $s \in S_o$ is allocated to copilot $o \in O$ and 0 otherwise. The variable r_v is 1 if VR $v \in V_l$ is unsatisfied for pilot $l \in L$ and 0 otherwise, and z_v is 1 if VR $v \in V_o$ is unsatisfied for copilot $o \in O$ and 0 otherwise. Let c_v be the penalty cost for unsatisfied VR $v \in V_l$ (or $v \in V_o$). The binary constant e^s_f is equal to 1 if flight $f \in F$ is covered by schedule $s \in S_l$ (or $s \in S_o$). The binary constant b^s_v is equal to 1 if VR $v \in V_l$ (or $v \in V_o$) is satisfied by schedule $s \in S_l$ (or $s \in S_o$). Let n_s be the number of PFs in schedule $s \in S_l$ (or $s \in S_o$), and let B be the bonus cost (a negative cost) for covering each PF. We denote by c_p the cost of pairing p. Let c_s be the cost of schedule $s \in S_l$ (or $s \in S_o$), where $c_s = \sum_{p \in P_s} c_p + B n_s$, and P_s is the set of all pairings covered by schedule s. Let A be all the feasible connection arcs between two successive flights. For every connection arc $(i, j) \in A$ and every schedule s, we set the binary constant a^s_{ij} to 1 if (i, j) between flights i and j is covered by s. The variable u_{ij} is 1 if $(i, j) \in A$ is covered by the pilot problem or copilot problem but not both, and 0 otherwise. Let R_{ij} be the penalty cost for $(i, j) \in A$ if u_{ij} is equal to 1. The value of the penalty cost R_{ij} is inversely proportional to the connection time between flights i and j. This forces the model to include these flights in the same pairing in both the pilot and copilot problems where possible.

Using this notation, the integrated model is as follows:

$$\mathbb{Z} = \min \sum_{l \in L} \sum_{s \in S_l} c_s x_s + \sum_{l \in L} \sum_{v \in V_l} c_v r_v + \sum_{o \in O} \sum_{s \in S_o} c_s y_s + \sum_{o \in O} \sum_{v \in V_o} c_v z_v + \sum_{(i,j) \in A} R_{ij} u_{ij} \tag{6.1}$$

Pilot constraints:

$$\sum_{l \in L} \sum_{s \in S_l} e^s_f x_s = 1 \qquad \forall f \in F \tag{6.2}$$

$$\sum_{s \in S_l} b^s_v x_s + r_v = 1 \qquad \forall l \in L, \forall v \in V_l \tag{6.3}$$

$$\sum_{s \in S_l} x_s \leq 1 \qquad \forall l \in L \tag{6.4}$$

$$\text{Linking constraints:} \quad \sum_{l\in L}\sum_{s\in S_l} a_{ij}^s x_s - \sum_{o\in O}\sum_{s\in S_o} a_{ij}^s y_s \le u_{ij} \qquad \forall (i,j)\in A \tag{6.5}$$

$$\sum_{o\in O}\sum_{s\in S_o} a_{ij}^s y_s - \sum_{l\in L}\sum_{s\in S_l} a_{ij}^s x_s \le u_{ij} \qquad \forall (i,j)\in A \tag{6.6}$$

$$\text{Copilot constraints:} \quad \sum_{o\in O}\sum_{s\in S_o} e_f^s y_s = 1 \qquad \forall f\in F \tag{6.7}$$

$$\sum_{s\in S_o} b_v^s y_s + z_v = 1 \qquad \forall o\in O, \forall v\in V_o \tag{6.8}$$

$$\sum_{s\in S_o} y_s \le 1 \qquad \forall o\in O \tag{6.9}$$

$$x_s \in \{0,1\} \qquad \forall l\in L, \forall s\in S_l \tag{6.10}$$

$$r_v \in \{0,1\} \qquad \forall l\in L, \forall v\in V_l \tag{6.11}$$

$$u_{ij} \in \{0,1\} \qquad \forall (i,j)\in A \tag{6.12}$$

$$y_s \in \{0,1\} \qquad \forall o\in O, \forall s\in S_o \tag{6.13}$$

$$z_v \in \{0,1\} \qquad \forall o\in O, \forall v\in V_o \tag{6.14}$$

The objective function (6.1) finds a trade-off between maximizing the number of satisfied VRs and PFs and minimizing the total cost of the pairings and the dissimilarity of the pilot and copilot pairings. Set-partitioning constraints (6.2) and (6.7) ensure that each flight is covered by exactly one pilot and one copilot. Constraints (6.3) and (6.8) impose the VRs for each pilot and copilot respectively. Constraints (6.4) and (6.9) guarantee that at most one schedule is chosen for each pilot and copilot respectively. Linking constraints (6.5) force variable u_{ij} to be 1 if $(i,j)\in A$ is covered only in the pilot problem. Linking constraint (6.6) forces variable u_{ij} to be 1 if $(i,j)\in A$ is covered only in the copilot problem. The integrality conditions are defined by constraints (6.10)–(6.14).

Remark 10 *The presence of differently covered connection arcs in the two problems leads to different pairings.*

6.4 Algorithm

In this section, we describe our algorithm. The integrated model (6.1)–(6.14) contains a large number of linking constraints; see Table 6.3. To handle these constraints, we propose a solution approach based on ALD. The algorithm iterates between two Lagrangian subproblems, the first for the pilots and the second for the copilots. In both subproblems, the number of variables grows exponentially as the number of flights increases, so we use a CG method. However, because of degeneracy, CG becomes inefficient when the number of set partitioning constraints is large and the columns are dense (more than 8–12 nonzero elements per column; see Elhallaoui et al. (2005)). In our problem, there are between 30 and 45 nonzeros per column. To overcome this difficulty and accelerate the solution process, we combine CG with DCA. In the following subsections, we explain our approach and then describe how to apply Lagrangian decomposition to model (6.1)–(6.14). Finally, we describe the integrated approach.

6.4.1 Column Generation

CG (Barnhart et al., 1998) is an iterative method for large problems with many variables. CG decomposes the main problem into a restricted master problem (RMP) and several pricing problems. The RMP is a restriction of the master problem and contains a subset of the columns. New variables for the master problem are then generated at each iteration by solving pricing problems, with the goal of improving the RMP objective function. At each CG iteration, we solve the RMP using an LP solver to produce primal and dual solutions. Based on the dual solution, we solve the pricing problems to find negative-reduced-cost variables. We then add these variables to the current RMP. We iterate until no negative-reduced-cost variables are identified: the current RMP primal solution is then optimal for the master problem. In our case, each pricing problem corresponds to a resource-constrained shortest path problem that is solved using a label-setting algorithm (Desrochers and Soumis, 1988).

Pricing Problems

There is one pricing problem for each pilot and copilot, defined on a directed acyclic time-space network. The network has six node types: *source, sink, departure, arrival, midnight*, and *waiting*. The source and sink nodes represent the start and end of the schedules. Each flight is defined by departure and arrival nodes, and there is a waiting node for each departing node. The midnight nodes are used to define the start and end of each day. The network has twelve arc types: *start of schedule, end of schedule, flight, deadhead, VR, rest, wait time, start of duty, start of pairing, day off, post-pairing, post-pairing rest*. The start-of-schedule arc connects the source node to the first midnight node of the horizon. Each flight arc and deadhead arc is defined by a departure/arrival node and fixed departure/arrival dates and times. 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. 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. An empty arc links each waiting node to its corresponding departure node. 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. A VR arc links two midnight nodes covering between one and ten days. To ensure schedule feasibility, eight resources are used: maximum pairing duration, maximum number of duties in a pairing, maximum number of landings per duty, maximum working time per duty, maximum total duty duration, minimum number of days off, maximum number of consecutive working days, and maximum credited flying time.

Labeling Algorithm

The pricing problems are solved using a labeling algorithm. Every feasible schedule corresponds to a path from the source node to the sink node in the above network. The feasibility constraints on the schedules are enforced via resource constraints in the network. For each feasibility constraint, there is a resource window at each node, and the value of every constrained resource must be within this interval. Resources are consumed on arcs of the network. The cost of a feasible path corresponding to a schedule from the source node to the sink node is the sum of the arc costs on the path. The arc costs are updated during the solution process based on the dual variables for the master problem constraints. A label is associated with each partial path starting from the source node. Each label has a set of resource components and a cost component. The resource components determine the value of the resources at the last node of the partial path, and the cost component calculates the reduced cost of this path. To reduce the time for the path generation, we eliminate labels using a dominance rule: label L_1 is dominated by label L_2 if the resource components and cost component of L_1 are less than or equal to the corresponding components of L_2. We use a heuristic and an exact version of the labeling algorithm. The heuristic version considers a subset of the resource components with the reduced-cost component in the dominance rule. If no negative-reduced-cost paths are found, the exact version uses all the resource components.

6.4.2 Dynamic Constraint Aggregation

DCA (Elhallaoui et al., 2005) is a new version of CG that reduces the number of set-partitioning constraints in the RMP by aggregating some of them. The approach aggregates clusters of the RMP set-partitioning constraints and retains one representative constraint for each cluster. DCA relies on an aggregated restricted master problem (ARMP) that considers smaller subsets of variables and constraints compared to traditional RMP. This new version speeds up the solution process. At the beginning of DCA, a partition is defined with respect to an initial set of clusters that is provided by a planned solution. During the CG solution process, this aggregation can change dynamically. A column is said to be compatible with the current partition if, for each cluster of the partition, it covers either all of the cluster's tasks or none. Otherwise, this column is incompatible. A newly generated column can be added to ARMP if it is compatible with the current partition. An incompatible column can be added to the ARMP only if the partition is modified. The ARMP is solved using the simplex algorithm to produce a primal solution and a dual solution for the aggregated constraints. To compute the dual values for each set partitioning constraint of the original problem, we use a dual variable disaggregation procedure. A modified version of DCA was developed by Elhallaoui et al. (2010) with multiple phases (MPDCA). MPDCA defines an incompatibility number (r) with respect to the current partition for each column. It estimates the minimum number of additional clusters needed to make an incompatible column compatible. In phase k, only the incompatible variables with $r \leq k$ are priced out. At the beginning of the solution process, we set the phase number to 0. We solve the ARMP and calculate the optimal primal and disaggregated dual solutions and the incompatibility number for each column. All variables with 0-incompatibilities for $0 \leq k$ are priced out by solving the pricing problems. When there is no negative-reduced-cost column, the algorithm continues to phase $(k + 1)$ or stops if k is the last phase. If negative-reduced-cost columns are found, we determine whether or not the current partition must be modified. If no change is required, we add some or all of the compatible columns to the ARMP, and the MPDCA moves to its next iteration. Otherwise, we perform a test to determine whether or not the current partition should be changed. We update the partition if the reduced cost of the least-reduced-cost compatible column is greater than or equal to the reduced cost of the least-reduced-cost incompatible column times a predetermined multiplier.

6.4.3 Alternating Lagrangian Decomposition

To reduce the complexity of the solution process, model (6.1)–(6.14) can be decomposed into two subproblems via Lagrangian decomposition (Guignard and Kim, 1987) and solved by combining CG and DCA. The two subproblems are called ALD-pilot and ALD-copilot. The classical Lagrangian decomposition using subgradient or bundle methods usually needs many iterations to reach an optimal solution (Crainic et al., 2001). We use Lagrangian decomposition in an alternating fashion by proposing a new way to update the Lagrangian multipliers. The algorithm iterates between ALD-pilot and ALD-copilot by estimating the effect of decisions made in one problem on the other. Without loss of generality, we first solve the integrated pilot model (6.2)–(6.4) considering the first two terms of function (6.1) as the objective function; let $\bar{x}_s$ $(s \in S_l, l \in L)$ be the optimal solution. Let $A^1_L = \{(i,j) \in A | \sum_{l \in L} \sum_{s \in S_l} a^s_{ij} \bar{x}_s = 1\}$ and $A^0_L = \{(i,j) \in A | \sum_{l \in L} \sum_{s \in S_l} a^s_{ij} \bar{x}_s = 0\}$ be the set of connection arcs that are covered and uncovered, respectively. The resulting linking constraints (6.5)–(6.6) are

$$1 - \sum_{o \in O} \sum_{s \in S_o} a^s_{ij} y_s \leq u_{ij} \qquad \forall (i,j) \in A^1_L \subseteq A \quad (\text{multipliers } (\lambda_{ij})_{pilot} \geq 0) \tag{6.15}$$

$$\sum_{o \in O} \sum_{s \in S_o} a^s_{ij} y_s - 1 \leq u_{ij} \qquad \forall (i,j) \in A^1_L \subseteq A \tag{6.16}$$

$$0 - \sum_{o \in O} \sum_{s \in S_o} a^s_{ij} y_s \leq u_{ij} \qquad \forall (i,j) \in A^0_L \subseteq A \tag{6.17}$$

$$\sum_{o \in O} \sum_{s \in S_o} a^s_{ij} y_s - 0 \leq u_{ij} \qquad \forall (i,j) \in A^0_L \subseteq A \quad (\text{multipliers } (\pi_{ij})_{pilot} \geq 0) \tag{6.18}$$

Lemma 1 *Constraints* (6.16) *and* (6.17) *are redundant for the integrated copilot problem.*

Proof. Clearly $0 \leq \sum_{o \in O} \sum_{s \in S_o} a^s_{ij} y_s \leq 1$, so $-1 \leq \sum_{o \in O} \sum_{s \in S_o} a^s_{ij} y_s - 1 \leq 0$ and $-1 \leq -\sum_{o \in O} \sum_{s \in S_o} a^s_{ij} y_s \leq 0$. Since $u_{ij} \geq 0$, constraints (6.16) and (6.17) are redundant. □

Let $\boldsymbol{(\lambda)_{Pilot}} = \{(\lambda_{ij})_{pilot} \geq 0 | (i,j) \in A^1_L\}$ and $\boldsymbol{(\pi)_{Pilot}} = \{(\pi_{ij})_{pilot} \geq 0 | (i,j) \in A^0_L\}$ be Lagrangian multipliers associated with constraints (6.15) and (6.18), respectively, and calculated from the pilot problem as explained in Section 6.4.4. The Lagrangian decomposition is obtained by dualizing these linking constraints. The resulting copilot subproblem is

$$\text{ALD-copilot } (\lambda, \pi)_{Pilot} = \min \sum_{o \in O} \sum_{s \in S_o} c_s y_s + \sum_{o \in O} \sum_{v \in V_o} c_v z_v + \sum_{(i,j) \in A} R_{ij} u_{ij} +$$

$$\sum_{(i,j) \in A^1_L} (\lambda_{ij})_{pilot} (1 - \sum_{o \in O} \sum_{s \in S_o} a^s_{ij} y_s - u_{ij}) + \sum_{(i,j) \in A^0_L} (\pi_{ij})_{pilot} (\sum_{o \in O} \sum_{s \in S_o} a^s_{ij} y_s - 0 - u_{ij}) \tag{6.19}$$

subject to (6.7)–(6.9) and (6.12)–(6.14).

After we solve ALD-copilot and fix variable y_s in model (6.1)–(6.14), the resulting linking constraints (6.5)–(6.6) are

$$\sum_{o\in O}\sum_{s\in S_o} a^s_{ij}x_s - 1 \le u_{ij} \qquad \forall (i,j)\in A^1_O \subseteq A \tag{6.20}$$

$$1-\sum_{o\in O}\sum_{s\in S_o} a^s_{ij}x_s \le u_{ij} \qquad \forall (i,j)\in A^1_O \subseteq A \quad \text{(multipliers } (\lambda_{ij})_{copilot}\ge 0) \tag{6.21}$$

$$\sum_{o\in O}\sum_{s\in S_o} a^s_{ij}x_s - 0 \le u_{ij} \qquad \forall (i,j)\in A^0_O \subseteq A \quad \text{(multipliers } (\pi_{ij})_{copilot}\ge 0) \tag{6.22}$$

$$0-\sum_{o\in O}\sum_{s\in S_o} a^s_{ij}x_s \le u_{ij} \qquad \forall (i,j)\in A^0_O \subseteq A \tag{6.23}$$

Here $A^1_O = \{(i,j)\in A | \sum_{o\in O}\sum_{s\in S_o} a^s_{ij}\bar{y}_s = 1\}$ and $A^0_O = \{(i,j)\in A | \sum_{o\in O}\sum_{s\in S_o} a^s_{ij}\bar{y}_s = 0\}$ are the sets of connection arcs that are covered and uncovered by ALD-copilot respectively, and $\bar{y}_s$ $(s\in S_o; o\in O)$ is the optimal solution.

Lemma 2 *Constraints* (6.20) *and* (6.23) *are redundant for the integrated pilot problem.*

The proof is similar to that for Lemma 1.

Now let $\boldsymbol{(\lambda)_{Copilot}} = \{(\lambda_{ij})_{copilot}\ge 0 | (i,j)\in A^1_O\}$ and $\boldsymbol{(\pi)_{Copilot}} = \{(\pi_{ij})_{copilot}\ge 0 | (i,j)\in A^0_O\}$ be Lagrangian multipliers associated with constraints (6.21) and (6.22), respectively, and calculated from ALD-copilot. The resulting pilot subproblem is

$$\text{ALD-pilot } (\lambda,\pi)_{Copilot} = \min \sum_{l\in L}\sum_{s\in S_l} c_s x_s + \sum_{l\in L}\sum_{v\in V_l} c_v r_v + \sum_{(i,j)\in A} R_{ij}u_{ij} +$$
$$\sum_{(i,j)\in A^1_O} (\lambda_{ij})_{copilot}(1-\sum_{l\in L}\sum_{s\in S_l} a^s_{ij}x_s - u_{ij}) + \sum_{(i,j)\in A^0_O} (\pi_{ij})_{copilot}(\sum_{l\in L}\sum_{s\in S_l} a^s_{ij}x_s - 0 - u_{ij}) \tag{6.24}$$

subject to (6.2)–(6.4) and (6.10)–(6.12).

6.4.4 Lagrangian Multipliers

To compute the Lagrangian multipliers, we introduce the concept of an arc reduced cost to benefit from information connecting the two problems. In a traditional CG, the arc reduced costs are not directly available, and we require additional information from the pricing problem to compute them. Consider two consecutive flights i and j. We calculate the arc reduced cost on the arc (i,j) between these two flights using the following formula:

$$\bar{c}_{ij} = \Pi_j - \Pi_i. \tag{6.25}$$

Here Π_i and Π_j are the minimum costs from the source node to the arrival node of flight i and the departure node of flight j, respectively. Π_i and Π_j are calculated using the labeling algorithm described in Section 6.4.1. Consider the example shown in Figure 6.2 (for clarity, some arcs and labels are truncated or omitted). Assume that we have three pilots: A, B, and C. Without loss of generality, we assume that in the optimal solution of ALD-pilot, flights i, j, and k are covered by pilot A; flights m, n, o, and p are covered by pilot B; and flights q, s, and r are covered by pilot C. Let Π_i^A and Π_j^A respectively be the minimum cost from the source node to the arrival node of flight i and the departure node of flight j, obtained from the pricing problem corresponding to pilot A. The arc reduced cost on (i, j) is then given by $\bar{\Pi}_j^A - \bar{\Pi}_i^A$; see Figure 6.3.

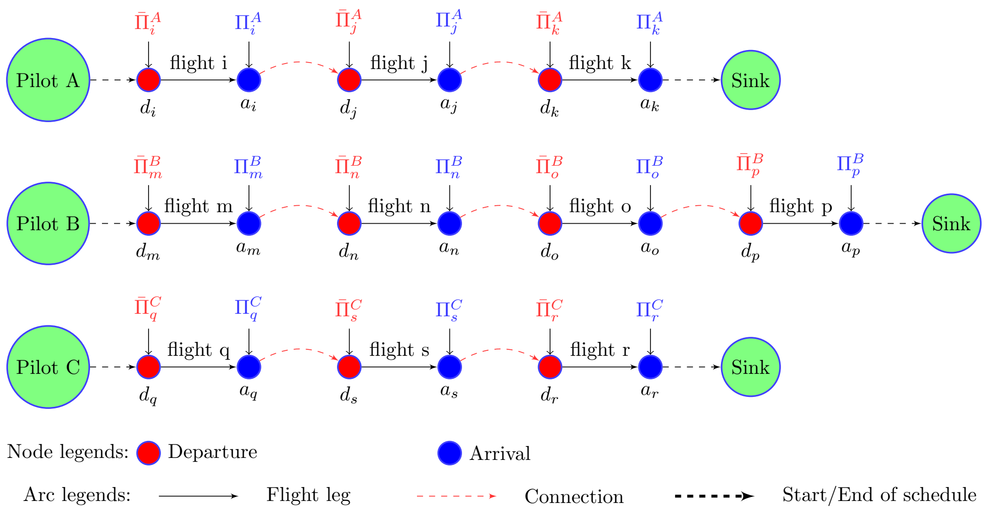

Figure 6.2 Finding minimum cost from source node to each departure and arrival node in pilot pricing problems.

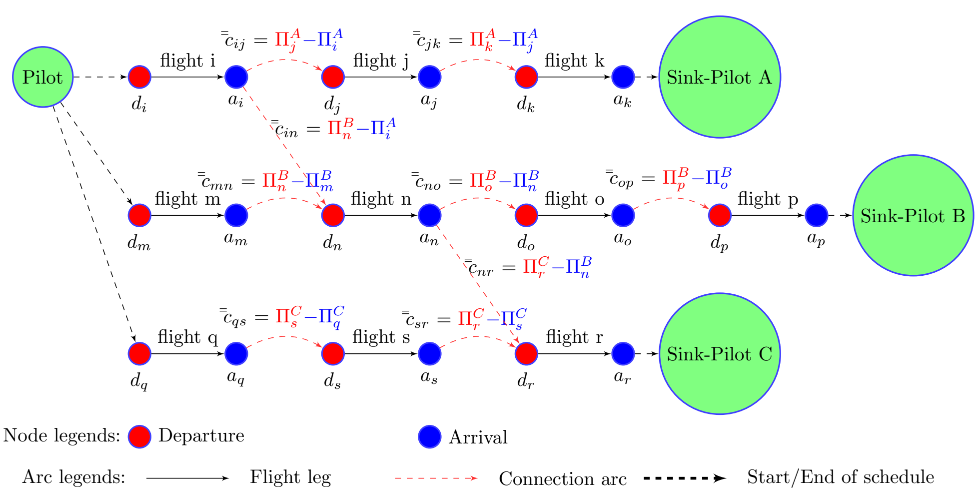

Figure 6.3 Calculate arc reduced cost on each connection arc in pilot pricing problems.

Remark 11 *There is one pricing problem for each pilot (or copilot), so arc $(i,j) \in A$ may exist in more than one pricing problem. Therefore, there is more than one way to calculate the arc reduced cost on $(i,j) \in A$. We use the pilot (copilot) who covered the first flight and the pilot (copilot) who covered the second flight related to this arc in the optimal solution.*

Let $(\bar{c}_{ij})^k_{pilot}$ be the arc reduced cost of $(i,j) \in A$ obtained by ALD-pilot at iteration k. We pay a minimum penalty cost to ALD-copilot to have pairings similar to those obtained by ALD-pilot. For each $(i,j) \in A$, we calculate the Lagrangian multipliers as the minimum of the arc reduced cost $(\bar{c}_{ij})^k_{pilot}$ and the penalty cost R_{ij}. The Lagrangian multipliers for ALD-pilot (ALD-copilot) at iteration k are computed as follows:

$$(\lambda_{ij})^k_{pilot} = min\{max\{(\bar{c}_{ij})^k_{pilot}, 0\}, R_{ij}\} \qquad \forall (i,j) \in A^1_L \tag{6.26}$$

$$(\pi_{ij})^k_{pilot} = min\{max\{(\bar{c}_{ij})^k_{pilot}, 0\}, R_{ij}\} \qquad \forall (i,j) \in A^0_L \tag{6.27}$$

The Lagrangian multipliers can be considered dual information. Before we solve ALD-copilot, we must project the calculated multipliers on each connection arc in ALD-pilot onto the corresponding arcs in the copilot's pricing problems. We then solve ALD-copilot using CG combined with DCA. Since $(\lambda_{ij})_{pilot}$ is negated in the arc covering term of ALD-copilot (i.e., $(\lambda_{ij})_{pilot}(1 - \sum_{o \in O} \sum_{s \in S_o} a^s_{ij} y_s - u_{ij})$) we add the value $-(\lambda_{ij})_{pilot}$ on the arc (i,j) in all the ALD-copilot's pricing problems that include this arc. On the other hand, since $(\pi_{ij})_{pilot}$ is positive in the arc covering term of ALD-copilot (i.e., $(\pi_{ij})_{pilot}(\sum_{o \in O} \sum_{s \in S_o} a^s_{ij} y_s - 0 - u_{ij})$) we add the value $+(\pi_{ij})_{pilot}$ on the arc (i,j) in all the ALD-copilot's pricing problems that include this arc; see Figure 6.4. In fact, we add a bonus cost $-(\lambda_{ij})_{pilot}$ to $(i,j) \in A$ in ALD-copilot's pricing problems if the arc is covered by ALD-pilot, and otherwise we add the penalty cost $+(\pi_{ij})_{pilot}$; see Figure 6.4. This encourages ALD-copilot to cover the connection arcs that are covered by ALD-pilot. The same strategy is used for ALD-pilot.

Remark 12 *Let $(\bar{\bar{C}}_{ij})^k_{copilot}$ be the Lagrangian arc reduced cost of ALD-copilot, calculated via*

$$(\bar{\bar{C}}_{ij})^k_{copilot} = \begin{cases} (\bar{c}_{ij})^k_{copilot} - (\lambda_{ij})^k_{pilot}, & \text{if } (i,j) \in A^1_L \\ (\bar{c}_{ij})^k_{copilot} + (\pi_{ij})^k_{pilot}, & \text{if } (i,j) \in A^0_L. \end{cases}$$

Since $(\lambda_{ij})^k_{pilot}$ and $(\pi_{ij})^k_{pilot}$ are known, we can calculate $(\bar{c}_{ij})^k_{copilot}$ at each iteration to update the Lagrangian multipliers. The same is true for ALD-pilot.

We now have all the ingredients to derive an iterative algorithm for model (6.1)–(6.14).

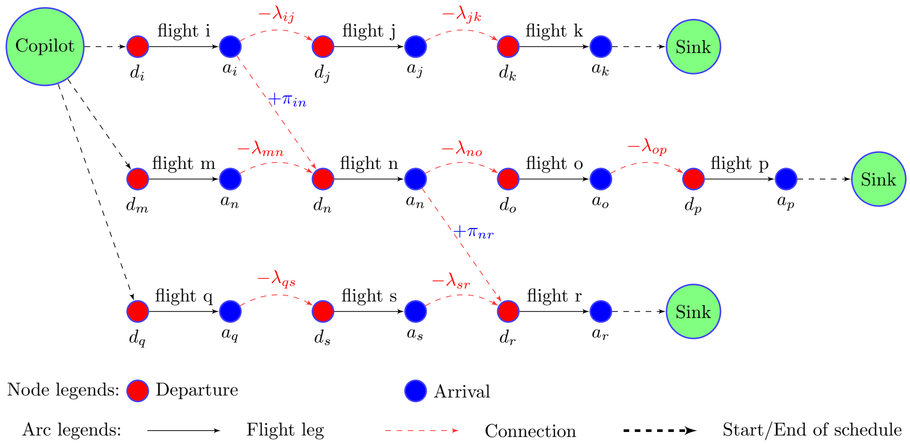

Figure 6.4 Projection of Lagrangian multipliers for pilot onto corresponding arcs in copilot pricing problems.

6.4.5 Integrated Approach

In this section, we describe the integrated approach (INT). INT solves the Lagrangian subproblems alternatively by transferring *primal* and *dual information* between them. The subproblems are solved by CG and DCA. Figure 6.5 illustrates the INT approach; the green line represents the Lagrangian terms in the objective function of ALD-copilot obtained via the Lagrangian multipliers from ALD-pilot.

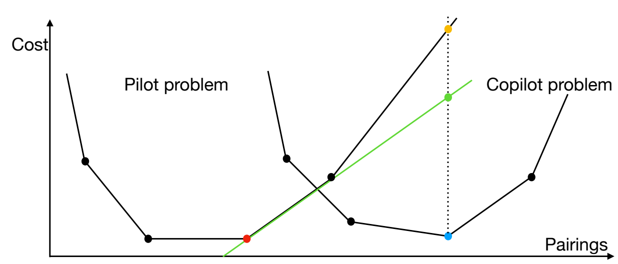

Figure 6.5 Schematic of INT approach.

Primal information: To apply DCA to each subproblem we need to define a set of clusters as an initial partition. We define a cluster as a nonempty subset of consecutive flights. A partition is a set of clusters that covers all the flights; in our case, a cluster is a pairing. At each INT iteration, we use the optimal solution obtained by ALD-pilot as an initial partition for ALD-copilot, and vice versa.

Dual information: At each INT iteration, to estimate the effect of decisions made in ALD-pilot on ALD-copilot and define the Lagrangian terms in the objective function, ALD-copilot calculates the multipliers $(\lambda, \pi)_{pilot}$ and information about the coverage of each $(i, j) \in A$ (A_l^1, A_l^0) from ALD-pilot. ALD-pilot must do likewise.

Figure 6.6 gives a flowchart for INT. INT starts with the set of initial pairings constructed by Zeighami and Soumis (2017) as an initial partition. The initial values of the multipliers are 0. The algorithm starts by solving ALD-pilot using CG and DCA and constructs personalized monthly schedules for the pilots. It then solves ALD-copilot using CG and DCA and taking into account the primal and dual information from ALD-pilot, and it constructs personalized monthly schedules for the copilots. It alternates between the two subproblems until the relative difference between the lower and upper bounds is less than or equal to a predetermined threshold.

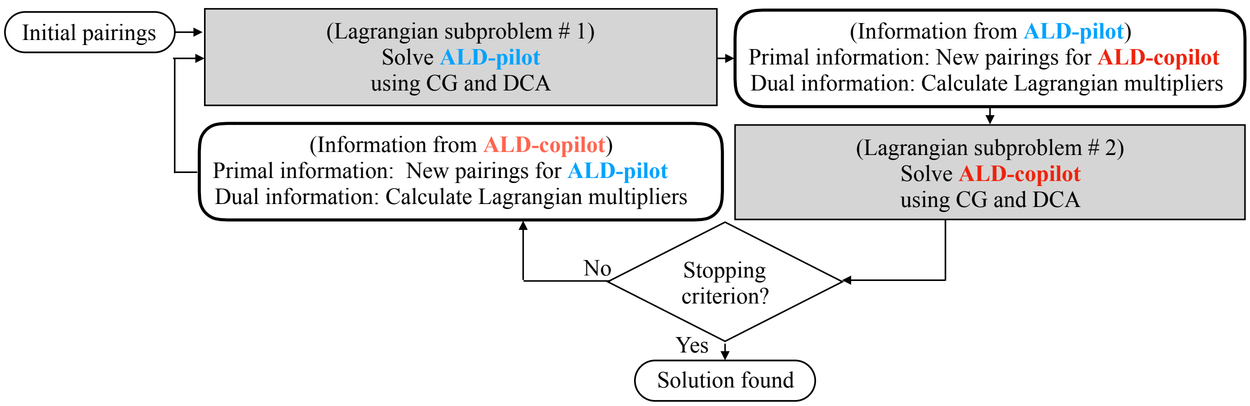

Figure 6.6 Flowchart of INT approach.

6.4.6 Lower and Upper Bounds

Computing lower and upper bounds allows us to assess the quality of the solutions generated by the INT approach. The upper bounds can be computed at each iteration by using the solutions of ALD-pilot $(\bar{x}_s, \bar{r}_v)$ and ALD-copilot $(\bar{y}_s, \bar{z}_v)$ that are calculated during INT. Clearly, $(\bar{x}_s, \bar{r}_v)$ satisfies constraints (6.2)–(6.4) and $(\bar{y}_s, \bar{z}_v)$ satisfies constraints (6.7)–(6.9). By substituting $\bar{x}_s$ and $\bar{y}_s$ into linking constraints (6.5) and (6.6), we obtain a value of variable u_{ij} that is feasible for model (6.1)–(6.14). Using this feasible solution, i.e., $(\bar{x}_s, \bar{r}_v, \bar{y}_s, \bar{z}_v, \bar{u}_{ij})$, we compute an upper bound. To compute lower bounds we need the following definition.

Definition 2 *Let $C(y_s, z_v)$ be the cost function of ALD-copilot and $P(x_s^*, r_v^*)$ the optimal cost of ALD-pilot considering the first two terms of objective functions* (6.19) *and* (6.24)*,* *respectively. Let* $L(y_s)_{(\lambda,\pi)_{pilot}}$ *be the cost function of the Lagrangian term obtained by solving ALD-copilot with respect to the Lagrangian multipliers obtained from ALD-pilot. Let* $LB(y_s, z_v)_{(\lambda,\pi)_{pilot}} = P(x_s^*, r_v^*) + L(y_s)_{(\lambda,\pi)_{pilot}} + C(y_s, z_v)$.

Proposition 5 $LB(y_s, z_v)_{(\lambda,\pi)_{pilot}}$ *is a lower bound on the optimal objective function value of model* (6.1)–(6.14).

Proof. $C(y_s, z_v)$ is the cost of the integrated copilot problem, and $P(x_s^*, r_v^*) + L(y_s)_{(\lambda,\pi)_{pilot}}$ (the green line in Figure 6.5) is a lower bound on the cost of the integrated pilot problem and the penalty cost of the dissimilarity of pairings between the two problems. Therefore, $LB(y_s, z_v)_{(\lambda,\pi)_{pilot}}$ is a lower bound on the optimal objective function value of model (6.1)–(6.14); see Figure 6.5. □

Definition 3 *Let* $P(x_s, r_v)$ *be the cost function of ALD-pilot and* $C(y_s^*, z_v^*)$ *the optimal cost of ALD-copilot considering the first two terms of objective functions* (6.24) *and* (6.19)*, respectively. Let* $L(x_s)_{(\lambda,\pi)_{copilot}}$ *be the cost function of the Lagrangian term obtained by solving ALD-pilot with respect to the Lagrangian multipliers obtained from ALD-copilot. Let* $LB(x_s, r_v)_{(\lambda,\pi)_{copilot}} = P(x_s, r_v) + L(x_s)_{(\lambda,\pi)_{copilot}} + C(y_s^*, z_v^*)$.

Proposition 6 $LB(x_s, r_v)_{(\lambda,\pi)_{copilot}}$ *is a lower bound on the optimal objective function value of model* (6.1)–(6.14).

Proof. See the proof of Proposition 6.1. □

In the following two propositions, we show that INT converges to an optimal solution of model (6.1)–(6.14) if the similarity of the pairings between the two problems is 100%.

Proposition 7 *The following model is equivalent to the ALD-copilot model (6.19):*

$$min \sum_{o \in O} \sum_{s \in S_o} c_s y_s + \sum_{o \in O} \sum_{v \in V_o} c_v z_v + \sum_{(i,j) \in A_L^1} (\lambda_{ij})_{pilot} (1 - \sum_{o \in O} \sum_{s \in S_o} a_{ij}^s y_s) + \sum_{(i,j) \in A_L^0} (\pi_{ij})_{pilot} (\sum_{o \in O} \sum_{s \in S_o} a_{ij}^s y_s - 0)$$

$$\textit{subject to } (6.7)\text{–}(6.9) \textit{ and } (6.12)\text{–}(6.14). \tag{6.28}$$

Proof. Since $A = A_L^1 \cup A_L^0$, $\sum_{(i,j) \in A_L} R_{ij} u_{ij} = \sum_{(i,j) \in A_L^1} R_{ij} u_{ij} + \sum_{(i,j) \in A_L^0} R_{ij} u_{ij}$. Also,

$$\sum_{(i,j) \in A_L^1} (\lambda_{ij})_{pilot} (1 - \sum_{o \in O} \sum_{s \in S_o} a_{ij}^s y_s - u_{ij}) = \sum_{(i,j) \in A_L^1} (\lambda_{ij})_{pilot} (1 - \sum_{o \in O} \sum_{s \in S_o} a_{ij}^s y_s) - \sum_{(i,j) \in A_L^1} (\lambda_{ij})_{pilot} u_{ij} \tag{6.29}$$

and

$$\sum_{(i,j)\in A_L^0}(\pi_{ij})_{pilot}(\sum_{o\in O}\sum_{s\in S_o}a_{ij}^s y_s - 0 - u_{ij}) = \sum_{(i,j)\in A_L^0}(\pi_{ij})_{pilot}(\sum_{o\in O}\sum_{s\in S_o}a_{ij}^s y_s) - \sum_{(i,j)\in A_L^0}(\pi_{ij})_{pilot}u_{ij}. \tag{6.30}$$

By (6.29) and (6.30), the following model is equivalent to the ALD-copilot model (6.19):

$$\begin{aligned} \min \sum_{o\in O}\sum_{s\in S_o} c_s y_s + \sum_{o\in O}\sum_{v\in V_o} c_v z_v + \sum_{(i,j)\in A_L^1}(R_{ij} - (\lambda_{ij})_{pilot})u_{ij} + \sum_{(i,j)\in A_L^0}(R_{ij} - (\pi_{ij})_{pilot})u_{ij} + \\ \sum_{(i,j)\in A_L^1}(\lambda_{ij})_{pilot}(1 - \sum_{o\in O}\sum_{s\in S_o}a_{ij}^s y_s) + \sum_{(i,j)\in A_L^0}(\pi_{ij})_{pilot}(\sum_{o\in O}\sum_{s\in S_o}a_{ij}^s y_s - 0) \\ \text{subject to (6.7)–(6.9) and (6.12)–(6.14).} \end{aligned} \tag{6.31}$$

From equations (6.26) and (6.27) we have $R_{ij} \geq (\lambda_{ij})_{pilot}$ and $R_{ij} \geq (\pi_{ij})_{pilot}$, so $R_{ij} - (\lambda_{ij})_{pilot} \geq 0$ and $R_{ij} - (\pi_{ij})_{pilot} \geq 0$. Model (6.31) is a minimization problem, so in the optimal solution $\sum_{(i,j)\in A_L^1}(R_{ij} - (\lambda_{ij})_{pilot})u_{ij} = \sum_{(i,j)\in A_L^0}(R_{ij} - (\pi_{ij})_{pilot})u_{ij} = 0$. This completes the proof. □

Remark 13 *We have an equivalent result for the ALD-pilot model* (6.24)*.*

Proposition 8 *The solution calculated by INT is an optimal solution for model* (6.1)–(6.14) *if the similarity of the pairings between the two problems is 100%.*

Proof. It suffices to show that the cost of the lower bound obtained by INT is equal to the cost of the upper bound. Let $(\bar{x}_s, \bar{r}_v, \bar{y}_s, \bar{z}_v)$ be an INT solution in which the similarity of the pairings between the two problems is 100%. Hence, $\sum_{(i,j)\in A_L^1}(\lambda_{ij})_{pilot}(1-\sum_{o\in O}\sum_{s\in S_o}a_{ij}^s\bar{y}_s) = 0$ and $\sum_{(i,j)\in A_L^0}(\pi_{ij})_{pilot}(\sum_{o\in O}\sum_{s\in S_o}a_{ij}^s\bar{y}_s - 0) = 0$ (or $\sum_{(i,j)\in A_O^1}(\lambda_{ij})_{copilot}(1-\sum_{l\in L}\sum_{s\in S_l}a_{ij}^s\bar{x}_s) = 0$ and $\sum_{(i,j)\in A_O^0}(\pi_{ij})_{copilot}(\sum_{l\in L}\sum_{s\in S_l}a_{ij}^s\bar{x}_s - 0) = 0$), and from Proposition 7 the cost function of INT at the relevant iteration is

$$ALD = \sum_{l\in L}\sum_{s\in S_l} c_s\bar{x}_s + \sum_{l\in L}\sum_{v\in V_l} c_v\bar{r}_v + \sum_{o\in O}\sum_{s\in S_o} c_s\bar{y}_s + \sum_{o\in O}\sum_{v\in V_o} c_v\bar{z}_v. \tag{6.32}$$

Clearly, $(\bar{x}_s, \bar{r}_v)$ satisfies constraints (6.2)–(6.4) and $(\bar{y}_s, \bar{z}_v)$ satisfies constraints (6.7)–(6.9). Substituting $(\bar{x}_s, \bar{y}_s)$ into the linking constraints (6.5)–(6.6) in model (6.1)–(6.14) gives $u_{ij} = 0$. If we put this solution into (6.1)–(6.14), the cost is

$$\mathbb{Z} = \sum_{l\in L}\sum_{s\in S_l} c_s\bar{x}_s + \sum_{l\in L}\sum_{v\in V_l} c_v\bar{r}_v + \sum_{o\in O}\sum_{s\in S_o} c_s\bar{y}_s + \sum_{o\in O}\sum_{v\in V_o} c_v\bar{z}_v. \tag{6.33}$$

Hence, $ALD = \mathbb{Z}$. □

6.4.7 Integer Solution

To obtain integer solutions, the CG is embedded in a branch-and-bound framework. We use two branching heuristics, column fixing and inter-task fixing. In the column fixing strategy, 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. For each ordered pair of flights f_1 and f_2, let $r_{f_1 f_2} = \sum_{s \in S(f_1,f_2)} x_s$ be the total flow between the pair. Here $S(f_1, f_2)$ is the subset of schedules covering flight f_2 immediately after f_1. We select the maximum fractional values, $r_{f_1 f_2}$. By setting $r_{f_1 f_2} = 1$ we force these two flights to be covered consecutively in the same schedule. The numerical results show that the solutions produced with this heuristic branching are close to optimality; see Table 6.6.

6.5 Computational Experiments

We carried out computational experiments on four instances derived from a one-month flight schedule for a major North American airline. The characteristics of the instances are summarized in Table 6.1. They all contain three crew bases. Each pilot and copilot requests two vacations and a set of PFs each month. The number of pilots and copilots per base is equal and given in Table 6.2. For example, the last instance contains 5613 flights, 49 airports, 145 pilots and 145 copilots, 290 VRs for pilots and 290 VRs for copilots, and 2500 PFs for pilots and 2500 PFs for copilots. The pairing and schedule feasibility are restricted by the rules stated in Section 6.2. To compare the performance of INT with the traditional sequential approach (SEQ), we solve each of the instances of Table 6.1 with both methods. SEQ first solves the crew pairing problem of Zeighami and Soumis (2017) and then the personalized crew assignment problem using CG for a given set of pilots and copilots independently. To define a sequential model for the personalized pilot crew assignment problem, we replace constraint (6.2) with $\sum_{l \in L} \sum_{s \in S_l} e^s_p x_s = 1 \; \forall p \in P$ by considering constraints (6.3)–(6.4) and (6.10)–(6.11) and the first two terms of function (6.1) as the objective function, where P is the set of pairings constructed by the crew pairing problem. The copilot problem is defined similarly. Table 6.3 reports the number of constraints for each instance for SEQ and INT. For example, for SEQ and the fourth instance, the crew pairing model has 5616 constraints, the pilot assignment model has 1514, and the copilot assignment model has 1514. In INT for the same instance, there are 6051 pilot and copilot constraints, 152187 linking constraints, and 164289 constraints in total.

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. In practice, when CG gets close to the optimal value, new columns 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 0.01% over 25 iterations for small problems and 15 for large problems. We chose these parameter values based on the results of preliminary tests. Also, we use the improved version of DCA, i.e., MPDCA, as described in Section 6.4.2. MPDCA is exact when the final phase number k is sufficiently large to ensure the pricing of all feasible columns. Because of the complexity of ALD-copilot and ALD-pilot, we use three phases of DCA ($k = 2$). We also ran a test with $k = 3$, and it gave no significant improvement over $k = 2$.

Table 6.1 Instance Characteristics

	#			No. VRs		No. PFs	
Instance	Flights	Airports	Bases	Pilots	Copilots	Pilots	Copilots
1	1013	26	3	66	66	600	600
2	1500	35	3	68	68	750	750
3	1854	41	3	94	94	850	850
4	5613	49	3	290	290	2500	2500

Table 6.2 Number of Pilots and Copilots per Base

	Base 1		Base 2		Base 3		Total	
Instance	No. Pilots	No. Copilots	No. Pilots	No. Copilots	No. Pilots	No. Copilots	No. Pilots	No. Copilots
1	7	7	20	20	6	6	33	33
2	10	10	9	9	15	15	34	34
3	10	10	30	30	7	7	47	47
4	42	42	78	78	25	25	145	145

Table 6.3 Number of Constraints

	Sequential model			Integrated model			
Instance	Pairing problem	Pilot problem	Copilot problem	Pilot problem	Copilot problem	Linking constraints	Total
1	1016	287	287	1112	1112	6514	8738
2	1503	405	405	1602	1602	9465	12669
3	1857	440	440	1995	1995	16796	20786
4	5616	1514	1514	6051	6051	152187	164289

6.5.1 Solution Process of Integrated Approach

In this experiment, we study the solution process of INT at each iteration. Table 6.4 reports the similarity percentage of the pairings for the two problems, the percentage of satisfied VRs and PFs, the CPU time, and the gap at each iteration for each instance.

Using SEQ to solve these problems produces the same pairings for the pilots and copilots. INT produces the same pairings in all instances except the second. The percentage of dissimilarity for the second instance is only 0.35%. The good performance of INT is in part explained by the fact that for all four instances the similarity percentage increases after each iteration while the percentage of satisfied VRs and PFs slightly decreases.

The reported CPU time is the total effort needed to solve each instance at each iteration. For all the instances, most of the CPU time is spent on the first iteration, and this can be explained as follows. First, choosing a good initial solution has a positive influence on the performance of DCA; it can result in fewer CG iterations and hence a lower CPU time. We warm-start the pilot and copilot problems after the first iteration. After solving each problem we use the current optimal pairings as the initial solution for the other problem; see Section 6.4.3. Second, using good initial columns reduces the CPU time of CG. We use the optimal solution of the current iteration of ALD-pilot as the initial solution for the next iteration of this problem. When adding initial columns to the master problem of CG, we take into account that some of these columns will subsequently be removed by DCA if they are not compatible with the pairings found by ALD-copilot. They can be re-introduced when the initial partition is reconsidered by DCA. The same strategy applies to ALD-copilot. In particular, the computational times are reduced considerably after the first iteration.

The gap corresponds to the relative difference between the values of the lower bound and upper bound obtained by INT for the global objective. The results show that the percentage gaps are between 0.00% and 0.01%, and this justifies the use of INT.

INT reaches a (nearly) optimal solution in three or four iterations for the large and complex problem. For example, in the third instance, the similarity is 91.35% at the first iteration and 100% at the end of the fourth iteration. The percentage of satisfied VRs is unchanged for the pilots and 0.04% lower for the copilots. The percentage of satisfied PFs is 0.15% lower for the pilots and 0.43% lower for the copilots. The global optimality gap is 1.52% at the first iteration and 0% at the last iteration.

Table 6.4 Solution Process of INT

Instance	Iteration	Similarity percentage	Pilots: Satisfied VRs (%)	Pilots: Satisfied PFs (%)	Copilots: Satisfied VRs (%)	Copilots: Satisfied PFs (%)	CPU (min)	Gap (%)
1	WT	82.14	96.96	37.66	98.48	39.00	3.37	-
	1	92.44	96.96	37.66	98.48	38.66	3.55	1.45
	2	97.61	96.96	37.16	98.48	38.50	1.24	0.20
	3	100.00	96.96	37.00	98.48	38.00	1.05	0.00
2	WT	83.49	92.64	43.65	97.05	41.73	12.02	-
	1	89.50	92.64	43.65	97.05	41.50	12.15	1.75
	2	95.22	92.64	43.55	97.05	41.46	4.06	0.45
	3	98.00	92.64	43.33	97.05	41.46	3.52	0.12
	4	99.65	92.46	43.33	97.05	41.06	3.41	0.01
3	WT	81.75	84.00	50.75	82.97	51.17	26.23	-
	1	91.35	84.00	50.75	81.95	50.90	26.40	1.52
	2	97.43	84.00	50.72	81.95	50.88	7.17	0.18
	3	99.27	84.00	50.70	81.95	50.88	7.26	0.05
	4	100.00	84.00	50.70	81.91	50.47	6.55	0.00
4	WT	80.15	98.62	50.12	98.62	50.68	1206.45	-
	1	96.47	98.62	50.00	98.62	50.56	1224.33	0.22
	2	99.96	98.27	49.40	98.62	50.28	175.23	0.03
	3	100.00	98.27	49.40	98.62	50.28	150.38	0.00

6.5.2 Impact of Lagrangian Multipliers on the Similarity of Pairings

To highlight the impact of the Lagrangian multipliers on the similarity percentage, we solve each problem using CG and DCA without transferring these values (labeled WT in Table 6.4). The results clearly show that there is a significant reduction in the similarity percentage when we ignore the Lagrangian multiplier values. For example, for the first iteration of the largest instance, the similarity percentage is 80.15% without the Lagrangian values and 96.47% at the first iteration with them. That is an improvement of 16.32% at the first iteration while the percentage of satisfied PFs and VRs decreases by less than 1%.

6.5.3 Computational Time, Iterations, and Cost for Integrated and Sequential Approaches

In Table 6.5, we compare the total CPU time (in minutes) to solve each of the four instances with the two approaches. For SEQ, the total CPU time consists of the time to solve the pairing problem and the time to solve the pilot and copilot assignment problems. For INT, the total CPU time consists of the time to compute the initial partition, i.e., the time to solve the crew pairing problem of Zeighami and Soumis (2017), and the time to solve ALD-pilot and ALD-copilot. We report the total solution cost obtained with both approaches. This includes the cost of the pairings, bonuses for satisfied PFs, penalties for unsatisfied VRs, and the cost associated with the similarity percentage. The last two columns of Table 6.5 give the ratios of the CPU times of INT and SEQ and the percentage cost savings achieved by the INT solutions. These results show that INT can yield significant cost savings: between 12.76% and 24.20%, with an average of 18.89%. Just a few iterations suffice to find a good solution, and the computational time is close to that for SEQ: 1.65 times longer on average.

Table 6.5 Computational Time and Cost for Integrated and Sequential Approaches

Instance	SEQ approach		INT approach			CPU INT/SEQ	Cost saving by INT (%)
	CPU (min)	Cost	CPU (min)	Cost	Number of iterations		
1	5.47	1551144	11.42	1250844	3	2.08	19.35
2	19.56	1903930	36.28	1660930	4	1.85	12.76
3	61.50	2442300	80.15	2047500	4	1.30	19.28
4	1393.58	3672890	1943.52	2783900	3	1.39	24.20
Average						1.65	18.89

6.5.4 Computational Gap in Integrated and Sequential Approaches

In this experiment, we study the optimality gap of each instance. The gap is the percentage difference between the best lower bound and the best integer solution obtained by CG. To find a lower bound, we stop CG at each branching node if the optimal value of the RMP improves by less than 0.01% over 25 iterations for small problems and 15 for large problems. The lower bound improves during the branch and bound process since we explore many nodes. To obtain integer solutions, we use the branching heuristics discussed in Section 6.4.7. Table 6.6 gives the results of this experiment. For SEQ, three optimality gaps are reported: the crew pairing gap corresponds to the crew pairing problem, and the pilot and copilot gaps correspond to the pilot and copilot assignment problems. For INT, two optimality gaps are reported, for ALD-pilot and ALD-copilot. The percentage gaps are small, indicating that the two approaches produce good solutions. On average the percentage gap for SEQ varies between 0.18 and 0.47, and that for INT varies between 0.03 and 0.10.

Table 6.6 Optimality Gap (%) for Integrated and Sequential Approaches

Instance	SEQ approach			INT approach	
	Pairing problem	Pilot problem	Copilot problem	Pilot problem	Copilot problem
1	0.07	0.10	0.00	0.02	0.00
2	0.86	0.00	0.22	0.06	0.05
3	0.88	0.55	0.21	0.00	0.00
4	0.10	0.38	0.30	0.06	0.35
Average	0.47	0.25	0.18	0.03	0.10

6.5.5 Comparisons of Pairing Costs and Coverage of VRs and PFs

In this experiment, we study the performance of INT and SEQ in terms of pairing costs and the percentage of satisfied VRs and PFs. The objective function (6.1) finds a trade-off between maximizing the number of satisfied VRs and PFs and minimizing the total cost of the pairings and their dissimilarity. Table 6.7 reports the total pairing cost obtained by the two approaches. On average, INT increases the pairing cost by only 0.55% above that of SEQ. This increase might be necessary to increase the number of satisfied VRs and PFs. However, to force the model to stay close to the SEQ pairing cost, we can add the constraints $\sum_{l\in L}\sum_{s\in S_l}(\sum_{p\in P_s} c_p)x_s \leq (1+\epsilon)N$ and $\sum_{o\in O}\sum_{s\in S_o}(\sum_{p\in P_s} c_p)y_s \leq (1+\epsilon)N$ in model (6.1)–(6.14), where N is the pairing cost obtained by SEQ. Tables 6.8 and 6.9 give the results: INT performs significantly better than SEQ. On average, INT increases the coverage of VRs by 43.50% for the pilots and 38.49% for the copilots. On average, it increases the coverage of PFs by 14.85% for the pilots and 12.67% for the copilots. These improvements can be explained by the fact that SEQ does not take into account the crew assignment constraints and objective during the construction of the pairings. The pairings generated by the pairing problem may not be suitable for the objective of the assignment problem, so it is difficult to maximize the satisfaction of VRs and PFs in SEQ.

Table 6.7 Comparisons of Pairing Cost

Instance	SEQ approach	INT approach	Improvement by INT (%)
1	1274922	1280044	-0.40
2	1678746	1688880	-0.60
3	1952742	1967500	-0.75
4	2966890	2980900	-0.47
Average			-0.55

Table 6.8 Comparisons of Satisfied VRs (%)

Instance	SEQ approach		INT approach		Improvement by INT (%)	
	Pilot problem	Copilot problem	Pilot problem	Copilot problem	Pilot problem	Copilot problem
1	45.00	60.00	96.96	98.48	53.00	38.00
2	42.00	52.00	92.46	97.05	50.00	45.00
3	42.00	39.00	84.00	81.91	42.00	42.00
4	68.96	69.65	98.26	98.62	29.30	28.97
Average					43.50	38.49

Table 6.9 Comparisons of Satisfied PFs (%)

Instance	SEQ approach		INT approach		Improvement by INT (%)	
	Pilot problem	Copilot problem	Pilot problem	Copilot problem	Pilot problem	Copilot problem
1	25.83	27.66	37.00	38.00	11.00	11.00
2	28.00	30.00	43.33	41.06	15.00	11.00
3	32.00	33.52	50.70	50.47	20.00	16.00
4	36.00	37.60	49.40	50.28	13.40	12.68
Average					14.85	12.67

Remark 7 *A (co)pilot may request flights that have conflicts in the departure or arrival times. Therefore, it is difficult to satisfy a high percentage of the PFs.*

6.6 Conclusion

We have introduced an integrated crew pairing and personalized crew assignment model to generate personalized monthly schedules for a given set of pilots and copilots simultaneously. Each pilot and copilot requests a set of preferred flights and vacations each month. Our model keeps the pairings in the two problems as similar as possible to reduce the propagation of perturbations. We have presented an integrated approach based on ALD, CG, and DCA. This novel approach uses Lagrangian decomposition in an alternating fashion. It introduces a new way to update the Lagrangian multipliers by estimating the effect of decisions made in one problem on the other. We proved that our approach finds an optimal solution when the sets of pairings are identical. We studied real-world instances from a major US carrier, and the results show that a few iterations suffice to find a (nearly) optimal solution for the large and complex problem. Also, the integrated approach yields significant improvements over the sequential approach in terms of satisfied VRs and PFs.

Table 6.10 Notation for chapters 4 et 5

Sets:	
B	Set of all bases.
P_b	Set of feasible pairings in base $b \in B$.
F	Set of scheduled flights to be covered.
F_p	Set of flights contained in pairing p.
J	$j = 1$ refers to the pilot problem and $j = 2$ refers to the copilot problem.
L_b^j	Set of pilots (for $j = 1$) and copilots (for $j = 2$) at base $b \in B$.
V_l	Set of VRs for (co)pilot $l \in L_b^j$ at base $b \in B$.
A_b	Sets of aggregated arrival nodes in base $b \in B$.
D_b	Sets of aggregated departure nodes in base $b \in B$.
I	Set of midnight nodes.
E_b^j	Set of outgoing arcs from midnight nodes I in base $b \in B$ to nodes in D_b.
G_b^j	Set of outgoing arcs from nodes in A_b in base $b \in B$ to midnight nodes I.
M_b^j	Set of outgoing arcs from arrival nodes to departure nodes D_b.
S_l	Set of feasible schedules for (co)pilot $l \in L_b^j$.
Variables:	
y_p	Equals to 1 if pairing $p \in P_b$ is chosen and 0 otherwise.
e_f	Equals to 1 if flight $f \in F$ is not covered and 0 otherwise.
u_e	Indicate the flow on the outgoing arcs $e \in E_b^j$.
r_v	Equals to 1 if $v \in V_l$ is covered and 0 otherwise.
w_i^j	Flow on the waiting arc i between departure nodes i and $i+1$ in base $b \in B$.
h_g	Indicates the flow on the outgoing arcs $g \in G_b^j$.
t_m	Indicates the flow on the outgoing arcs $m \in M_b^j$.
x_s	Equals to 1 if schedule $s \in S_l$ is chosen and 0 otherwise.
e_p^j	Equals to 1 if $p \in P$ is not covered by (co)pilots and 0 otherwise.
e_v	Equals to 1 if $v \in V_l$ is not covered and 0 otherwise.
Parameters:	
c_f	Penalty cost for uncovered flight $f \in F$.
ρ_f	Duration of flight f.
ψ_f^p	Equals to 1 if flight f is covered by pairing p and 0 otherwise.
Q_b	Maximum available flying time for base $b \in B$.
c_v	Negative cost (bonus) for the coverage of $v \in V_l$.
c_p	Cost of pairing $p \in P$.
k_b	Number of available (co)pilots at base $b \in B$.
α_i^p	Equals to 1 if end of pairing $p \in P_b$ is incoming at node $i \in A_b$.
η_i^p	Equals to 1 if start of pairing $p \in P_b$ is outgoing at node $i \in D_b$.
γ_i^g	Equals to 1 if arc $g \in G_b^j$ is outgoing from node $i \in A_b$.
π_i^g	Equals to 1 if arc $g \in G_b^j$ is incoming at midnight node $i \in I$.
τ_i^m	Equals to 1 if arc $m \in M_b^j$ is outgoing from node $i \in A_b$.
ζ_i^m	Equals to 1 if arc $m \in M_b^j$ is incoming at node $i \in D_b$.
ϑ_i^e	Equals to 1 if arc $e \in E_b^j$ is incoming at node $i \in D_b$.
μ_i^e	Equals to 1 if arc $e \in E_b^j$ is outgoing from midnight node $i \in I$.
σ_i^v	Equals to 1 if $v \in V_l$ starts on day $i \in I$ and -1 if it finishes on this day.
$\bar{c}_p$	Penalty cost for uncovered pairing $p \in P$.
$\bar{c}_v$	Penalty cost for uncovered $v \in V_l$.
a_p^s	Equals to 1 if $p \in P$ is covered by personalized schedule $s \in S_l$.
v_v^s	Equals to 1 if $v \in V_l$ is covered by schedule $s \in S_l$.

Table 6.11 Notation for chapter 6

Sets:	
F	Set of scheduled flights to be covered.
L	Set of pilots.
O	Set of copilots .
V_l	Set of VRs for pilot $l \in L$.
V_o	Set of VRs for copilot $o \in O$.
S_l	Set of feasible schedules for pilot $l \in L$.
S_o	Set of feasible schedules for copilot $o \in O$.
A	Set of all the feasible connection arcs between two successive flights.
Variables:	
x_s	Equal to 1 if schedule $s \in S_l$ is allocated to pilot $l \in L$ and 0 otherwise.
y_s	Equal to 1 if schedule $s \in S_o$ is allocated to copilot $o \in O$ and 0 otherwise.
r_v	Equal to 1 if VR $v \in V_l$ is unsatisfied for pilot $l \in L$ and 0 otherwise.
z_v	Equal to 1 if VR $v \in V_o$ is unsatisfied for copilot $o \in O$ and 0 otherwise.
u_{ij}	Equal to 1 if $(i, j) \in A$ is covered by the pilot problem or copilot problem but not both, and 0 otherwise.
Parameters:	
c_v	Penalty cost for unsatisfied VR $v \in V_l$ (or $v \in V_o$).
e_f^s	Equal to 1 if flight $f \in F$ is covered by schedule $s \in S_l$ (or $s \in S_o$).
b_v^s	Equal to 1 if VR $v \in V_l$ (or $v \in V_o$) is satisfied by schedule $s \in S_l$ (or $s \in S_o$).
n_s	Number of PFs in schedule $s \in S_l$ (or $s \in S_o$).
B	Bonus cost (a negative cost) for covering each PF.
c_p	Cost of pairing p.
c_s	Cost of schedule $s \in S_l$ (or $s \in S_o$).
a_{ij}^s	Equal to 1 if connection arc $(i, j) \in A$ between flights i and j is covered by schedule s.
R_{ij}	Penalty cost for $(i, j) \in A$ if u_{ij} is equal to 1.

CHAPITRE 7 SYNTHÉSE DU TRAVAIL

Dans cette thèse, nous nous concentrons sur le problème de l'affectation personnalisée des équipes dans un contexte personnalisé avec un objectif global, où chaque vol doit être couvert par un pilote et un copilote. Chaque pilote et copilote demande un ensemble de vacances et de vols préférés par mois. L'utilisation d'une approche séquentielle pour résoudre ce problème réduit considérablement la complexité du processus, mais peut conduire à des solutions nettement sous-optimales, car les contraintes associées aux horaires et les objectifs ne sont pas pris en compte lors de la construction des rotations. Dans le premier objectif de cette thèse, nous avons développé un algorithme intégré basé sur la combinaison de la décomposition de Benders et la génération de colonnes traitant le cas ou il y a seulement des vacances préférées. Dans le deuxième objectif de cette thèse, nous avons étudié la difficulté d'utiliser l'algorithme proposée si nous voulons considérer un ensemble de vols préférés pour chaque pilote et copilote. Aussi, nous avons étudié la difficulté d'utiliser la relaxation lagrangienne pour résoudre ce problème. Enfin, dans le dernier objectif, nous avons développé une approche intégrée en combinant la décomposition Lagrangienne, génération de colonnes et l'agrégation dynamique de contraintes.

CHAPITRE 8 DISCUSSION GÉNÉRALE

Pour l'industrie aérienne, les coûts d'équipage sont les plus élevés après les coûts de carburant (Barnhart et al., 2003a). Par conséquent, même de légères améliorations dans la qualité des horaires peuvent avoir un avantage financier significatif. Ainsi, le problème de la planification équipage aérien est l'un des problèmes les plus importants et difficiles dans la planification des compagnies aériennes. En raison de sa taille et de sa complexité, ce problème est généralement résolu séquentiellement en deux étapes: la construction des rotations suivie par la construction des blocs mensuels. Cependant, il peut s'avérer impossible l'approche séquentielle d'obtenir avec une solution globale optimale car le domaine de décision du problème d'affectation des membres d'équipage est réduit par les décisions précédemment prises dans le problème de construction des rotations des membres d'équipage. En fait, la phase de construction des rotations ne peut pas déterminer les meilleurs rotations pour la phase d'affectation des équipages. Par conséquent, il est difficile de maximiser la satisfaction des préférences de l'équipage dans l'étape d'affectation de l'équipage et la solution l'optimalité de du problème de planification de l'équipage n'est pas garantie. Dans cette thèse, pour surmonter cette faiblesse, nous avons proposé deux algorithmes intégrés basés sur la combinaison de plusieurs méthodes de résolution afin de résoudre le problème considéré pour un ensemble donné de pilotes et de copilotes. Chaque pilote et copilote demande un ensemble de vols et de vacances préférés par mois. Pour réduise la propagation des perturbations au cours de l'opération, nous gardons les rotations des pilotes et des copilotes aussi semblables que possible. Nous optimisons les horaires pour les pilotes et copilotes simultanément, en tenant compte de leurs préférences.

CHAPITRE 9 CONCLUSION ET RECOMMANDATIONS

En résumé, cette thèse a fourni diverses contributions numériques et théoriques pour résoudre l'un des problèmes les plus importants et les plus difficiles du problème de la planification des compagnies aériennes, appelé *le problème de planification de l'équipage.*

Comme premier objectif, nous avons proposé deux nouveaux modèles mathématiques. Premièrement, une extension du problème de construction des rotations des membres d'équipage qui intègre les demandes de vacances du pilote et du copilote au stade de rotation de l'équipage afin de déterminer les bonnes rotations pour affectation des équipages. Deuxièmement, nous avons introduit un modèle qui intègre complètement les problèmes de construction des rotations et le problème d'affectation des équipes simultanément pour les pilotes et les copilotes. Nous avons présenté deux approches de solution séquentielles et une approche intégrée. Les approches séquentielles sont basées sur la génération de colonnes. L'approche intégrée est basée sur la combinaison la décomposition de Benders et la génération de colonne. Les résultats théoriques de cette dernière montrent aussi que l'optimalité peut être atteinte. Les résultats montrent que l'approche intégrée produit des améliorations significatives par rapport aux approches séquentielles.

Dans le deuxième objectif, nous avons étudié la difficulté d'utiliser la combinaison de la décomposition de Benders et la génération de colonnes développée dans le premier objectif en considérant un ensemble de PFs et de VRs pour chaque pilote et copilote. Nous avons envisagé trois types de coupe Benders: coupes fortes de Benders définies par les variables duales des rotations, coupes faibles de Benders et coupes fortes de Benders définies par les variables duales des vols. Nous avons montré qu'aucun d'entre eux ne pouvait s'occuper des PFs. Pour surmonter cette difficulté, nous avons proposé un nouveau modèle et une nouvelle méthodologie dans un troisième objectif. Le dernier a introduit un modèle qui intègre les rotations d'équipage et l'affectation personnalisée des blocs mensuels pour un ensemble donné de pilotes et copilotes simultanément. Nous avons présenté une approche intégrée basée sur la combinaison de trois méthodes ALD, CG et DCA. Cette nouvelle approche utilise la décomposition lagrangienne en alternance, proposant une nouvelle façon de mettre à jour les multiplicateurs lagrangiens. Nous avons prouvé que l'algorithme avait atteint une solution optimale lorsque l'ensemble des rotations sont identiques entre pilotes et copilotes. Les résultats de l'approche intégrée montrent que très peu d'itérations suffisent pour trouver une solution optimale (dans la plupart des cas), 3 ou 4 itérations, pour le problème complexe. De plus, les résultats ont monté que l'approche intégrée produit des améliorations significatives par rapport aux approches séquentielles en termes de satisfaction des VRs et des PFs. L'approche intégrée proposée peut être utilisée pour résoudre le problème général de la décomposition lagrangienne où les solutions identiques doivent être trouvées pour les copies de la solution dans les deux parties de la décomposition; voir la formulation générale suivante:

$$\min \quad Z = CX + \bar{C}Y \tag{9.1}$$

$$X \in D_1 \tag{9.2}$$

$$Y \in D_2 \tag{9.3}$$

$$X = Y \tag{9.4}$$

RÉFÉRENCE

M. R. Akhavan Kazemzadeh, T. G. Crainic, et B. Gendron, “A Lagrangian based matheuristic for multilayer single flow type multicommodity capacitated fixed charge network design”, Publication CIRRELT, Rapp. tech., 2018.

R. Anbil, E. Gelman, B. Patty, et R. Tanga, “Recent advances in crew-pairing optimization at American Airlines”, *Interfaces*, vol. 21, no. 1, pp. 62–74, 1991.

R. Anbil, R. Tanga, et E. L. Johnson, “A global approach to crew-pairing optimization”, *IBM Systems Journal*, vol. 31, no. 1, pp. 71–78, 1992.

A. Azadeh, M. H. Farahani, H. Eivazy, S. Nazari-Shirkouhi, et G. Asadipour, “A hybrid meta-heuristic algorithm for optimization of crew scheduling”, *Applied Soft Computing*, vol. 13, no. 1, pp. 158–164, 2013.

C. Barnhart et R. G. Shenoi, “An approximate model and solution approach for the long-haul crew pairing problem”, *Transportation Science*, vol. 32, no. 3, pp. 221–231, 1998.

C. Barnhart, E. L. Johnson, G. L. Nemhauser, M. W. Savelsbergh, et P. H. Vance, “Branch-and-price: Column generation for solving huge integer programs”, *Operations Research*, vol. 46, no. 3, pp. 316–329, 1998.

C. Barnhart, P. Belobaba, et A. R. Odoni, “Applications of operations research in the air transport industry”, *Transportation science*, vol. 37, no. 4, pp. 368–391, 2003.

C. Barnhart, A. M. Cohn, E. L. Johnson, D. Klabjan, G. L. Nemhauser, et P. H. Vance, “Airline crew scheduling”, dans *Handbook of Transportation Science*. Springer, 2003, pp. 517–560.

J. E. Beasley et B. Cao, “A tree search algorithm for the crew scheduling problem”, *European Journal of Operational Research*, vol. 94, no. 3, pp. 517–526, 1996.

J. F. Benders, “Partitioning procedures for solving mixed-variables programming problems”, *Numerische Mathematik*, vol. 4, no. 1, pp. 238–252, 1962.

R. E. Bixby, J. W. Gregory, I. J. Lustig, R. E. Marsten, et D. F. Shanno, “Very large-scale linear programming: a case study in combining interior point and simplex methods”, *Operations Research*, vol. 40, no. 5, pp. 885–897, 1992.

K. Boubaker, G. Desaulniers, et I. Elhallaoui, “Bidline scheduling with equity by heuristic dynamic constraint aggregation”, *Transportation Research Part B: Methodological*, vol. 44, no. 1, pp. 50–61, 2010.

V. Cacchiani et J.-J. Salazar-González, “Optimal solutions to a real-world integrated airline scheduling problem”, *Transportation Science*, vol. 51, no. 1, pp. 250–268, 2016.

K. W. Campbell, R. B. Durfee, et G. S. Hines, “Fedex generates bid lines using simulated annealing”, *Interfaces*, vol. 27, no. 2, pp. 1–16, 1997.

C.-H. Chen, T.-K. Liu, J.-H. Chou, et C.-C. Wang, “Multiobjective airline scheduling: An integrated approach”, dans *SICE Annual Conference (SICE), 2012 Proceedings of.* IEEE, 2012, pp. 1266–1270.

I. T. Christou, A. Zakarian, J.-M. Liu, et H. Carter, “A two-phase genetic algorithm for large-scale bidline-generation problems at Delta Air Lines”, *Interfaces*, vol. 29, no. 5, pp. 51–65, 1999.

A. M. Cohn et C. Barnhart, “Improving crew scheduling by incorporating key maintenance routing decisions”, *Operations Research*, vol. 51, no. 3, pp. 387–396, 2003.

J.-F. Cordeau, G. Stojković, F. Soumis, et J. Desrosiers, “Benders decomposition for simultaneous aircraft routing and crew scheduling”, *Transportation Science*, vol. 35, no. 4, pp. 375–388, 2001.

T. G. Crainic, A. Frangioni, et B. Gendron, “Bundle-based relaxation methods for multicommodity capacitated fixed charge network design”, *Discrete Applied Mathematics*, vol. 112, no. 1-3, pp. 73–99, 2001.

G. Desaulniers, J. Desrosiers, Y. Dumas, S. Marc, B. Rioux, M. M. Solomon, et F. Soumis, “Crew pairing at Air France”, *European Journal of Operational Research*, vol. 97, no. 2, pp. 245–259, 1997.

M. Desrochers et F. Soumis, "A generalized permanent labelling algorithm for the shortest path problem with time windows", *INFOR: Information Systems and Operational Research*, vol. 26, no. 3, pp. 191–212, 1988.

V. Dück, F. Wesselmann, et L. Suhl, "Implementing a branch and price and cut method for the airline crew pairing optimization problem", *Public Transport*, vol. 3, no. 1, p. 43, 2011.

W. El Moudani, C. A. N. Cosenza, M. De Coligny, et F. Mora-Camino, "A bi-criterion approach for the airlines crew rostering problem", dans *International Conference on Evolutionary Multi-Criterion Optimization.* Springer, 2001, pp. 486–500.

I. Elhallaoui, D. Villeneuve, F. Soumis, et G. Desaulniers, "Dynamic aggregation of set-partitioning constraints in column generation", *Operations Research*, vol. 53, no. 4, pp. 632–645, 2005.

I. Elhallaoui, A. Metrane, F. Soumis, et G. Desaulniers, "Multi-phase dynamic constraint aggregation for set partitioning type problems", *Mathematical Programming*, vol. 123, no. 2, pp. 345–370, 2010.

M. Gamache, F. Soumis, D. Villeneuve, J. Desrosiers, et E. Gelinas, "The preferential bidding system at air canada", *Transportation Science*, vol. 32, no. 3, pp. 246–255, 1998.

M. Gamache, F. Soumis, G. Marquis, et J. Desrosiers, "A column generation approach for large-scale aircrew rostering problems", *Operations research*, vol. 47, no. 2, pp. 247–263, 1999.

C. Gao, E. Johnson, et B. Smith, "Integrated airline fleet and crew robust planning", *Transportation Science*, vol. 43, no. 1, pp. 2–16, 2009.

I. Gershkoff, "Optimizing flight crew schedules", *Interfaces*, vol. 19, no. 4, pp. 29–43, 1989.

B. Gopalakrishnan et E. L. Johnson, "Airline crew scheduling: State-of-the-art", *Annals of Operations Research*, vol. 140, no. 1, pp. 305–337, 2005.

M. Guignard et S. Kim, "Lagrangean decomposition: A model yielding stronger Lagrangean bounds", *Mathematical Programming*, vol. 39, no. 2, pp. 215–228, 1987.

Y. Guo, T. Mellouli, L. Suhl, et M. P. Thiel, “A partially integrated airline crew scheduling approach with time-dependent crew capacities and multiple home bases”, *European Journal of Operational Research*, vol. 171, no. 3, pp. 1169–1181, 2006.

K. L. Hoffman et M. Padberg, “Solving airline crew scheduling problems by branch-and-cut”, *Management Science*, vol. 39, no. 6, pp. 657–682, 1993.

S. Irnich et G. Desaulniers, “Shortest path problems with resource constraints”, dans *Column Generation.* Springer, 2005, pp. 33–65.

A. I. Jarrah et J. T. Diamond, “The problem of generating crew bidlines”, *Interfaces*, vol. 27, no. 4, pp. 49–64, 1997.

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 Montreal*, 2015.

A. Kasirzadeh, M. Saddoune, et F. Soumis, “Airline crew scheduling: Models, algorithms, and data sets”, *EURO Journal on Transportation and Logistics*, pp. 1–27, 2014.

——, “Airline crew scheduling: Models, algorithms, and data sets”, *EURO Journal on Transportation and Logistics*, vol. 6, no. 2, pp. 111–137, 2017.

D. Klabjan, E. L. Johnson, G. L. Nemhauser, E. Gelman, et S. Ramaswamy, “Solving large airline crew scheduling problems: Random pairing generation and strong branching”, *Computational Optimization and Applications*, vol. 20, no. 1, pp. 73–91, 2001.

——, “Airline crew scheduling with time windows and plane-count constraints”, *Transportation Science*, vol. 36, no. 3, pp. 337–348, 2002.

B. Maenhout et M. Vanhoucke, “A hybrid scatter search heuristic for personalized crew rostering in the airline industry”, *European Journal of Operational Research*, vol. 206, no. 1, pp. 155–167, 2010.

R. E. Marsten et F. Shepardson, “Exact solution of crew scheduling problems using the set partitioning model: Recent successful applications”, *Networks*, vol. 11, no. 2, pp. 165–177, 1981.

A. Mercier et F. Soumis, “An integrated aircraft routing, crew scheduling and flight retiming model”, *Computers & Operations Research*, vol. 34, no. 8, pp. 2251–2265, 2007.

A. Mercier, J.-F. Cordeau, et F. Soumis, “A computational study of Benders decomposition for the integrated aircraft routing and crew scheduling problem”, *Computers & Operations Research*, vol. 32, no. 6, pp. 1451–1476, 2005.

N. Papadakos, “Integrated airline scheduling”, *Computers & Operations Research*, vol. 36, no. 1, pp. 176–195, 2009.

F. Quesnel, G. Desaulniers, et F. Soumis, “A new heuristic branching scheme for the crew pairing problem with base constraints”, *Computers & Operations Research*, vol. 80, pp. 159–172, 2016.

R. Rahmaniani, T. G. Crainic, M. Gendreau, et W. Rei, “The Benders decomposition algorithm: A literature review”, *European Journal of Operational Research*, vol. 259, no. 3, pp. 801–817, 2017.

M. Saddoune, G. Desaulniers, I. Elhallaoui, et F. Soumis, “Integrated airline crew scheduling: A bi-dynamic constraint aggregation method using neighborhoods”, *European Journal of Operational Research*, vol. 212, no. 3, pp. 445–454, 2011.

——, “Integrated airline crew pairing and crew assignment by dynamic constraint aggregation”, *Transportation Science*, vol. 46, no. 1, pp. 39–55, 2012.

M. Saddoune, G. Desaulniers, et F. Soumis, “Aircrew pairings with possible repetitions of the same flight number”, *Computers & Operations Research*, vol. 40, no. 3, pp. 805–814, 2013.

R. Sandhu et D. Klabjan, “Integrated airline fleeting and crew-pairing decisions”, *Operations Research*, vol. 55, no. 3, pp. 439–456, 2007.

S. Shao, H. D. Sherali, et M. Haouari, “A novel model and decomposition approach for the integrated airline fleet assignment, aircraft routing, and crew pairing problem”, *Transportation Science*, vol. 51, no. 1, pp. 233–249, 2015.

N. Souai et J. Teghem, “Genetic algorithm based approach for the integrated airline crew-pairing and rostering problem”, *European Journal of Operational Research*, vol. 199, no. 3, pp. 674–683, 2009.

S. Subramanian et H. D. Sherali, “An effective deflected subgradient optimization scheme for implementing column generation for large-scale airline crew scheduling problems”, *INFORMS Journal on Computing*, vol. 20, no. 4, pp. 565–578, 2008.

P. H. Vance, C. Barnhart, E. L. Johnson, et G. L. Nemhauser, “Airline crew scheduling: A new formulation and decomposition algorithm”, *Operations Research*, vol. 45, no. 2, pp. 188–200, 1997.

J. D. Weir et E. L. Johnson, “A three-phase approach to solving the bidline problem”, *Annals of Operations Research*, vol. 127, no. 1-4, pp. 283–308, 2004.

F. Zeghal et M. Minoux, “Modeling and solving a crew assignment problem in air transportation”, *European Journal of Operational Research*, vol. 175, no. 1, pp. 187–209, 2006.

V. Zeighami et F. Soumis, “Combining Benders decomposition and column generation for integrated crew pairing and personalized crew assignment problems”, Cahiers du GERAD, Rapp. tech. G–2017–41, 2017.