Minimising propagated delay in an integrated aircraft routing and crew pairing framework

Author:
Dunbar, Michelle Elizabeth

Publication Date:
2012

DOI:
https://doi.org/10.26190/unsworks/15571

MINIMISING PROPAGATED DELAY IN AN INTEGRATED AIRCRAFT ROUTING AND CREW PAIRING FRAMEWORK

Michelle Elizabeth Dunbar

Supervisor: Associate Professor Gary Froyland

School of Mathematics and Statistics,

The University of New South Wales.

January 2012

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY AT THE UNIVERSITY OF NEW SOUTH WALES, JANUARY 2012.

ORIGINALITY STATEMENT

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

Signed: . . . .

Date: . . . .

COPYRIGHT STATEMENT

Signed: . . . .

Date: . . . .

AUTHENTICITY STATEMENT

I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.

Signed: . . . .

Date: . . . .

Abstract

For reasons of tractability, the airline scheduling problem has traditionally been sequentially decomposed into various stages (eg. schedule generation, fleet assignment, aircraft routing, and crew pairing), with the decisions from one stage imposed upon the decision making process in subsequent stages. Whilst this approach greatly simplifies the solution process, it unfortunately fails to capture the many dependencies between the various stages, most notably between those of aircraft routing and crew pairing, and how these dependencies affect the propagation of delays through the flight network. As delays are commonly transferred between late running aircraft and crew, it is important that aircraft routing and crew pairing decisions are made together. The propagated delay may then be accurately estimated to minimise the overall propagated delay for the network and produce a robust solution for both aircraft and crew.

In this thesis we introduce a new approach to accurately calculate and minimise the cost of propagated delay, in a framework that integrates aircraft routing and crew pairing. Additionally, we propose an extension on this model, in which we incorporate scheduling decisions; allowing higher quality aircraft and crew assignments to be obtained. Finally, we propose a new re-timing heuristic that may be used in conjunction with an incumbent aircraft and crew assignment, capable of simultaneously re-timing aircraft and crew whilst retaining the solution structure. We apply our approaches on a real-world airline network and provide numerical results for a number of test instances. Our results indicate that our new approaches perform very well on the test instances and outperform a number of existing models in a number of areas.

Contents

Acknowledgements

I would like to take this opportunity to thank a number of people who have contributed in the production of this thesis. First and foremost, I would like to thank the School of Mathematics and Statistics at UNSW for providing me with the opportunity to undertake this Ph.D. I'd especially like to thank my supervisor A/Prof. Gary Froyland for his guidance, insight and advice over the past four years and my co-supervisor from Aviation, Dr. Richard Wu for explaining the many intricacies involved with the airline scheduling process; it has been a privilege to work with both of them.

I'd like to thank the various members of staff who additionally allowed me to take part in the School's various teaching and outreach programs, allowing me to hone my teaching and lecturing skills and ensuring that my Ph.D. experience was not only well-rounded, but also extremely enjoyable. The funding support of the Australian Research Council under its Centres of Excellence program is also gratefully acknowledged.

To the Ph.D. students with whom I shared an office, Boris, Ognjen and Tim, thank you for the enjoyable conversations, support and encouragement. I have enjoyed every minute in your company. Finally, I must thank my family for their continual support and encouragement over the last four years.

Michelle Dunbar,

27 January, 2012.

CHAPTER

ONE

Introduction

Since the dawn of flight, the airline industry has continually developed and expanded internationally, with a vast array of large international carriers and small domestic carriers making it possible to travel to all major capital cities and even the most remote locations of the world. The airline industry plays a pivotal role in not only shrinking our world through tourism, but in driving economic and social progress through both the transport of business passengers and the transport of freight. The world's airlines have provided access to global markets, allowing for the development of world economies and have no doubt played a key role in enabling the rise of globalisation. The enormity of this influence on the global economy was perhaps no more evident than after the shut down of UK and European airspace throughout the eruption of the Icelandic volcano, Eyjafjallajökull in 2010, with economies taking many months to recover.

In order to capitalise on these opportunities, airlines must best position themselves within the global arena. To achieve this, airlines have continually sought to gain a competitive advantage over their rivals by servicing the most profitable markets, competing for slot times at major hubs and reducing overall costs. With the onset of deregulation in the late 1970s (most notably in the USA) and in the 1990s (for Europe) competition between the major airlines has become increasingly fierce, with established airlines having to compete with an increasing number of low-cost carriers. Thus the traditional focus of airline schedule planners has been largely restricted to cost-cutting and consequently the primary tenet of airline schedule planning has for many years been one of maximising profit.

Favouring such an approach however has come at a price, as the schedules generated using this approach do not perform well in an operational setting and are often considered to be "over-optimised". This stems from the fact that maximising profit at the exclusion of other key factors has a tendency to generate schedules that are very "tight", with limited connection time between flights; since aircraft and crew are most profitable whilst they are in the air. The resulting schedules are therefore highly brittle in practice; that is, they have the capacity to perform well when all flights depart and arrive as planned, but collapse rather dramatically when an unexpected disruption occurs, allowing delays to propagate rapidly throughout the network. This is evidenced by the recent statistics from the Bureau of Transportation Statistics (BTS) [76] demonstrating that approximately 21.5% of flights in the U.S. between the months of July 2010 and July 2011, were delayed and failed to meet On-Time Performance (OTP) measures, with delay propagation between late-arriving aircraft being the pre-dominant cause.

In recent years, this dramatic increase in schedule disruption has resulted in an ever increasing discrepancy between planned costs and realised operational costs. As aircraft networks continue to grow, this trend is set to continue with AhmadBeygi *et al.* [4] reporting that in 2006, it was estimated that the US airline industry experienced a total of 116.5 million minutes of delay; translating into a $7.7 billion increase in operating costs. Such large discrepancies have prompted airline schedule planners to shift their focus from maximising profit to maximising expected profits under uncertainty, by including various types of costs arising from unplanned events.

In 2010 the number of airlines worldwide totalled 1,629, with 27,271 aircraft and 29.6 million departures per year. Coupled with this, an increase in passenger numbers in the last few years has resulted in approximately 2.8 billion passengers travelling worldwide annually (IATA Fact Sheet, 2010) [49]. With passenger growth expected to rise in the next 10 - 15 years, the effects of delay propagation and network robustness can no longer be ignored. Moreover, there is a growing need for schedules to possess the ability to recover quickly from disruptions.

The last few years have witnessed a growing body of research addressing the need to capture uncertainty within the airline scheduling problem. In addition, attempts to improve the robustness of airline schedules and their ability to recover quickly from disruption is also an area of growing interest. A detailed summary of the recent literature can be found in Chapter 2 and 3.

1.1 Tactical Planning In The Airline Industry

The airline scheduling problem in its entirety is incredibly complex as airlines are not only required to construct schedules, that is a set of origin-destination flight pairs with departure and arrival times for each flight (e.g. SYD→MEL, departing SYD at 0800 and arriving in MEL at 0930) but are additionally required to assign aircraft and crew to cover each of these flight legs. This is complicated further by the fact that airlines typically possess many aircraft of different fleet type (e.g. Boeing 767 or A380) with different seating capacities and ranges; together with multiple crew groups that are required to be utilised to their fullest extent with the restriction that they may only be assigned to certain fleet type(s). Moreover, both aircraft and crew must satisfy certain requirements for maintenance, the return of crew to their crew-base of origin and numerous crewing restrictions[1] predominantly relating to the maximum number of flying hours in a given duty.

The aforementioned constraints are but a few of the real-world difficulties faced by airlines in the process of constructing a feasible schedule; not to mention defining an objective and eventually obtaining an optimal schedule (and assignment). Consequently, in order to retain any degree of tractability, the airline scheduling problem has been traditionally decomposed sequentially in the following manner (e.g. schedule generation, fleet assignment, aircraft routing, and crew pairing), with the decisions from one stage imposed upon the decision making process in subsequent stages. A more detailed description of each of these stages and methods used to solve them is covered in Chapter 2.

Whilst the sequential approach greatly simplifies the solution process, it unfortunately fails to capture the many dependencies between the various stages, most notably between those of aircraft routing and crew pairing, and how these dependencies affect the propagation of delays through the flight network. As delays are commonly transferred between late running aircraft and crew, it is important that

[1] Mutually agreed upon rules between an airline and its crew union(s).

aircraft routing and crew pairing decisions are made together. The propagated delay may then be accurately estimated to minimise the overall propagated delay for the network and produce a robust solution for both aircraft and crew.

In this thesis we introduce new approaches for accurately calculating and minimising the cost of propagated delay, in an integrated aircraft routing and crew pairing framework. Additionally we seek to ensure that we do not increase running costs above that of the profit-maximisation case and attempt to achieve this by ensuring that we use the same number of aircraft and crew as used in the profit-maximisation case.

1.2 Causes Of Disruption

According to the Bureau of Transportation Statistics [76], approximately 21.5% of all flight legs in the U.S. did not satisfy key On-Time Performance (OTP) measures between the months of July 2010 and July 2011. A flight is classified as arriving "on-time" if it does not arrive more than 15 minutes after its specified arrival time. The various causes of disruption are listed in Figures 1.1 and 1.2 below.

Propagated delay transferred between late-arriving aircraft accounts for 7.21% of delays, and approximately 2% of delays are the result of flight cancellations. It is also clear that the National Aviation System accounts for a large proportion of the delays, just over 6%; demonstrating that even the best laid plans for an individual airline may be thwarted by the disruptions caused by other airlines in the airspace.

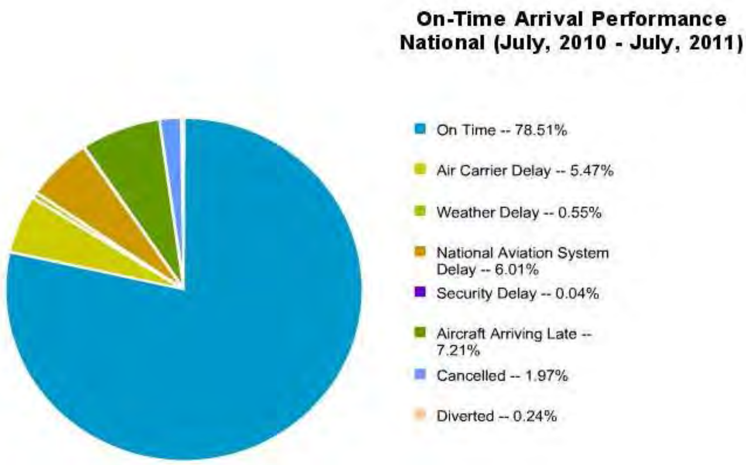


Figure 1.1: The Causes of Disruption [76].

This may be the result of adverse weather such as fog or snow, congestion at busy airports and aircraft (or computer) breakdowns to name a few. In Figure 1.2 below, it may be observed that weather accounts for approximately 62.11% of delay for the National Aviation System, with volume the second largest cause at 28.36%.

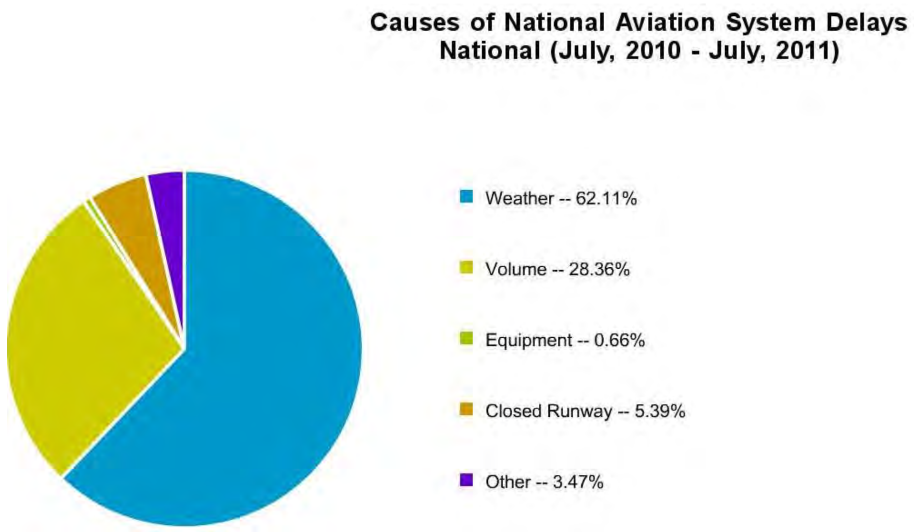


Figure 1.2: Cause of National Aviation System Delays [76].

This demonstrates that weather may have a significant effect in influencing delays and although weather disruptions may be somewhat expected at certain times of the year (e.g. throughout the northern Winter), other disruptions such as fog and thunderstorms may not be as predictable. Therefore it may not be possible to prepare for such disruptions during the long-term planning stages and thus it becomes necessary to design schedules that are better equipped to quickly recover from unexpected delays, in order to minimise the effect of delay propagation throughout the network. Furthermore, as volume is the second largest factor contributing to delay for the National Aviation System, reducing the effect of delay propagation will become increasingly important as traffic volume continues to increase into the future.

1.3 Outline Of Thesis

We now briefly outline the content of each Chapter and the contributions of this Thesis.

Chapter 2: The Airline Scheduling Problem

In Chapter 2, we introduce the Airline Scheduling Problem and discuss each of the individual planning stages that form the sequential solution approach to tactical planning. We also introduce and outline recent attempts at integrating the various stages.

Chapter 3: Capturing Uncertainty: Improving Robustness

In Chapter 3, we discuss recent approaches in the literature for including uncertainty and improving the robustness of various aspects of the airline scheduling problem. We briefly outline a number of current models that seek to produce robust, integrated solutions.

Chapter 4: Minimising Propagated Delay In an Integrated Aircraft Routing and Crew Pairing Framework

In Chapter 4 we outline in detail our proposed model for minimising propagated delay in an integrated aircraft routing and crew pairing framework. We discuss the mathematical formulation behind our model and describe how this work improves upon existing models in the field. We then obtain results for our model on data from a real airline network and demonstrate that our model outperforms existing models in a number of key areas.

Chapter 5: The Re-timing Heuristic

In Chapter 5 we explore the possibility of obtaining a potentially more operationally robust solution by proposing a heuristic capable of simultaneously re-timing aircraft and crew whilst attempting to minimise overall delay propagation in the network. We demonstrate that, despite its simplicity, the heuristic performs very well on a number of datasets.

Chapter 6: Integration of Aircraft Routing and Crew Pairing with the Re-timing Heuristic: Including Scenarios within the Subproblem and Re-timing Heuristic

In Chapter 6, we propose two new methods of embedding delay scenarios within the aircraft routing and crew pairing subproblems of the model proposed in Chapter 4 and additionally utilise these scenarios within the re-timing heuristic of Chapter 5. We demonstrate that embedding each of these methods within the subproblem is beneficial and when further combined with the improved re-timing heuristic, the solution obtained out-performs the solution obtained using the heuristic (based on mean delays) of Chapter 5. Additionally we investigate whether it is possible to improve upon this re-timed solution by re-solving each approach using the new departure times obtained from the heuristic. We demonstrate that for certain delay instances this approach is highly effective and able to achieve significant improvements with relatively short computation times.

Chapter 7: Integration of Aircraft Routing, Crew Pairing and Re-timing: Using Scenarios in the Subproblem

Finally, In Chapter 7 we propose a fully integrated model that captures the aspects of aircraft routing, crew pairing and re-timing without introducing any extra complexity within the master problem. All three decisions are made simultaneously by allowing re-timing decisions to be made within the aircraft routing subproblem. We demonstrate that this proposed method improves the solution quality further and has the capacity to achieve significant delay improvements; out-performing all others used in the thesis.

Chapter 8: Summary of Contributions

Finally, In Chapter 8, we provide a summary of our contributions and some possible areas for future research.

CHAPTER

TWO

The Airline Scheduling Problem

2.1 The Sequential Solution Approach

As mentioned in Chapter 1, in order to retain any degree of tractability, the Airline Scheduling Problem has traditionally been decomposed into four broad stages; namely, Schedule Generation (SG), Fleet Assignment (FA), Aircraft Routing (AR) and Crew Pairing (CP). The need for such a decomposition stems directly from the fact that the Airline Scheduling Problem in its entirety, is inherently very complex.

In the modern era of travel, airlines face an almost overwhelming task of meeting all (sometimes conflicting) objectives. In particular, airlines must contend with the rising cost of fuel, aircraft and crew costs, fluctuations in passenger numbers, competition with other airlines for established routes, airport congestion, brand loyalty and the need to fulfill basic key performance indicators (such as On-Time-Performance (OTP)). Such examples are but a few of the many obstacles placed before an airline, even before an aircraft leaves the ground.

Moreover, building a mathematical model capable of capturing and accurately expressing all desirable objectives would be incredibly difficult, if not intractable to solve. For this reason, airline schedule planners have modelled the most necessary aspects and sequentially decomposed the problem into the four stages mentioned above, with the output from each stage providing the input for subsequent stages. A general overview of the various stages is provided in the following papers [38, 43, 8, 54].

In recent years, increases in computing power have enabled the partial integration of two or more stages of the airline scheduling problem, allowing for higher quality solutions and the elimination of potential incompatibilities. Indeed, research into the improvement of integration techniques has been a significant area of growth in the last few years, with results demonstrating that even partial integration of key components can deliver solutions that are not only closer to cost-optimal but potentially more operationally robust.

In this Chapter, we outline each of these stages and briefly mention some of the ways in which various authors have addressed these problems. We then outline a few attempts at integrating one or more of these stages to provide a better solution. Finally, we conclude the Chapter by introducing the solution technique known as column generation, a solution technique extensively used within the airline schedule planning literature.

2.1.1 Schedule Generation

The first stage of the Airline Scheduling Process involves the design and generation of the schedule. An airline schedule is typically generated several months prior to the day of operations, and there often exists a separate summer and a winter schedule in order to meet fluctuations in passenger demand and cater to specific origin-destination flight pairs.

In planning a schedule, an airline must determine the *markets* it wishes to serve, the *frequency* with which it serves these markets and lastly the specific *departure and arrival times* of each flight leg (Lan *et al.* (2006) [57]). The choice of markets is crucial as it determines the choice of origin-destination pairs and the resulting profit from these pairs. The choice of markets to serve also depends on the fleet of aircraft the airline has available and whether they intend to favour business travellers, domestic or international flights, or compete with other airlines for established routes.

In practice, there are often many restrictions placed on the amount by which the schedule may be changed from year-to-year as airlines prefer to purchase and retain favourable slot times to ensure schedule saleability. Furthermore, slots may be expensive to purchase at certain airports, or major hubs may be more congested at certain times of the day. A more comprehensive summary of restrictions is provided in Ageeva (2000) [3]. As the schedule generation is performed first, it has the potential to affect every subsequent operational decision, and as such, has the capacity to dramatically affect the airlines profitability. The following paper by Feo and Bard (1989) outlines a model used to construct flight schedules [39].

2.1.2 Fleet Assignment

Once an airline has determined a suitable schedule, the next task is to assign a specific aircraft *type* to each flight leg in the schedule. The objective of the Fleet Assignment problem is to determine an optimal assignment, so as to closely match the capacity of each aircraft with the expected number of passengers for each flight, whilst simultaneously keeping operational costs to a minimum.

Matching the expected passenger demand with seating capacity is of importance to an airline, as operational costs associated with larger aircraft are typically much greater than that for aircraft with smaller seating capacity. However, despite having lower operational costs, smaller aircraft may incur larger passenger spill costs if incorrectly assigned to a flight leg with high passenger demand.

Definition 2.1.1 (Spill Cost). *Spill cost is defined as the amount of lost revenue resulting from insufficient seating capacity.*

Ageeva (2000) [3] states that a feasible assignment must satisfy multiple hard constraints such as limits on aircraft fuel capacity, aircraft range (longest distance the aircraft can travel) and limitations on gate availability and the number of aircraft on ground at certain airports, to name a few.

The fleet assignment problem has been well studied, with detailed surveys contained within the following papers [1, 46, 54, 89, 12, 10]. In the literature, the fleet assignment problem is usually solved as a multi-commodity flow problem with side constraints, where each aircraft may be considered to be a “commodity” that needs to “flow” across the network in a minimal cost way, whilst maintaining aircraft balance (amongst different aircraft fleets) and ensuring that each flight leg is covered by exactly one aircraft type (Sarmadi *et al.* (2004) [86]).

An optimal fleet assignment may be non-trivial for an airline to achieve, as major airlines typically possess multiple fleet types (e.g Boeing 767 or A380) and it may additionally be difficult to accurately estimate the passenger demand for particular flights many months prior to departure. For this reason, researchers have recently attempted to incorporate elements of stochasticity into their models as a means of better reflecting this uncertainty. Although this approach partially addresses the problem of uncertainty, there is often an over-compensation in an attempt to minimise passenger spill and consequently, such solutions may produce unfavourable solutions involving assignments with large numbers of empty seats, representing lost revenue for the airline. Additionally, in practice, airlines typically overbook flights in the expectation that some passengers may cancel bookings before departure. The obvious drawback of such an approach is the potential to cause passenger *spill*, as a result of insufficient seating capacity.

In an attempt to avoid the lost revenue associated with such solutions, a relatively new area of research involves the development of two-stage stochastic models in which the fleet assignment problem is typically solved a few *months* prior to the day of operations and then re-solved once passenger demand information is known with greater certainty a few *days* prior to the day of operations (Sherali *et al.* (2008) and Berge *et al.* (2008) [90, 19]). The concept of two-stage stochastic programming models will be discussed in greater detail in Chapter 3.

Another shortcoming of the traditional fleet assignment model is the use of simplistic assumptions for computing the revenue and passenger spill cost for each flight leg and consequently each route (or string of flights). Moreover, the basic fleet assignment model does not take into account the dependency between different legs within a string, nor the effect of spill and recapture from other parts of the network - both of which are important factors in calculating spill costs. In an attempt to address these issues Barnhart *et al.* (2000, 2008) [12, 11] proposed an *Itinerary-based Fleet Assignment Model* (IFAM) and an enhanced revenue model in which they consider network effects allowing them to more accurately assess the profitability of the fleet assignments. In these papers, the authors embed the spill optimisation problem within the fleet assignment problem. Various attempts to improve revenue management for the fleet assignment problem will be discussed further in Chapter 3.

Once a fleet type has been assigned to each of the flight legs, the flight legs are then typically grouped by fleet type and the subsequent stages of Aircraft Routing and Crew Pairing are solved *within* each fleet type. This approach not only reduces the complexity of the aircraft routing and crew pairing problems, but has the additional benefit of only requiring the enforcement of flow balance of aircraft and crew within each fleet. This decomposition of flights legs by fleet type often occurs in practice, as for example technical flight crew have restrictions[1] placed on the specific fleet type(s) on which they may fly. Therefore, such a decomposition has the additional benefit of simplifying the crew pairing problem further.

2.1.3 Aircraft Routing

The Aircraft Routing problem consists of two parts, namely *maintenance routing* and *tail-assignment*. Maintenance routing is performed a few weeks or months prior to the day of operations and involves an assignment of aircraft to flight legs in a minimal cost way; subject to satisfying relevant aircraft maintenance constraints. This assignment is performed without reference to specific tail numbers and so is feasible for any aircraft within the fleet. A few days prior to the day of operations, once more detailed information about aircraft availability may be determined, airline schedule planners specify an individual aircraft (along with its tail-number) to fly one of the routes previously chosen by the maintenance routing problem. This process continues until each of the aircraft routes determined by the maintenance routing problem have been assigned a specific tail-number. The process of assigning a particular flight tail-number to each aircraft route is known as tail-assignment. As this thesis is primarily concerned with long-term planning, we will hereafter focus our attention solely on the maintenance routing problem.

As the name suggests, maintenance routing requires the construction of a feasible assignment of aircraft to flights legs for which each aircraft route (sequence of flight legs) is maintenance feasible. As aircraft must be regularly maintained, there are four types of maintenance checks that each aircraft must undergo, with some checks occurring more frequently than others. These four checks are referred to as A-checks, B-checks, C-checks and D-checks. Of these, the A-check occurs most frequently, per-

[1] Technical flight crew (pilots) can only be certified for one type of aircraft at any given time, whilst cabin crew can hold multiple certifications for different aircraft types.

formed every three to four days, with the aircraft out of service for between 10-20 hours (usually overnight), occurring at the end of a rotation. The remaining three checks occur less frequently and involve a significant overhaul of the aircraft, with the aircraft taken out of service for a longer period of time. As a consequence, the maintenance routing problem typically only considers A-checks as part of the model. Maintenance feasibility is difficult to model, with most formulations opting for "plane-count constraints" and "ground arcs" to ensure the number of aircraft on the ground overnight is sufficiently high to "approximately" model overnight maintenance. This assumption is reasonable for major hubs for which the aircraft usually return to the hub at the end of the day.

The aircraft routing problem has received significant attention in recent years and the following papers provide a comprehensive summary: [26, 30, 42, 54, 25]. Clarke *et al.* (1997) [26] construct a flight-based formulation to solve the aircraft maintenance routing problem. They select a group of sub-tours that cover all flights, satisfy maintenance requirements and maximise through revenue. Gopalan and Talluri (1998) [43] propose an alternative approach in which the aircraft maintenance routing problem is modelled using graph theoretic techniques. The authors construct a directed graph $G = (V, E)$ where V denotes the set of stations/airports and E, the flying path. The problem is solved by identifying Euler tours within this directed graph.

More recently, the aircraft maintenance routing problem has been posed as a connection based model (or multi-commodity flow) (Cordeau *et al.* (2001) [29]) in which the decision variables model the connections between various flight legs. However such models may be prohibitively large to solve, containing thousands or even millions of integer variables, each describing the decision to include or exclude the feasible connection (feasible arc) from the solution. Moreover, many of these variables may be fractional, requiring branch and bound, resulting in lengthy computation times. In an attempt to "break the curse of dimensionality", authors have recently begun to model this problem as a set-partitioning problem for which the decision variables correspond to different aircraft routes, commonly referred to as "strings".

Definition 2.1.2 (Flight String)**.** *A flight string is a sequence of connected flights that start and end at a maintenance station, satisfy flow balance and incorporate maintenance requirements (such as maximum flying time between stations).*

This allows the problem to be solved via column generation for which the subproblem feeds 'maintenance feasible' strings into the master problem. This approach has the advantage that complicated constraints (such as maintenance feasibility constraints) can be captured relatively easily within the subproblem; whereas, such constraints would be somewhat more difficult to capture using a leg-based formulation. Moreover, although there may be many millions of feasible aircraft routes, only a small selection of them will allow coverage of all flights in a minimal cost way. The use of column generation allows the user to start with an initial feasible set of aircraft routes and only generate beneficial (cost-reducing) routes until an optimal solution is obtained. The reader is referred to Section 2.3 for a brief outline of the column generation approach and its applicability to aircraft routing (and crew pairing).

The objective of the Aircraft Routing problem is to achieve a minimal cost assignment of aircraft to flight legs. For many years, the focus was one of maximising *through revenue*. Through revenue is the revenue obtained by ensuring that passengers remain on the same aircraft where they would normally be required to meet a connecting aircraft. Until recently, this was considered an important aspect of planning, as passengers are sometimes willing to pay a premium to remain on the same aircraft, rather than make a connection. However, in practice it has been found that the increase in revenue is not overly significant (Lan *et al.* (2006) [57]), and "in reality many of these planned through itineraries are broken in the operation phase as a result of aircraft swapping" (Sarmadi (2004) [86]).

Recent increases in delay costs have prompted airline schedule planners to begin to include robustness measures as part of the objective cost. Authors such as Lan *et al.* (2006) [57] and Weide *et al.* (2007) [98] have proposed string based models that attempt to incorporate delay propagation costs and a non-robustness measure respectively. See Chapter 3 for a more detailed description of these and other approaches for including robustness within the aircraft routing problem. In Chapter 4 we outline our proposed integrated model for both aircraft routing and crew pairing that seeks to minimise propagated delay.

2.1.4 Crew Pairing

The last stage of the airline scheduling problem is that of crew pairing. The crew pairing problem is to determine a set of crew itineraries (known as pairings) that partition the flights in the network into disjoint sets so as to minimise crew cost, whilst satisfying regulatory agency requirements and collective bargaining agreements (Lan *et al.* (2006) [57]). As the crew are usually only allowed to be assigned to a particular fleet, the solutions to the fleet assignment and aircraft routing problems are used as input for this problem. As with the fleet assignment problem, the crew pairing problem has been well studied. The reader is referred to the following papers [97, 9, 56, 66, 55, 27], for a comprehensive survey.

The crew pairing problem, is often posed as set-partitioning problem in a similar manner in that of the aircraft routing problem. One of the distinct differences between the aircraft and the crew pairing problems is the requirement that crew must return to the crew base at which they started their duty and must comply with restrictions on working and flying time throughout their duty. These constraints are typically difficult to capture within a leg-based model, but as for the aircraft routing problem, are relatively simple to enforce within the subproblem of a column generation framework.

Definition 2.1.3 (Crew Base [85])**.** *A crew base is a designated airport in the network at which crews are stationed. A crew base is often a major hub for the airline.*

Definition 2.1.4 (Duty [85])**.** *A duty is a working day (or number of days) for a crew and consists of a sequence of flights that return to the crew-base of origin. A duty is subject to a number of regulatory and union rules (eg. Minimum/maximum sit-times (connection times) between two successive flights, maximum flying time, time away from base and elapsed time on duty).*

In practice, there are many restrictions enforced on crew, depending on the airline, the airline's objective and individual union rules. The reader is referred to Barnhart *et al.* (2003) [9] for a more detailed description of these rules and regulations. However, the following constraints are often included within a simplified crew pairing model: Namely, the maximum number of flying hours in a duty is 8 hours, with maximum duration of a duty set to 10 hours. When an aircraft has an overnight, the minimum duration for a night rest is 10 hours. It is possible to extend the maximum duration of a duty under certain circumstances, but we will make use of the aforementioned rules in our models.

Crew costs contribute a large part of an airline's expenses, and are second only to fuel costs. "The cost of a duty is often expressed as the maximum of three quantities: the flying time, a fraction of the elapsed time, and the minimum guaranteed pay" (Sandu *et al.* (2007) [85]). In some countries, as crew costs are essentially fixed, the objective is to minimise the number of crew pairings to cover all flights or additionally to minimise reserve crew. In some cases it may also be favourable to minimise the number of overnights (to minimise accommodation costs), or to avoid the requirement of deadheads (crew that have to be re-routed to make a connecting flight, but fly as passengers so as not to violate crew working hours). As with the aircraft routing problem, we will seek to minimise delay propagation for the crew pairing problem, see Chapter 4. In the next section we will outline attempts to integrate the various stages of the airline scheduling problem.

2.2 Integrated Solution Approaches

As was mentioned earlier in the Chapter, one of the significant drawbacks of the sequential solution approach is the sub-optimality of the overall solution; as decisions fixed earlier in the planning process limit the choices that can be made in the subsequent stages(s). In an attempt to rectify this problem, many authors have in recent years begun to integrate two or more of the planning stages in order to better model and integrate a variety of key decisions that should ideally be made simultaneously.

The growing interest in the integration of various planning stages has no doubt been fuelled by the growing need to improve operational performance and increase profit; as improvements as small as 1% in crew costs can translate into millions of dollars in revenue for an airline. As early as 2001, Cordeau *et al.* (2001) [29] reported that their integrated aircraft routing and crew pairing approach was capable of achieving a reduction in variable crew costs by 9.4% using data from a Canadian airline. With such improvements to be made, research into various methods of integration has been an area of significant growth in the last few years; aided in part by significant increases in computing power and improved solution techniques.

Integration of the various stages has occurred in a number of forms and began to feature more prominently in the literature around the late 1990s; Lu [64] and Barnhart [13] are two such authors pioneering this early approach to integration. Limited computing power made it necessary to avoid an explosion in size (both in terms of the number of variables and number of constraints) of an integrated model. Consequently, early authors adopted an approximate approach; opting for constraints that approximately modelled one of the stages whilst maintaining the sequential solution process.

In these papers, both Lu and Barnhart [64, 13] solved the fleet assignment aircraft (maintenance) routing and crew planning problems sequentially, whilst calculating the "approximate impact of crew scheduling on various fleetings" within the fleet assignment problem. This resulted in an approximate integrated model that still makes use of the sequential solution process. We now outline a few of the different integrated models that have been proposed in the last few years whilst briefly mentioning the rationale behind each approach and relevant solution techniques.

2.2.1 Fleet Assignment, Aircraft Routing Or Crew Pairing Combined With Schedule Generation

Recognising that schedule generation plays a key role in determining the feasibility of subsequent aircraft and crew assignments, a fairly large number of authors have attempted to combine (an approximation of) schedule generation with fleet assignment, aircraft routing and crew pairing.

Under the assumption that the availability of take-off and landing slots (and thus the schedule) does not vary significantly from year to year, authors have attempted to incorporate more flexibility (and potentially improve operational robustness) via the introduction of *time windows*. Time windows allow the departure time to fall anywhere within a discretised window, usually extending 10-15 minutes either side of the originally scheduled departure time. This is achieved by producing copies of the flight that each correspond to a choice of possible departure time within the discretised time window, along with corresponding connection arcs. Additionally, extra decision variables are introduced, each with an extra index representing one of the possible departure times for a particular flight.

By allowing greater flexibility in departure time, the number of possible connections is increased, allowing for a potentially more profitable solution via a reduction in the number of aircraft or crew required to cover all flights, or perhaps a potentially more operationally robust solution that performs well under different operational scenarios. It should be noted however, that it is important to choose the size of the time window carefully so as not to shift the departure time too significantly and potentially mis-represent the demand for a particular flight; thus losing revenue from under-utilisation or passenger spill. For example, an 8:00am flight between two cities may have significantly different passenger demand than an 8:40 flight between the same two cities.

Desaulniers *et al.* (1997) [30] introduce time windows on flight departures for the fleet assignment problem. The problem is modelled as a multi-commodity flow problem in which extra time variables are introduced. The authors solve this problem using branch-and-bound and column generation in which "the column generator is a time constrained shortest path problem". Rexing *et al.* (2000) [77] incorporate time windows within the fleet assignment problem, and the time windows are discretised into 5 (and 1) minute intervals.

Klabjan *et al.* (2002) [56] address the problem of airline crew scheduling, incorporating time windows to allow more flexibility in the crew-pairing solution. Lan *et al.* (2006) [57] attempt to incorporate robustness into the schedule by estimating delays for each flight leg and minimising expected delay. The authors also use time windows (referred to as re-timing) to address the issue of reducing missed connections for passengers. As in the above, they introduce flight copies, and estimate the number of disrupted passengers for each possible connection using a connection based model, leaving the fleeting and routing solutions unchanged.

In [86], Samardi extends the passenger connection model of Lan (above) and presents an integrated flight departure, re-timing and aircraft routing model that aims to minimise the expected number of misconnecting passengers. This model assists in providing any potential misconnecting passengers with alternative recovery options (such as increasing the number of connection opportunities to assist in re-capturing spilled passengers).

Lohatepanont *et al.* (2004) [63] extend the itinerary-based fleet assignment model (IFAM) of Barnhart *et al.* [12] to determine market service frequency, departure times and fleet assignments simultaneously. The authors make use of a set of flight legs that may be categorised an mandatory, and assess the worth of a particular itinerary and re-adjust flight leg demand if a particular itinerary is removed, or if the schedule is altered.

Belanger *et al.* (2006) [17], present an integrated model for fleet assignment with time windows for which they assume the schedule is periodic. The authors penalise short connections between flights and make use of profit estimations that integrate and capture both departure time and aircraft type; resulting in a potentially more profitable matching of fleet type with expected passenger demand.

2.2.2 Schedule Generation, Fleet Assignment Or Aircraft Routing Combined With Crew Pairing

Aside from fuel costs, crew costs are the second largest expense for an airline. Thus, another area of research growth is that of integrating aspects of the aircraft planning process with the crew pairing stage. Results have indicated that even relatively modest improvements have the potential to improve revenue for an airline.

Barnhart *et al.* (1998b) [15] propose an integrated model for the fleet assignment and crew pairing problems. The authors approximately integrate the fleet assignment with the crew pairing, by including a relaxation of the crew scheduling problem within the fleet assignment model. This relaxation is based on a duty network in Barnhart *et al.* (1998) [14], and ensures that each flight leg is covered by an eligible crew without imposing constraints on the maximum number of duties within a crew pairing, or the maximum time away from the crew base.

Cordeau *et al.* (2001) [29] integrate aircraft routing with crew pairing, using linking constraints to ensure that a crew does not swap aircraft if there is insufficient connection time. The authors use a Benders decomposition, for which the Benders master problem corresponds to the aircraft routing problem and the Benders subproblem corresponds to the crew pairing problem. The solution process then iterates between the master and subproblem allowing information to be passed between the aircraft routing and crew pairing problems. According to the authors, short connections are fixed by the master problem and the subproblem constructs minimum cost crew pairings using only the fixed set of short connections. The authors also make use of a heuristic solution approach (based on branch and bound) in order to obtain integer solutions. Results indicate that the integrated model allows for an improvement of 9.4% in variable crew costs.

Klabjan *et al.* (2002) [56] partially integrate aircraft routing with crew scheduling. Solving the problem sequentially, the authors add plane-count constraints to the crew scheduling model to allow a feasible aircraft routing problem to be obtained. The authors also include time windows to allow more flexibility within the crew scheduling problem.

Mercier *et al.* (2005) [68] improve upon the method proposed in Cordeau (2001) [29], through the introduction of so-called *restricted connections* (i.e. Connections that are longer than the minimum sit-time for crews but smaller than a given threshold). The authors allow restricted connections, but apply a penalty if both legs are covered in sequence by the same aircraft. The authors improve the speed of convergence by reversing the order in which the problems are solved so that the crew pairing is instead solved in the master problem.

Sandhu *et al.* (2007) [85] propose a model that integrates fleet assignment and crew pairing whilst maintaining the possibility of feasible aircraft routings by way of plane-count constraints. The author states that plane-count constraints are sufficient to ensure maintenance feasibility, and thus maintenance constraints are not explicitly modelled. The authors first solve the the crew-pairing problem, determining which crew connections require the crew to remain on the same aircraft (known as a forced turn). The authors state that "a set of forced turns can be extended into a plane-count feasible rotation if and only if the number of planes on the ground at any time imposed by the forces turns does not exceed the plane count constraints". Each station has n activities connected by ground arcs that are used to count the number of planes on the ground for flow balance purposes. Wrap-around arcs are used to allow the schedule to repeat daily.

2.2.3 Three Stages Combined

Mercier *et al.* (2007) [69] propose a model that integrates three of the components in the airline scheduling process, namely, re-timing, aircraft routing and crew pairing. Extending the models of Cordeau *et al.* (2001) [29] and Mercier *et al.* (2005) [68], the authors allow the departure times of flights to be chosen from within a certain time window. As before, the linking constraints ensure that the same schedule is chosen for the aircraft and crew pairings, preventing a crew from swapping aircraft if there is insufficient connection time. The authors propose a compact formulation, using Benders decomposition with dynamic constraint generation to obtain a solution.

An integrated model that incorporates these three stages is also addressed in Papadakos (2007) [74], with the emphasis placed on improving convergence/ solution times (see below).

Improving Convergence

The following three papers address computational techniques that may be used to improve convergence of the integrated models that have been discussed thus far. Papadakos (2007) [74], Haouari (2007) [47] and Mercier (2005) [68]. Papadakos [74] proposes an improved Benders decomposition combined with a column generation that may be accelerated using a termination heuristic to circumvent the so-called "tailing-off effect"[2] commonly experienced when solving with column generation. Papadakos notes that retaining the crew scheduling problem within the Benders subproblem leads to greater efficiency.

2.3 Multi-commodity Flow Problems, Set-partitioning Problems And Column Generation

As mentioned earlier in this Chapter, the aircraft routing problem has been modelled as a multi-commodity flow problem and more recently as a set-partitioning problem. In this section we will outline the multi-commodity flow problem as it relates to the aircraft routing problem and outline the corresponding set-partitioning problem.

[2] A term used to refer to the slow convergence experienced in many real-world column generation problems (See [54], page 349).

Finally, we will conclude by describing the column generation approach used to solve set-partitioning problems.

The multi-commodity flow problem for aircraft routing

Let N denote the number of available aircraft that may be used to cover all flights contained within the set $\mathcal{F}$, where $\mathcal{F} = \{1, 2, \ldots, F\}$. Define the (directed) arc set $\mathcal{A}$, as the set consisting of all allowable flight pairs (i, j), for which (i, j) denotes a direct connection between the flights i and j. For the n^{th} aircraft, each decision variable $x_{ij}^n \in \{0, 1\}$ corresponds to a connection and is defined as follows:

$$x_{ij}^n = \begin{cases} 1, & \text{if flight } i \text{ connects with flight } j, \\ 0, & \text{otherwise.} \end{cases}$$

The cost c_{ij}^n, associated with each decision variable x_{ij}^n, is assumed to be given (or already calculated). Let so denote a "dummy" flight originating from the "source" and t denote a flight whose destination is the "sink"; the optimisation problem may be stated as follows:

$$\begin{aligned} \text{Minimise:} \quad & \sum_{n=1}^{N} \sum_{(i,j)\in\mathcal{A}} c_{ij} x_{ij}^n && && (2.1) \\ \text{Subject to:} \quad & \sum_{n=1}^{N} \sum_{i\in\mathcal{F}} x_{ij}^n = 1 && \forall\, j \in \mathcal{F} \\ & \sum_{j\in\mathcal{F}} x_{so,j}^n \leq 1 && \forall\, n \in \{1, 2 \ldots, N\} \\ & \sum_{i\in\mathcal{F}} x_{i,t}^n \leq 1 && \forall\, n \in \{1, 2, \ldots, N\} \\ & \sum_{i\in\mathcal{F}} x_{ik}^n - \sum_{j\in\mathcal{F}} x_{kj}^n = 0 && \forall\, k \in \mathcal{F}, \forall\, n. \end{aligned}$$

The set-partitioning problem

In many practical problems, it may be more natural to cast the problem in terms of paths or columns, for which the columns correspond to certain tasks, or describe the order of certain processes. For example, in the traditional cutting stock problem which involves the cutting of rolls from paper drums, each entry i of column j corresponds to the number of rolls of width w_i produced by column j, and so the column $(3, 1, 0, \ldots, 0)$ represents 3 rolls of width w_1, 1 roll of width w_2 and no rolls of any other width.

Similarly, for the aircraft routing and crew pairing problems, it is often more natural to describe the aircraft routes and crew pairings in terms of strings or columns in which the entries of the columns $\mathbf{a}_j$ correspond to flights (sorted into topological order). The entries a_{ij}, then take the value 1 if the flight i is included in the string j and 0 otherwise. For example, the column $(1, 0, 1, 0, 0, 1, 1, 0, \ldots, 0)$ represents the flight string $1 \to 3 \to 6 \to 7$. Let I denote the set of all flights and J the set of all feasible aircraft routes (assumed to satisfy flow balance and maintenance requirements). Define the decision variable x_j as follows:

$$x_j = \begin{cases} 1, & \text{route } j \text{ is chosen in the optimal solution,} \\ 0, & \text{otherwise.} \end{cases}$$

where c_j represents the corresponding cost for column j. The set-partitioning problem for aircraft routing may be written as:

$$\begin{aligned} \text{Minimise:} \quad & \sum_{j \in J} c_j x_j \\ \text{Subject to:} \quad & \sum_{j \in J} a_{ij} x_j = 1 \quad \forall\, i \in I \\ & x_j \in \{0, 1\} \quad \forall\, j \in J \end{aligned} \tag{2.2}$$

The approach of simply writing down all the feasible columns may be relatively easy to accomplish in small scale problems, but as was one of the criticisms mentioned earlier, there may be many thousands or even millions of feasible columns corresponding to aircraft routes or crew pairings. Moreover, in a complicated network, finding feasible routes/pairings may itself be extremely difficult. In practice, although there may be thousands of feasible routes/pairings, only a very small selection of those are eventually chosen as optimal. Therefore, what we require is a method for generating *beneficial* feasible routes/pairings, without having to generate all possible feasible columns. This leads us to the idea of *Column Generation*.

Consider the following linear program for which the number of feasible columns $|J|$ is extremely large.:

$$\begin{aligned} \text{Minimise:} \quad & \sum_{j \in J} c_j x_j \\ \text{Subject to:} \quad & \sum_{j \in J} a_{ij} x_j = b_i \quad \forall\, i \in I \\ & x_j \geq 0 \quad \forall\, j \in J \end{aligned} \tag{2.3}$$

The idea behind column generation is to start with an initial subset of columns $J_0 \subset J$ to form a *Restricted Master Problem* (RMP), providing an initial basic feasible solution. The approach is to then solve a subproblem (commonly referred to as a pricing problem), using dual information from the master problem to produce *beneficial* feasible columns (i.e. columns with negative reduced cost) that may then be added into the master problem. The master problem is then re-solved to obtain new dual information. This process is continued in an iterative manner, until no further beneficial columns can be produced.

The subproblem is typically formulated as a path problem (such as a shortest path problem (SSP), resource constrained shortest path problem (RCSPP)), or may consist of several constraints that possess a special structure and are themselves well-solved problems (such as a knapsack problem, or the Travelling Salesman Problem). The objective in the subproblem is to obtain the column with the most negative reduced cost using the dual values π_i for each $i \in I$:

$$\text{Minimise :} \quad c_j - \sum_{i \in I} \pi_i a_{ij} \qquad \forall j \in J$$

The optimal column is then added into master problem, and the cost c_j is added to the objective for this new column. In some problems, the calculation of c_j is straightforward, or indeed, may be identical for all j. For example, if the objective is to minimise the number of aircraft, then c_j may be simply set to 1 for all $j \in J$. However, as we shall see in Chapter 4 in which c_j represents delay propagation cost, the calculation of c_j is significantly more complicated and must be calculated simultaneously with the column. To achieve this we propose a new label setting algorithm capable of calculating both the optimal column and its corresponding delay cost, simultaneously.

CHAPTER

THREE

Capturing Uncertainty: Improving Robustness

In the previous Chapter, we introduced the concept of integration and outlined the manner in which authors have integrated the traditional planning stages. The driving force behind this move toward integration is one of improving solution quality and potentially providing a more profitable and operationally robust solution.

Almost all of the integrated models described in the previous Chapter are deterministic. That is, there is no uncertainty in any of the parameters as it is assumed that flights will depart and arrive as planned. However, experience demonstrates that unforeseen events such as weather disruptions, delayed connecting resources (such as passengers and crew) and the interaction of flights in an increasingly complicated world network, mean that deterministic models have the tendency to become brittle in real-world operations. They are often referred to as "over-optimised" and have a tendency to collapse rather quickly, as delays propagate rapidly within such networks.

Consequently, in these times of increasing customer dissatisfaction with delays, airlines wish to construct solutions that possess operational robustness whilst continuing to deliver profitability. In an attempt to address this problem, researchers have stated to shift their focus to one of maximising profit whilst accounting for real-world uncertainty.

3.1 Stochastic Approaches

Two-stage stochastic optimisation problems have proven popular with researchers interested in both long-term and short-term planning, and increasingly airline recovery. In such models, the airline initially makes a (first-stage) decision based on the information currently known with certainty (such as number of available aircraft, expected arrival and departure time of flights, expected passenger demand etc.). Once the schedule has been implemented, the actual values associated with such parameters are realised and are used as input to implement a (second-stage) or recourse (recovery) decision - in which the objective is to determine the 'optimal' course of action once these real-world values are known.

This type of modelling is particularly useful in the context of fleet assignment for which fleet allocation decisions are required to be made several months prior to the day of operations, when passenger demand forecasts for each flight leg may only be approximately known. A few days prior to the day of operations, once an airline has a reasonably accurate estimate of demand, they may solve the second stage recourse decision problem to obtain an improved solution that better matches the fleet type with the passenger demand. Such a analysis was performed by Sherali (2008) [91]. The author implemented a two-stage stochastic program to adjust fleet assignment in the weeks leading up to the schedule's implementation.

Authors such as Yen and Birge (2006) [104] have adapted the two-stage stochastic programming model to assist in solving the crew pairing problem. In their model, the authors use the recourse model to reflect the long-term and short-term interactions. The authors capture interactions between crew schedules and use this to identify robust solutions that are capable of withstanding flight disruptions. To solve their problem, they propose a new branching algorithm capable of branching on multiple variables simultaneously.

Rosenberger *et al.* (2002) [79] use a discrete event semi-Markov process to stochastically model the daily operations of a domestic airline in a bid to evaluate crew schedules and recovery plans that are subject to uncertainty.

3.2 Robust Optimisation

Robust Optimisation is a branch of optimisation that seeks to achieve an optimal solution for a problem in which there is an element of data/variable uncertainty or for which problem information is not known *a priori*. Many of the results are based on the assumption that the magnitude by which the parameters (or a subset of the parameters) are allowed to vary lies within known *uncertainty sets*. The objective is often to minimise the cost of the worst case scenario.

The area of Robust Optimisation has expanded very rapidly within the last few years, with a number of authors such as Beyer (2007) [22], Janak (2007) [51], Bertsimas (2007) [20] providing optimality conditions for optimisation problems in which the data, resource vectors and cost coefficients are subject to uncertainty.

Marla *et al.* (2011) [67] propose two extensions of existing robust optimisation models; namely the extreme-value based model of Bertsimas *et al.* (2003) [21] and the probabilistic, chance-constrained model of Charnes *et al.* (1959) [24]. The authors apply their extensions to the airline routing problem with their models having the advantage that they can be solved in a single iteration with run-times comparable to that of the basic models. The authors observe via simulation that the extended extreme-value and probabilistic models have the potential to consistently lead to the generation of more robust solutions; albeit with different performance characteristics. The authors observed that the extreme-event model had the characteristic of being driven by extreme values of delay and thus ignored probabilistic information in some cases. This led to a large disparity in the performance of optimal solutions. The probabilistic approach was contrasted with the tailored approach of Lan *et al.* (2006) [57]; with both of these approaches performing in a similar fashion on defined robustness measures. The authors conclude that the probabilistic approach and the tailored approach capture more relevant information and focus on more likely delay events, leading to solutions that better satisfy the robustness metrics of interest.

Yu *et al.* (2000) [105] reformulate a stochastic management problem as an efficient robust optimisation model. The authors propose a method to transform the robust model into a linear program requiring only $n + m$ variables, where n represents the number of scenarios and m represents the number of total control constraints; smaller than that of existing methods that require $2n+2m$ such variables. The authors apply their new approach to various logistical management problems, including an aircraft routing problem for the U.S. Military.

List *et al.* (2003) [61] propose a robust optimisation approach to fleet planning under uncertainty for a general transportation problem. The authors propose a formulation that focuses on two different sources of uncertainty; namely, the future demands to be served by the vehicle fleet and the productivity of individual vehicles. Through the use of robust optimisation, the authors are able to examine the tradeoff between the level of fleet investment and the level of risk in the solution.

Qin *et al.* (2010) [75] propose a robust optimisation approach to the capacitated hub and spoke airline network design problem. The authors seek to control the deviation between solutions under different scenarios to within a pre-set range. The objective is to determine the optimal connection between different non-hub airports to the hub in the case of uncertain demand and cost. The authors make use of an ant colony algorithm to solve their model on different numerical examples.

Birbil *et al.* (2009) [23] propose a robust version of the standard and dynamic single-leg seat allocation models. Their model seeks to improve upon existing models that inaccurately estimate the probability distributions for the total demand for the different fare classes. The authors demonstrate that the new approach results in solutions with less variability in realised revenue.

3.3 Robust Methods

One of the potential difficulties involved with using the robust optimisation techniques outlined in the previous section, is that it is first necessary to make assumptions on the level of uncertainty of each variable or parameter. If the choice of uncertainty set or type of distribution is not made appropriately, the solution quality may vary significantly when applied to the real world. In an airline network, delays may interact in a multitude of ways between aircraft, crew and passengers and consequently it may be very difficult to determine precise delay information for certain parameters *a priori*. To sidestep this problem, a number of authors have turned to robust methods in an attempt to improve the way in which certain parts of the airline scheduling problem are modelled.

One of the key problems with trying to develop a robust solution is that there is no general consensus or systematic way to define robustness (Lan *et al.* (2006) [57]). A very natural definition of robustness is offered by Ahmed (2008) [6], which we will also take as our definition:

Definition 3.3.1 (Robustness [6]). *Robustness is a fast recovery from disruptions with dampening; thereby limiting the growth of disruptions.*

Despite the lack of consensus on the definition of robustness, there are a variety of ways in which to improve the robustness of a solution. Primarily, the focus of robust methods and within the literature itself is one of embedding robustness into the model in the planning stage. One way is to take advantage of particular network structure (eg. Hub connectivity, aircraft swap opportunities) and secondly to include uncertainty in the initial model parameters to produce a solution that performs well on average; either with respect to delay propagation, or another key performance indicator (or combination thereof). Typical key performance indicators used by airlines are [6]: percentage of flights cancelled, percentage of late arrivals, on-time performance distribution of delays, number of delayed passengers/average passenger delay and average delay per flight.

We now outline a few different approaches for including robustness within the planning stage.

3.3.1 Aircraft-Swap Opportunities: Incorporating Fault-Tolerant Recovery Paths

Ageeva (2000) [3] proposes a model that assists in providing flexibility for airline planners in the event of schedule recovery. By identifying and maximising the number of times different aircraft routes 'meet' more than once, the author's model provides opportunities for the aircraft to swap routes and return to their original route the next time the routes meet. This may be of assistance if the passenger demand on a particular leg (or sequence of legs) changes, and may provide alternative routes that both arrive at a particular destination.

3.3.2 Robust Fleet Assignment: Hub Isolation And Short Cycles

Rosenberger, Johnson and Nemhauser (2001) [78] propose a robust fleet assignment and aircraft routing model developed to produce a solution that consists of a large number of short cycles with low hub connectivity. The authors found that decreasing hub connectivity assisted in producing larger numbers of short cycles. This may be beneficial, as when airlines cancel flights they usually need to cancel a cycle (to maintain flow balance). In this case, the cancellation of shorter cycles has less of an impact on passengers.

3.3.3 Robust Crew Scheduling

Simulations

Schaefer et al. (2001) [87] propose a heuristic capable of generating robust crew schedules. The authors achieve this by initially calculating an approximate expected cost for each crew pairing, making assumptions on the possible crew recovery procedure and utilising their simulation tool named SimAir. The authors use the approximate crew costs to solve the deterministic crew pairing problem in order to find solutions that perform well with respect to both planned costs and operational costs.

Move-Up Crews

Chebalov and Klabjan (2002) [88] also propose a model to improve the robustness of crew pairings. To achieve this, the authors seek to maximise the number of opportunities for the crews to be swapped (moved-up) in operations. The authors use Lagrangian relaxation to solve the model, but do not provide a measure of robustness for their solution.

3.3.4 Degradable Airline Scheduling

Kang [52] proposes a methodology in which each of the required flights is partitioned in such a way that each flight is assigned to a unique 'layer'. These layers correspond to different levels of 'service' from the airline, and are given different (recovery) priorities in the event of a disruption. To compensate for a different level of service, passengers in the lowest priority layer pay the least amount for their ticket, whilst those in the highest priority layer pay a premium for their level of service. By separating flights into different layers, this model reduces the effect of propagation delay throughout the network, as it is assumed that these layers do not interact with one another.

The author presents three different ways to incorporate degradability into the scheduling process. The first is to include degradability between the flight scheduling and fleet assignment (a degradable schedule partitioning model), the second within the fleet assignment (degradable fleet assignment model) and the third and main focus is within the aircraft routing (degradable aircraft routing model).

3.3.5 Robust Aircraft Routing

Lan *et al.* (2006) [57] develop a robust aircraft routing model to minimise the expected propagated delay along aircraft routes. The authors assume that the delays along each leg are additive, and do not consider interactions between aircraft and crew or delay effects from other parts of the network (eg. late running passengers from connecting flights are not considered). The authors use an approximate delay distribution to model propagation delay along each flight string and use a branch-and-bound technique to used to solve their MIP. They authors also propose a connection-based re-timing model to minimise the number of misconnecting passengers (as mentioned previously).

3.3.6 Imposing Station Purity

Smith *et al.* (2006) [93] propose a model in which the number of different fleet types that are allowed to serve each airport is limited - this is referred to by the authors as *imposing station purity*. The authors demonstrate that imposing station purity within the fleet assignment model leads to solutions that are more robust for crew planning, maintenance planning and for operations in general. One potential disadvantage of such an approach is the potential for excessive computational time required to solve such a model. To address this problem, the authors develop a "station decomposition" solution approach that takes advantage of network structure.

The authors also propose a primal-dual method to improve solution quality and model efficiency.

3.3.7 Bi-criteria Optimisation

Ehrgott *et al.* (2002) [36] and Bassy *et al.* (2007) [95] seek the multiple objectives of minimising cost whilst maximising robustness for crew scheduling. This multi-objective is addressed within the bi-criteria optimisation framework. The authors develop a bi-criteria optimisation framework to generate Pareto optimal schedules. Results indicate that for a small increase in cost, one may gain a significant increase in robustness (as a result of an increase in overnights, leading to a reduction in the number of aircraft changes).

3.3.8 Re-booking, Revenue Management And Stochastic Passenger Demand

As mentioned at the beginning of this Chapter, two-stage stochastic programming models have proven popular with researchers aiming to improve the way in which fleet types are assigned to individual flights legs when passenger demand may be uncertain. In this section, we outline a few additional improvements that can be made to such a model.

According to Barnhart (2002) [12] and Jacobs (2008) [50], there are several disadvantages of traditional fleet assignment models. Firstly, spill and recapture are ignored or modelled only approximately and estimates of recaptured revenue are achieved without knowledge of capacity or passenger flow on the network. Secondly, such models consider only aggregate demand and average fares for different fare classes, leading to an inaccurate representation of estimated spills and spill costs. Finally, the majority of traditional fleet assignment models assume that passenger demand is static over time.

Several approaches to incorporating improved revenue management aspects into the fleet assignment model have been investigated over the past 10 years. In Barnhart et al. (2002) [12] the authors develop a new fleet assignment model, termed an

Itinerary-based Fleet Assignment Model (IFAM) capable of capturing network effects and in particular, accurately estimating passenger spill and re-capture for revenue management purposes. The authors achieve this by way of a proposed "passenger mix model" used to determine optimal traffic and revenue when given as input, a schedule with known flight capabilities and passenger demands with known fare. According to the authors, the passenger mix model allows for customer choice modelling and recapture to be included. As there may exist many thousands of possible passenger itineraries, the authors reduce the problem complexity by making use of *key paths*, defined as the originally desired itinerary for each passenger; thus, alternative itineraries are only necessary in the case where passengers are spilled from their desired itinerary.

In Jacobs (2008) [50] the author presents a model that addresses both network effects and the stochastic nature of demand. The authors use Benders decomposition to integrate the fleet assignment model with the origin and destination fleet assignment revenue management model. The author refers to this proposed approach as "O and D fleet assignment". The model takes as input a given fleet assignment solution and captures the O and D revenue management aspect within a subproblem. A revenue function for the entire network is approximated in the fleet assignment master problem using a series of Benders cuts, with each cut improving the accuracy of the revenue approximation in the master problem. Once an approximation has been obtained within the desired accuracy, the model is then solved as a MIP. The author claims that this approach is appealing as it addresses both passenger flows within the network and passenger demand uncertainty. Additionally, it also provides a method of incorporating the passenger mix optimisation model used for revenue management directly into the fleet assignment process.

A number of authors have addressed the problem of including passenger uncertainty and improving revenue management techniques for fleet assignment. The reader is referred to the following papers for a more comprehensive summary. Sherali *et al.* (2008) [91] in which the authors construct a two-stage stochastic model for fleet assignment. Dumas *et al.* (2008) [33] proposes a new passenger flow model capable of modelling spill and recapture between itineraries and accounts for the inter-dependencies between legs and the effect of this inter-dependency on revenue. The authors incorporate this passenger flow model into the model proposed in Dumas *et al.* (2008) [32], in which they seek to improve the objective of the fleet assignment model. To achieve this, the authors utilise their passenger flow model and incorporate it within an iterative model. The objective function is iteratively improved via a process of alternately generating fleet assignments and analysing their profit potential via the (modified) passenger flow model.

Sandhu *et al.* (2006) [84] proposes a new fleet assignment model that incorporates both passenger and cargo revenue. According to the authors, in the last few years, a significant decrease in the number of business class passengers has resulted in a decline in passenger revenue. This decline has prompted airlines to consider revenue derived from the transportation of cargo to compensate for these losses. The authors present a fleet assignment model that captures both cargo and passenger revenue, making use of a cargo mix bid-price model that considers demand, weight and volume constraints and assigns optimal cargo allocations to a given fleeting solution in order to maximise cargo revenue. To account for passenger revenue, the authors utilise a passenger mix bid-price model, commonly used in revenue management systems to allocate passengers to each itinerary whilst maximising passenger revenue.

Barnhart *et al.* (2009) [11] seek to improve the revenue function for fleet assignment problems stating that the standard assumptions placed on the standard revenue function are too simplistic. The authors propose a new sub-network fleet assignment model (SFAM) modelled as a mixed integer program for which composite decision variables represent the simultaneous assignment of fleet types to subnetworks. The authors construct local subnetwork fare structures and for each subnetwork, the local fare of a given itinerary is computed as the sum of the fares allocated to the constrained flight legs (where demand exceeds capacity) in that subnetwork. Furthermore, it provides a good approximation to the Itinerary Based Fleet Assignment model with the additional benefit of providing tighter LP bounds.

3.4 Integrated And Robust Methods

Up until fairly recently, the majority of authors have attempted to include robustness within one aspect of the airline planning stages. In the last few years (since 2007) there has been significant growth in the area of integrated robustness; that is, including robustness within a model that integrates two or more stages of the planning problem. We briefly mention a few of these below.

3.4.1 Iterative Airline Scheduling

Weide *et al.* (2007) [98] propose an integrated aircraft routing and crew pairing model for which the solution is obtained via an iterative process that passes information between the aircraft routing and crew pairing problems. The authors note that solving the aircraft routing and crew pairing problems individually may result in a sub-optimal solution as the decisions made in the preceding stages may restrict the number of feasible choices in subsequent stages.

To solve their model, the authors propose a non-robustness measure to keep track of the number of restricted aircraft changes. The authors wish to reduce the number of restricted aircraft changes (for which crew change aircraft over a restricted connection) in an attempt to minimise delay propagation resulting from insufficient connection time. The objective is to keep the aircraft and crew together for as long as possible over restricted connections, so as to minimise this effect. The authors achieve this in their model by seeking to maximise the number of restricted connections contained in aircraft solution that are operated in the current crew pairing solution, then seeking to minimise the number of restricted aircraft *changes*. This process continues iteratively, while increasing the crew penalty at each iteration and continues until the non-robustness measure cannot be improved further.

The advantage of this approach is that the computational complexity is not increased as for the other integrated models (as the original set-partitioning formulations of both problems remain unchanged) - allowing the problem to be solved efficiently. The authors extend this model to account for multiple crew groups.

3.4.2 Delay Propagation Calculation With Propagation Trees

AhmadBeygi *et al.* (2008) [4] proposed both a single-layer model and a multi-layer model to minimise delay propagation by re-timing flights in such a way that the slack present in the network is re-allocated to where it is required most. The single layer model considers the total delay propagated to a particular flight to be dependent only on the delay accumulated on the preceding leg, whereas the multi-layer model takes into account the delay from all other legs that precede a particular flight. This is accomplished by way of a propagation tree for which the nodes represent flight legs and the branches, the flight connections. There are a number of disadvantages associated with this model, namely that it not only overestimates (double counts) delay in some cases, but underestimates the delay in others.

3.4.3 Integrated Aircraft Routing, Crew Pairing And Tail Assignment

Ruther *et al.* (2011) [82] outline an integrated model for aircraft routing, crew pairing and tail number assignment. The authors acknowledge that data available in the traditional planning stages may be inaccurate, leading to frequent re-scheduling of resources. Therefore, they wish to delay as many decisions for as long as possible until accurate data becomes available. Using their model they are able to solve the integrated model four days prior to the day of operations, allowing them to remove the uncertainty that is usually present in the standard long-term planning process. The authors claim that their model leads to less re-scheduling of resources, specific routes for each aircraft and pairings for each crew with a planning horizon of one week.

3.4.4 Minimising Delay Propagation In An Integrated Aircraft Routing And Crew Pairing Framework

In the next Chapter we outline in detail the contributions of this Thesis. Our aim in this work is to contribute to the emerging field of Integrated Robustness. Specifically, we seek to minimise delay propagation within an integrated aircraft routing and crew pairing framework. Secondly, we propose a heuristic re-timing algorithm that may be used in conjunction with the integrated aircraft routing and crew pairing problem to re-schedule aircraft and crew (within a limited time window) to further minimise delay propagation costs for the network. Finally, we extend our integrated aircraft routing and crew pairing model to include re-timing, so that all three decisions can be made simultaneously.

CHAPTER

FOUR

Minimising Propagated Delay In An Integrated Aircraft Routing And Crew Pairing Framework

In Chapter 1 we introduced the airline scheduling problem, highlighted its complexity and the subsequent motivation of Airline Schedule Planners to sequentially decompose the problem into manageable stages to retain any degree of tractability. In Chapter 2, we briefly outlined and explained the four broad stages of the the airline scheduling problem, traditionally decomposed sequentially in the following manner (eg. schedule generation, fleet assignment, aircraft routing, and crew pairing), with the decisions from one stage imposed upon the decision making process in subsequent stages. Whilst this approach greatly simplifies the solution process, it unfortunately fails to capture the many dependencies between the various stages, most notably between those of aircraft routing and crew pairing, and how these dependencies affect the propagation of delays through the flight network. As delays are commonly transferred between late running aircraft and crew, it is important that aircraft routing and crew pairing decisions are made together. The propagated delay may then be accurately estimated to minimise the overall propagated delay for the network and produce a robust solution for both aircraft and crew. In this Chapter we introduce a new approach to accurately calculate and minimise the cost of propagated delay, in a framework that integrates aircraft routing and crew pairing. Most of the material in this Chapter is joint work with Gary Froyland and Richard Wu at the University of New South Wales and appeared in the paper Dunbar *et al.* (2012) [35].

4.1 Motivation And Key Contributions

The airline scheduling problem involves the construction of timetables for an airline's major resources, namely aircraft and crew. Traditionally, this has been undertaken with a view towards maximising an airline's overall profit, often with limited consideration given to the stability of such a schedule, or indeed its operational robustness. Such an approach has a tendency to generate schedules that are highly brittle, performing poorly in practice as delays propagate rapidly throughout the network. The Bureau of Transportation Statistics [76] states that in 2009, approximately 23% of flight legs operated by a major US airline were delayed – with late arrivals and cancellations combined accounting for more than 7.5% of this delay. In recent years, this has resulted in an ever increasing discrepancy between planned costs and realised operational costs. As aircraft networks continue to grow, this trend is set to continue with AhmadBeygi *et al.* [4] reporting that in 2006, it was estimated that the US airline industry experienced a total of 116.5 million minutes of delay; translating into a $7.7 billion increase in operating costs. Such large discrepancies have prompted airline schedule planners to shift their focus from maximising profit to maximising expected profits under uncertainty, by including various types of costs arising from unplanned events.

Lan *et al.* [57] develop a robust aircraft routing model to minimise the expected propagated delay along aircraft routes. They use an approximate delay distribution to model the delay propagation along each string and use a branch and bound technique to solve their MIP. Lan *et al.* calculate propagated delay along individual strings when determining costs for the restricted master problem, but omit considerations of delay when solving the subproblem. The effect of connecting resources (such as crew and passengers) are not considered. Instead of estimating delay propagation, Wu [100] used a simulation model to calculate random ground operational delays and airborne delays in an airline network. Wu [100, 103] shows that delays are inherent in airline operations due to stochastic delay causes, e.g. passenger connections and late baggage loading. By adjusting flight times without changing aircraft routing, Wu [103] revealed that significant delay (cost) savings can be achieved via robust scheduling. Weide *et al.* [98] propose an integrated aircraft routing model for which the solution is obtained iteratively. The authors propose a non-robustness measure and initially solve the crew pairing problem without taking into account an aircraft routing solution. Their model then seeks to maximise the number of restricted connections contained in the aircraft solution that are also operated in the current crew pairing solution. Once this solution has been obtained, they minimise the number of restricted aircraft changes. This process continues iteratively, increasing the crew penalty at each iteration until the non-robustness measure cannot be improved further. The advantage of this approach is that the computational complexity is not increased as in other integrated models. AhmadBeygi *et al.* [4] make use of a propagation tree to minimise delay propagation due to flights and crew pairs in an existing routing and crew pairing solution, by re-timing flights so that the slack present in the network is re-allocated to where it is required most. Their approach is limited to re-timing and both under and overestimates the delay propagation in certain cases.

Key Contributions Of Our Model

Our aim is to improve upon the following shortcomings of AhmadBeygi *et al.* [4], Lan *et al.* [57] and Weide *et al.* [98]. Firstly, while Lan *et al.* correctly calculate propagated delay of aircraft strings in their master problem, the *selection* of these new columns is carried out more crudely: new columns are generated within the subproblem *without* considering the delay cost of the new column. The authors only make use of the dual variables from the master problem when determining the minimal cost column. Once a column has been generated they then calculate the propagated delay cost along the string and decide whether to add it to the restricted master problem. Furthermore, they ignore the effect of connecting resources such as crew and passengers. Secondly, while AhmadBeygi *et al.* [4] consider (in a re-timing setting) the combined delay effects from crew and from aircraft, their approach imperfectly calculates how delays are propagated, resulting in possible under or overestimates of the true propagated delay. Their improvements are also limited to those achievable by re-timing.

Finally, Weide *et al.* [98] treat the interactions of crew and aircraft in an iterative fashion, optimising a robustness measure, which is an *indirect* means of assessing the *true cost* due to total propagated delays of aircraft and crew. The model in [98] attempts to keep aircraft and crew together over restricted connections, to try to minimise the number of restricted aircraft changes. Although [98] takes into account the connection time, penalising shorter restricted aircraft changes more severely, the Weide *et al.* model penalties are time-of-day independent, independent of historical information for the network, and do not quantitatively assess the propagated delay from the interactive connectivity of the routing and crewing networks. For example, there may be relatively predictable large primary delays over certain connections or at certain times of the day, or the effects of delays for some connections are much worse in a propagated sense than for other connections, depending on the interactive network topology. Our approach explicitly utilises time-of-day historical primary delays and explicitly calculates and minimises the downstream effect of delay in the combined routing and crewing network. Solutions developed from our approach may (for example) mismatch aircraft and crew on a restricted connection if later connections have ample slack to absorb delays. This mismatch may free up the possibility to match crew and aircraft on a critical connection that has tight connections further downstream. We provide a quantitative comparison of our approach and the approach of [98] in Section 4.5.

The key ingredients of our approach are (i) the accurate calculation of the *combined effects* of propagation of delay along aircraft routing strings *and* crew pairing strings and (ii) the use of this information for both the calculation of the *cost* of columns and the dynamic *selection* of optimal columns. In Sections 4.2 and 4.3 we briefly outline standard column generation approaches to finding minimum cost aircraft routings and crew pairings, respectively. In Section 4.4 we describe our approach for accurately calculating the propagated delay of routing and crewing strings and in section 4.4.2 we describe the setup of our pricing problems. Sections 4.4.3 and 4.4.4 describe our numerical approaches for solving the master and pricing problems, respectively. Computational results are presented in Section 4.5 and we conclude with suggestions for future work in Section 4.6.

In this Section we describe our formulation for the integrated aircraft routing and crew pairing problem; the objective is to minimise the total cost associated with propagated delay. We first outline the mathematical formulation of the aircraft routing and crew pairing problems individually and then discuss estimation of propagated delay and the corresponding pricing problem. We concentrate solely on costs due to delays with the understanding that in practice, the additional costs due to unplanned delays can form part of an overall model of cost for the airline. We thus view our proposed methodology as a potential add-on to existing connection-based optimisation models to better reflect planned costs under uncertainty.

4.2 The Aircraft Routing Problem

The aircraft routing problem is performed separately for each specific fleet type. We seek a minimal cost assignment of aircraft to flights where each flight is covered exactly once by exactly one aircraft. The costs will represent the cost of the total delay incurred by the aircraft over a 24 hour period.

In the following routing model, we calculate a one day schedule where each aircraft begins and ends its day at a maintenance base. Maintenance feasible routings are represented as columns of an $m \times n_R$ binary matrix A^R, where m is the number of flights and n_R is the total number of feasible routings. The $(i,j)^{th}$ element of A^R takes the value 1 if flight i is contained in routing j and 0 otherwise. In practice there may be an extremely large number of feasible columns, so column generation is used to generate only the beneficial columns. For each flight (node) we assign a dollar cost per unit of delay arriving at that flight, and the cost c_j^R of column j is the sum of the costs of the delays along string j. The decision variable x_j^R takes the value 1 if routing j is included in the optimal solution and 0 otherwise. There is also an upper bound on the number of aircraft N. Thus we may state the aircraft routing problem as follows:

$$\begin{aligned} \text{minimise:} \quad & (\mathbf{c}^R)^T \mathbf{x}^R \\ \text{Subject to:} \quad & A^R \mathbf{x}^R = \mathbf{e} \\ & \sum_{i=1}^{n_R} x_i^R \leq N \\ & \mathbf{x}^R \in \{0,1\}^{n_R} \end{aligned} \tag{4.1}$$

where $\mathbf{e}$ is an m-dimensional column vector of 1s.

4.3 The Crew Pairing Problem

The crew pairing problem is also performed separately for each fleet type, as crew typically may only fly on board a specific fleet. The objective of crew pairing is to find a minimal cost assignment of crew to flights. As in the routing problem, the costs will represent the dollar cost of the total propagated delay incurred by the crew. The airline from which we source our data uses both pay-and-credits (for cabin crew) and flying hours (for pilots) as crew payment bases. For the purposes of this paper we use the flying-hour based crew costing model, which simplifies our crew costing model. A feasible set of crew pairings must satisfy union regulations (such as the 8-in-24 rule) and ensure each flight is covered exactly once by exactly one crew group (cabin crew, pilots etc.). In the following crew pairing model, we assume a one day schedule where the crew are restricted to flying a total of less than eight hours in each pairing (8-in-24 rule) and ensure that at the end of its duty, each crew pairing returns to the crew base at which it started. This modified 8-in-24 assumption for a one-day schedule simplifies our crew pairing model. One could relax this assumption and expand the schedule to one week during implementation. As for the aircraft routing problem, the pairings may be represented as columns of an $m \times n_P$ matrix A^P, where m is the number of flights and n_P is the total number of feasible crew pairings. We use column generation to generate the most beneficial columns. The element c_j^P denotes the cost of column j and is defined as in the aircraft routing problem above. Thus, we may state the crew pairing problem as follows:

$$\begin{aligned} \text{minimise:} \quad & (\mathbf{c}^P)^T \mathbf{x}^P \\ \text{Subject to:} \quad & A^P \mathbf{x}^P = \mathbf{e} \\ & \sum_{i=1}^{n_P} x_i^P \leq M \\ & \mathbf{x}^P \in \{0,1\}^{n_P} \end{aligned} \tag{4.2}$$

where $\mathbf{e}$ is an m-dimensional column vector of 1s. There is typically no upper bound placed on the number of crews in the standard crew pairing problem and we therefore do not include this constraint in our model.

4.4 Estimation Of Propagated Delay

The calculation of total propagated delay along an aircraft string in an aircraft connection network or along a crew string in a crew connection network is non-trivial. The model of delay propagation we use for individual strings is based on a simplified version of Wu [100, 103] and is similar to the calculation of delay cost in individual strings used by Lan *et al.* We outline our modelling approach for calculation of propagated delay in the isolated routing and crewing networks before describing how to calculate propagated delay in a combined network in the next subsection.

Let $G = (\mathcal{N}, \mathcal{A})$ be a directed acyclic graph with a single source node *so*, and a single terminal node t. The source and terminal nodes are dummy nodes that link to both the morning and evening flights, respectively. In this graph, nodes correspond to flights and arcs correspond to possible feasible connections between flight nodes. For simplicity of exposition, we use the same connection network for both aircraft and crews, although one may use different arc sets if necessary.

Each connection $(i, j) \in \mathcal{A}$, will have associated with it two *primary delays*. The primary delay for aircraft connection (i, j) is denoted p^R_{ij} and is the sum of the expected en-route delay for flight i (estimated from historical data), and expected primary delays during aircraft turnaround operations, such as passenger connection delay, and ground handling delay. Note $p^R_{jt} = 0$ for all $(j, t) \in \mathcal{A}$. The primary delay for crew connection (i, j) is denoted p^P_{ij} and is the the sum of the expected en-route delay for flight i, and other crew related expected primary delays during aircraft turnaround time, such as late crew boarding and crewing procedures. En-route delays and turnaround delays occur for a variety of reasons such as weather conditions, air traffic flow management, passenger delays, equipment failure, and so on. These delays and their causes are documented by airlines by using the IATA delay coding system or its in-house variant [48]. Note $p^P_{jt} = 0$ for all $(j, t) \in \mathcal{A}$.

The flight schedule is the starting point for calculating *slack* for individual connections. The slack s_{ij} for a connection (i, j) is the difference between the scheduled arrival time of flight i and the scheduled departure time of flight j, minus the mean turn-around time for the relevant aircraft type under the specific ground handling procedure of the airline. The value of the mean turn-around time is determined by the standard aircraft ground operating procedures of a specific fleet by an airline. Airlines design aircraft turn-around time based on the mean turn-around time and buffer allowance. For simplicity we have used the same turn-around time for all connections, as all aircraft belong to the same fleet and operate on a domestic network. It is however, straightforward to specify specific turn-around times for individual connections should this be required for an alternative network. All slacks $s_{so,i} = 0, (so, i) \in \mathcal{A}$, and $s_{jt} = 0, (j, t) \in \mathcal{A}$.

We now come to the *propagated delay* at node i, denoted d_i. We fix the initial delay at the source node $d_{so} = 0$ and inductively apply the formulae below to calculate propagated delay along a path in the aircraft connection network:

$$d_j^R = \max\left\{d_i^R - (s_{ij} - p_{ij}^R), 0\right\}, \quad j \neq so, \tag{4.3}$$

and in the crew connection network:

$$d_j^P = \max\left\{d_i^P - (s_{ij} - p_{ij}^P), 0\right\}, \quad j \neq so. \tag{4.4}$$

For computational tractability we assume (as in Lan *et al.* [57]) that the primary delay is independent of the propagated delay.

4.4.1 Estimation Of Combined Propagated Delay

In the previous section we saw how to calculate propagated delay along a path from the source node *so*. The delays along an aircraft string were only affected by aircraft delays in that string and not by delays due to connecting crew. Similarly, delays along a crew pairing were only affected by crew delays in that string and not delays due to connecting aircraft.

Firstly, we consider the effects of crew delays on the aircraft connection network. We assume that we are presented with a feasible set of crew strings and that propagated delays due to the crew have been calculated (to initialise the procedure, we will use (4.4) to calculate the d_i^P, $i \in \mathcal{N}$). To calculate the propagated delay along an aircraft string, *taking into account propagated delays from crew* we inductively apply:

$$d_j^R = \max\left\{d_i^R - (s_{ij} - p_{ij}^R), d_k^P - (s_{kj} - p_{kj}^P), 0\right\}, \quad j \neq so, \tag{4.5}$$

where the connection (i, j) is part of the aircraft string and the connection (k, j) is part of the crew string that includes flight j.

Thus, if flight j uses the same aircraft as flight i and the same crew as flight k, the delay propagated to flight j is the maximum of the delays of the aircraft and crew (or zero, if both delays are negative); see Figure 4.1 for an example.

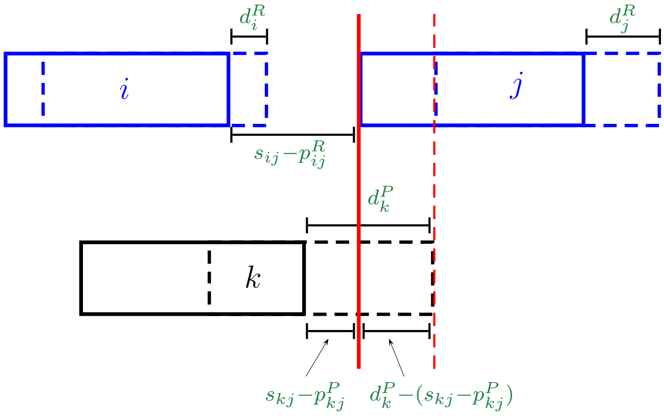


Figure 4.1: Illustration of the requirement of the maximum in equation (4.5). Aircraft and crew are denoted by blue and black boxes respectively. The bold red line denotes the scheduled departure time for flight j. Dashed lines represent the amount by which the aircraft and crew are delayed. Notice that although flight i is delayed, there is enough slack between flights i and j to absorb this delay. However, there is not enough slack between flights k and j for the crew on flight k to arrive in time for flight j. Thus, $d_k^P - (s_{kj} - p_{kj}^P) > 0$ and $d_j^R > 0$.

Secondly, we consider the effects of aircraft delays on the crew connection network. We assume that we are presented with a feasible set of aircraft strings and that propagated delays due to the aircraft have been calculated (to initialise the procedure, we will use (4.3) to calculate the d_i^R, $i \in \mathcal{N}$). As above, to calculate the propagated delay along a crew string, *taking into account propagated delays from aircraft* we inductively apply:

$$d_j^P = \max\left\{d_i^P - (s_{ij} - p_{ij}^P), d_k^R - (s_{kj} - p_{kj}^R), 0\right\}, \quad j \neq so, \tag{4.6}$$

where the connection (i, j) is part of the crew string and the connection (k, j) is part of the aircraft string that includes flight j.

4.4.2 The Pricing Problems

We now describe the pricing problems for the routing and crewing master problems. When solving the routing subproblem the propagated routing delays d_i^R, $i \in \mathcal{N}$ will be calculated dynamically as part of the subproblem, using fixed pre-calculated propagated crewing delays d_i^P, $i \in \mathcal{N}$. When solving the crewing subproblem, the reverse is true; the propagated crewing delays d_i^P are dynamically calculated and the crewing delays d_i^R are pre-calculated and fixed.

Each node i possesses a weight $-w_i$, corresponding to the dual multiplier for constraint i in the master problem; we denote by $-w_i^R$ the weights from the routing master and by $-w_i^P$ the weights from the pairing master. We assume that for every unit of time an aircraft (resp. crew) is late at node i, a dollar cost $a_i^R > 0$ (resp. $a_i^P > 0$) is incurred. These costs are combinations of costs associated with excess fuel consumption, overtime pay for crew, and costs associated with re-accommodating misconnecting passengers [4].

Let $\pi(i)$ denote an ordered collection of nodes in the aircraft path π, truncated so that node i is the final node in the list. Additionally, we use the notation $\pi^-(i)$ to denote the node prior to node i in path π. Similarly, we make use of the same notation for a crew path ξ. Finally, for the route pricing (resp. crew pricing) we add approximate reduced cost terms to represent the impact of inserting a particular route (resp. crew string) on overall crew delay (resp. routing delay). We describe these ideas for the routing pricing problem; the approach for the crew pricing problem is completely analogous. Consider node i and suppose that our incumbent routing solution has a connection (ℓ, i) and our incumbent crewing solution has a connection (k, i). The combined propagated routing and crewing delays at node i are given by

$$d_{\pi(i)}^R = \max\left\{d_{\pi(\ell)}^R - (s_{\ell i} - p_{\ell i}^R),\ d_{\xi(k)}^P - (s_{ki} - p_{ki}^P),\ 0\right\}, \tag{4.7}$$

$$d_{\xi(i)}^P = \max\left\{d_{\xi(k)}^P - (s_{ki} - p_{ki}^P),\ d_{\pi(\ell)}^R - (s_{\ell i} - p_{\ell i}^R),\ 0\right\}, \tag{4.8}$$

Suppose that in the current routing pricing problem we consider replacing the aircraft connection (ℓ, i) with (j, i). We calculate $d_{\pi(i)}^R$ along the routing string being constructed using (4.5). If this potential replacement string is inserted into master problem basis, there will be an impact on the crew delays. Using (4.6), at node i, the new (locally calculated) crew delay is given by

$$\tilde{d}^P_{i;\pi^-(i)} = \max\left\{d^P_{\xi(k)} - (s_{ki} - p^P_{ki}),\ d^R_{\underline{\pi}(j)} - (s_{ji} - p^R_{ji}),\ 0\right\}; \tag{4.9}$$

where the tilde is used to denote a temporary calculation local to node i, using the information that $j = \pi^-(i)$ is the prior node. Additionally, $\underline{\pi}(j)$ denotes the approximation for the aircraft delay at node j, using the previous aircraft routing solution. We will use $a^P_i(\tilde{d}^P_{i;\pi^-(i)} - d^P_{\xi(i)})$ as an *estimate* of the reduced cost for crew delay attributable to node i, for the routing string under construction.

Thus, for the aircraft routing pricing problem we wish to find a path $\pi = \{so, i_1, i_2, \ldots, t\}$ from *so* to t that minimises

$$z^R = \min\left\{\sum_{i\in\pi}\left(a^R_i d^R_{\pi(i)} + w^R_i + a^P_i(\tilde{d}^P_{i;\pi^-(i)} - d^P_{\xi(i)})\right)\ :\ \pi \text{ is a path from } so \text{ to } t\right\}, \tag{4.10}$$

with the further restriction that the path π begins and ends at a maintenance base. For the crew pricing problem, we analogously suppose that we consider replacing the crew connection (k, i) with (j, i). We calculate $d^P_{\xi(i)}$ along the crew string being constructed. We (locally calculate) the aircraft delay by using:

$$\tilde{d}^R_{i;\xi^-(i)} = \max\left\{d^R_{\pi(\ell)} - (s_{\ell i} - p^R_{\ell i}),\ d^P_{\underline{\xi}(j)} - (s_{ji} - p^P_{ji}),\ 0\right\}; \tag{4.11}$$

We similarly use $a^R_i(\tilde{d}^R_{i;\xi^-(i)} - d^R_{\pi(i)})$ as an estimate of the reduced cost for the aircraft delay attributable to node i for the crew string under construction. For the crew pairing pricing problem, we impose the additional upper limit H on the number of hours worked.

$$z^P = \min\left\{\sum_{i\in\xi}\left(a^P_i d^P_{\xi(i)} + w^P_i + a^R_i(\tilde{d}^R_{i;\xi^-(i)} - d^R_{\pi(i)})\right)\ :\ \begin{array}{c}\xi \text{ is a path from } so \text{ to } t,\\ \text{total hours worked} \le H.\end{array}\right\}, \tag{4.12}$$

with the further restriction that the path ξ begins and ends at the same crew base.

Upon obtaining a solution to (4.10) (resp. (4.12)), the minimising path (or string) forms a column A_j of the matrix A^R (resp. A^P). A routing string is assigned a cost of

$$\begin{aligned} c^R_j &= z^R - \sum_{i\in\pi} w^R_i,\\ &= \sum_{i\in\pi}\left(a^R_i d^R_{\pi(i)} + a^P_i(\tilde{d}^P_{i;\pi^-(i)} - d^P_{\xi(i)})\right). \end{aligned} \tag{4.13}$$

and a crew pairing string is assigned a cost of

$$\begin{aligned} c_j^P &= z^P - \sum_{i\in\pi} w_i^P, \\ &= \sum_{i\in\xi}\left(a_i^P d_{\xi(i)}^P + a_i^R(\tilde{d}_{i;\xi^-(i)}^R - d_{\pi(i)}^R)\right). \end{aligned} \tag{4.14}$$

In Section 4.4.4 the z^R and z^P - minimising paths are determined by a modified label setting algorithm that simultaneously calculates both the reduced cost of the path and the propagated delays.

In this section we describe our iterative approach for handling the two master problems of aircraft routing and crew pairing, and our computational approach for solving the pricing problem.

4.4.3 Integration Of Aircraft Routing And Crew Pairing

We seek a minimal propagated delay cost solution to the integrated aircraft routing and crew pairing problem. It is well known (eg. [15, 98]) that both the aircraft routing and crew pairing problems are individually $\mathcal{NP}$-hard. To avoid any additional complexity, we adopt the theme of modelling the interactions between the aircraft and the crew in an iterative way from Weide *et al.* [98]. In the first version of our approach, we solve the integrated problem iteratively, beginning with the aircraft routing problem, linked to output from a crew pairing problem and then switching to the crew pairing problem linked to new output from the aircraft routing problem, and so on. We call this first approach Iterative Case A. This approach is not exact, however we have carefully modelled the crew and aircraft delay interactions and expect to obtain solutions of good quality. In Section 4.5 we demonstrate that we achieve significant improvements over standard approaches and our solutions also compare well against a rigorous lower bound. We also study Iterative Case B, where the initial iteration begins with the crew pairing problem linked to output from an aircraft routing problem, and then proceeds to iterate as in Case A. The pricing problem solution approach is described in the next subsection.

We begin by introducing an updating algorithm (4.4.1) that ensures the (propagated) delay for each node is updated with respect to the incumbent aircraft routing and crew pairing solution.

Algorithm 4.4.1: Propagated Delay Evaluation

Input: An incumbent aircraft routing and crew pairing solution. The delay and slack information for each connection in the network

Output: The total propagated delay in the network.

begin

1. Perform a topological sorting of the flight nodes so that the flights are sorted from earliest to latest.

2. Using the strings from the incumbent routing and crewing solution, update d_j^R and d_j^P together by inductively applying equations (4.5) and (4.6) moving strictly forwards throughout the day.

end

Algorithm 4.4.2: Iterative AR and CP (Case A)

1 **INITIALISATION:**

(a). Solve problems (4.1) and (4.2) respectively with the objective of determining the minimum number of aircraft N and the minimum number of crew required M, to cover all flights exactly once. We now have incumbent routing and crewing solutions.

(b). For each arc $(i, j) \in \mathcal{A}$, assign expected primary delays p_{ij}^R and p_{ij}^R.

(c). Set $d_k^P = 0$, $d_k^R = 0$ for all $k \in \mathcal{N}$ and $d_{so}^R = 0$, $d_{so}^P = 0$. Set an iteration counter $c = 0$.

2 **MINIMUM DELAY AIRCRAFT ROUTING:**

(a). Apply Algorithm 4.4.1.

(b). Assign delay costs to strings using (4.13). Solve problem (4.1) via column generation to produce a new incumbent routing solution.

3 **MINIMUM DELAY CREW PAIRING:**

(a). Apply Algorithm 4.4.1.

(b). Assign delay costs to strings using (4.14). Solve problem (4.2) via column generation to produce a new incumbent crew pairing solution.

4 If either the aircraft routing or crew pairing solution has changed, increment iteration counter $c \to c + 1$ and return to Step 2. Otherwise, goto Step 5.

5 Return $\sum_{n=1}^{N} \sum_{i \in \pi_n^R} a_i^R d_i^R + \sum_{m=1}^{M} \sum_{i \in \pi_m^P} a_i^P d_i^P$, where π_n^R is the routing string for the n^{th} aircraft, $n = 1, \ldots, N$ and π_m^P is the crew pairing string for the m^{th} crew, $m = 1, \ldots, M$.

Algorithm 4.4.3: Iterative AR and CP (Case B)

As for Algorithm 4.4.2, interchanging Steps 2 and 3, replacing Step 1(c) with:

Set $d_k^R = 0$ for all $k \in \mathcal{N}$ and $d_{so}^P = 0$. Set an iteration counter $c = 0$.

4.4.4 Solving The Pricing Problems

We describe the methodology to solve the pricing problem (4.10); the problem (4.12) requires straight-forward modifications described at the conclusion of this section. For each $i \in \mathcal{N}$, we are given a dual multiplier $-w_i^R$ ($-w_{so}^R = -w_t^R = 0$), a per unit delay cost a_i^R ($a_{so}^R = a_t^R = 0$), and propagated delays for crew pairings d_i^P. We wish to solve (4.10), where the d_i^R are calculated via (4.5). Because the delay d_i^R is not a simple sum of delays along the path from so to i, the problem (4.10) is not easily cast as a minimum cost network flow. We propose a label setting algorithm, augmented by a notion of label dominance, modified from related problems in Desrochers and Soumis [31] and Dumitrescu *et al.* [34], that works efficiently in the cases tested.

Let π be a (full) path in G (an ordered collection of nodes $\{so, i_1, i_2, \ldots, i_q, t\}$ in $\mathcal{N}$ with $(so, i_1), (i_q, t) \in \mathcal{A}$ and $(i_\ell, i_{\ell+1}) \in \mathcal{A}$ for all $\ell = 1, \ldots, q-1$). For $i \in \pi$, let $\pi(i)$ denote the ordered collection of nodes in the path π truncated so that the final node in the list is i; we will also call $\pi(i)$ a path. Define $W_{\pi(i)}^R = \sum_{j\in\pi(i)} \left(w_j^R + a_j^P (\tilde{d}_{j;\pi^-(j)}^P - d_{\xi(j)}^P) \right)$. Denote by $d_{\pi(i)}^R$ the propagated expected routing delay at node i, computed along path $\pi(i)$ using (4.5), and define $A_{\pi(i)}^R = \sum_{j\in\pi(i)} a_j^R d_{\pi(j)}^R$.

In this terminology, we may rewrite (4.10) as

$$z^R = \min \left\{ A_{\pi(t)}^R + W_{\pi(t)}^R \ : \ \pi \text{ is a path from } so \text{ to } t \right\}. \tag{4.15}$$

Because of the nonlinear nature of the propagated routing delay formula (4.5), our labels must track both the accumulated cost $A_{\pi(i)}^R + W_{\pi(i)}^R$ at node i along path π, *and* the propagated delay $d_{\pi(i)}^R$. This motivates the following dominance conditions for labels.

Definition 4.4.1. *(Dominance condition)*
The pair (or label) $(A_{\pi(i)}^R + W_{\pi(i)}^R,\ d_{\pi(i)}^R)$ *dominates* $(A_{\eta(i)}^R + W_{\eta(i)}^R,\ d_{\eta(i)}^R)$ *if*

$$A_{\pi(i)}^R + W_{\pi(i)}^R \le A_{\eta(i)}^R + W_{\eta(i)}^R \qquad \textit{and} \qquad d_{\pi(i)}^R \le d_{\eta(i)}^R$$

and the labels are not identical.

Lemma 4.4.2. *Let ϖ be a path from i to j, where $(i,j) \in \mathcal{A}$. If $(A_{\pi(i)}^R + W_{\pi(i)}^R,\ d_{\pi(i)}^R)$ dominates $(A_{\eta(i)}^R + W_{\eta(i)}^R,\ d_{\eta(i)}^R)$, then $(A_{\{\pi(i),\varpi\}}^R + W_{\{\pi(i),\varpi\}}^R,\ d_{\{\pi(i),\varpi\}}^R)$ dominates $(A_{\{\eta(i),\varpi\}}^R + W_{\{\eta(i),\varpi\}}^R,\ d_{\{\eta(i),\varpi\}}^R)$.*

Proof: We show that this is true if i connects to j by a single arc (the path ϖ consists of a single node $\{j\}$); the result then follows by induction. Recall we are given a fixed set of crew pairing strings. Let ξ denote the crew pairing string that includes flight node j and let k be the node in ξ preceding j. Thus,

$$d^R_{\{\pi(i),j\}} = \max\left\{d^R_{\pi(i)} - (s_{ij} - p^R_{ij}), d^P_{\xi(k)} - (s_{kj} - p^P_{kj}), 0\right\}, \quad \text{and}$$
$$d^R_{\{\eta(i),j\}} = \max\left\{d^R_{\eta(i)} - (s_{ij} - p^R_{ij}), d^P_{\xi(k)} - (s_{kj} - p^P_{kj}), 0\right\}.$$

Since $d^R_{\pi(i)} \leq d^R_{\eta(i)}$, one has $d^R_{\{\pi(i),j\}} \leq d^R_{\{\eta(i),j\}}$.

Now

$$A^R_{\{\pi(i),j\}} + W^R_{\{\pi(i),j\}} = A^R_{\pi(i)} + W^R_{\pi(i)} + a^R_j d^R_{\{\pi(i),j\}} + w^R_j + a^P_j(\tilde{d}^P_{j;\pi^-(j)} - d^P_{\xi(j)}) \quad \text{and}$$
$$A^R_{\{\eta(i),j\}} + W^R_{\{\eta(i),j\}} = A^R_{\eta(i)} + W^R_{\eta(i)} + a^R_j d^R_{\{\eta(i),j\}} + w^R_j + a^P_j(\tilde{d}^P_{j;\pi^-(j)} - d^P_{\xi(j)}),$$

and we are done.

■

In particular, if ϖ terminates at t, the above lemma shows that $A^R_{\{\pi(i),\varpi\}} + W^R_{\{\pi(i),\varpi\}} \leq A^R_{\{\eta(i),\varpi\}} + W^R_{\{\eta(i),\varpi\}}$. In our labelling algorithm described below, we may therefore at each node only create labels for those paths which are not dominated by any other path at that node. We call such labels *efficient*.

Definition 4.4.3. *A label* $(A^R_{\pi(i)} + W^R_{\pi(i)},\ d^R_{\pi(i)})$ *at node* i *is said to be* efficient *if it is not dominated by any other label at node* i*. A path* $\pi(i)$ *is said to be efficient if the label it corresponds to at node* i *is efficient.*

We now describe the label setting algorithm we use to solve the problem (4.15). At a node $i \in \mathcal{N}$, the current collection of labels are denoted I_i and the current collection of treated labels we denote by M_i. Because the dominance condition does not allow identical labels at a node i, each label in I_i will correspond to a unique path (say $\pi(i)$) from *so* to i. For brevity, we will therefore denote individual elements of I_i and M_i as paths such as $\pi(i)$.

Algorithm 4.4.4: Label Setting Algorithm for the Aircraft Routing Problem

1. *Initialisation:*

   Set $I_{so} = \{so\}$ and $I_i = \emptyset$ for all $i \in \mathcal{N} \backslash \{so\}$.

   Set $M_i = \emptyset$ for each $i \in \mathcal{N}$.

2. *Selection of the label to be treated:*

   if $\bigcup_{i \in \mathcal{N}} (I_i \backslash M_i) = \emptyset$ **then go to** Step 4; all efficient labels have been generated.

   else choose $i \in \mathcal{N}$ and $\pi(i) \in I_i \backslash M_i$ so that $A^R_{\pi(i)} + W^R_{\pi(i)}$ is minimal.

3. *Treatment of label* $(A^R_{\pi(i)} + W^R_{\pi(i)}, d^R_{\pi(i)})$

   forall $(i, j) \in \mathcal{A}$

   if $(A^R_{\{\pi(i),j\}} + W^R_{\{\pi(i),j\}}, d^R_{\{\pi(i),j\}})$ is not dominated by $(A^R_{\eta(j)} + W^R_{\eta(j)}, d^R_{\eta(j)})$ for any $\eta(j) \in I_j$ **then**

   set $I_j = I_j \cup \{\pi(i), j\}$

   end do

   Set $M_i := M_i \cup \{\pi(i)\}$.

   Go to Step 2.

4. Return $\arg\min_{\pi(t) \in I_t} A^R_{\pi(t)} + W^R_{\pi(t)}$.

We now describe the modifications required to solve the corresponding problem for the crew. Define $T_{\pi(i)} = \sum_{j \in \pi(i)} t_j$, where t_j is the scheduled time that crew work on flight j. We denote the allowed upper limit of continuous scheduled crew work time by H. Equation (4.12) can be written as

$$z^P = \min \left\{ A^P_{\pi(t)} + W^P_{\pi(t)} \ : \ \pi \text{ is a path from } so \text{ to } t,\, T_{\pi(t)} \leq H \right\}. \tag{4.16}$$

Definition 4.4.4. *(Dominance condition)*
The pair (or label) $(A^R_{\pi(i)} + W^R_{\pi(i)},\ d^R_{\pi(i)},\ T_{\pi(i)})$ *dominates* $(A^R_{\eta(i)} + W^R_{\eta(i)},\ d^R_{\eta(i)},\ T_{\eta(i)})$ *if*

$$A^R_{\pi(i)} + W^R_{\pi(i)} \leq A^R_{\eta(i)} + W^R_{\eta(i)} \quad \textit{and} \quad d^R_{\pi(i)} \leq d^R_{\eta(i)} \quad \textit{and} \quad T_{\pi(i)} \leq T_{\eta(i)}$$

and the labels are not identical.

In Algorithm 4.4.5 we do not propagate paths to a node i if $T_{\pi(i)} > H$.

Algorithm 4.4.5: Label Setting Algorithm for the Crew Pairing Problem

As in Algorithm 4.4.4, replacing R superscripts by P superscripts throughout and replacing the **if** clause in Step 3 with:

if $T_{\{\pi(i),j\}} \leq H$ hours **and** $(A^P_{\{\pi(i),j\}} + W^P_{\{\pi(i),j\}}, d^P_{\{\pi(i),j\}}, T_{\{\pi(i),j\}})$ is not dominated by $(A^P_{\eta(j)} + W^P_{\eta(j)}, d^P_{\eta(j)}, T_{\eta(j)})$ for any $\eta(j) \in I_j$ **then**
set $I_j = I_j \cup \{\pi(i), j\}$

One could try to improve the efficiency of Algorithms 4.4.4 and 4.4.5, by for example using ideas from [34] for Algorithm 4.4.5. We found the algorithms to be efficient on the instances tested and therefore have not explored further possible improvements.

4.5 Numerical Results

To evaluate the effectiveness of our proposed iterative approach, we apply Algorithm 4.4.2 to a short-haul, one-day (domestic) schedule on a real airline network consisting of 54 flights and 128 feasible connections.

We determine that the minimum number of aircraft and crew pairs required to cover this network are 10 and 16, respectively, by solving (4.1) and (4.2). For simplicity we assume that all aircraft, crew and connections incur similar operating costs, and thus the minimum number of aircraft and crew pairs solution represents a cost minimisation without regard for costs due to unforeseen delays. We use the corresponding aircraft routings and crew pairings to form our Base Case to which we apply our iterative integrated approach to reduce total propagated delay. We use 10 aircraft and 16 crew pairs in all instances and all algorithms tested.

The mean primary aircraft and crew pairing delays p_{ij}^R and p_{ij}^P are randomly sampled from four different probability distributions. In practice, primary aircraft and crew pairing delays rarely correspond to a specific distribution, but are rather a composite of several causes of delays with different individual distributions that may vary throughout different times of the day [96, 102]. It is often difficult to extract bias free, accurate historical data for the expected primary aircraft and crew delay over a specific connection. Thus, precise delay distributions (and their means) for all connections are very difficult to determine analytically. We therefore sample a set of delays and use the values obtained to represent a possible mean delay for each connection. To capture the asymmetric nature of the aircraft and crew delays, we sample from an exponential distribution $E(\lambda)$ with mean $1/\lambda$ in minutes and a truncated normal distribution (truncated to non-negative delays), denoted $tN(\mu, \sigma^2)$ with mean μ and variance σ^2, both in minutes. We test our new computational approach on 12 random instances: 3 instances from $E(1/5)$, 3 from $E(1/10)$, 3 from $tN(5, 100)$, and 3 from $tN(10, 25)$. We use unit costs per unit delay for all connections.

We study two simplified approaches (SSD) and (SSP) in addition to our base case (B) and proposed approach (IPD). We also compare our results with the method of [98] (W) as well as a proposed improvement to the method of [98] (WI):

1. **Base (B):**

   - Step 1 of Algorithm 4.4.2, followed by Algorithm 4.4.1 and Step 5 of Algorithm 4.4.2.

2. **Routing and Crewing Solved Sequentially, Simple Delay (SSD):**

   - Steps 1, 2, 3 of Algorithm 4.4.2, followed immediately by Algorithm 4.4.1 and Step 5 of Algorithm 4.4.2. In Algorithm 4.4.1, (4.5) is replaced with $d_j^R = d_i^R - (s_{ij} - p_{ij}^R)$ and (4.6) is replaced with $d_j^P = d_k^R - (s_{kj} - p_{kj}^R)$. In Algorithm 4.4.2, (4.13) is replaced with $c_j^R = \sum_{i \in \pi} a_i^R d_i^R$ and (4.14) is replaced with $c_j^P = \sum_{i \in \pi} a_i^P d_i^P$.

3. **Routing and Crewing Solved Sequentially, Propagated Delay (SPD):**

   - Steps 1, 2, 3 of Algorithm 4.4.2, followed immediately by Algorithm 4.4.1 and Step 5 of Algorithm 4.4.2.

4. **Routing and Crewing Integrated, Propagated Delay (IPD):**

   - Algorithm 4.4.2.

5. **The Algorithm of Weide *et al.* [98] (W)**

   - The algorithm as described in Weide *et al.* In the absence of cost-differentiation for different crew pairings, we set the crew pairing cost to zero.

6. **An Improved version of the Algorithm of Weide *et al.* [98] (WI)**

   - The algorithm W, with an attempt to incorporate a "time-of-day" aspect based on expected primary delay. We Compute restricted connections using the scheduled slack minus the expected primary delay, instead of scheduled slack.

The SPD approach will demonstrate the value of calculating the more accurate, nonlinear, *propagated delay* over the simpler, less accurate linear delay of the SSD approach. Our proposed IPD approach will demonstrate the value of *integrating* routing and crewing, rather than simply performing them sequentially as in the SPD approach. The SPD approach may be viewed as an improvement over Lan *et al.* [57] because we use the correct calculation of propagated delay in column selection and also model interaction of aircraft and crew (see discussion in Section 4.1). The IPD approach is an improvement over AhmadBeygi *et al.* [4] as we correctly calculate the combined propagated delay due to aircraft and crew; moreover, we develop routing and crewing connections, rather than re-timing existing connections. We also view IPD as an improvement over Weide *et al.* [98] as our objective is in terms of a dollar cost, which can be easily added to other operating cost terms in a more sophisticated cost model. We compare our IPD approach with the model of Weide *et al.* W and also with the "improved" model WI.

For each instance and each of the approaches SSD, SPD, and IPD, we record in minutes the aircraft delay, crew delay, total delay, and improvement in total delay relative to the total delay incurred by the Base Case. In each approach we apply the evaluation Algorithm 4.4.1 to provide a consistent means of comparison between each of the approaches. Algorithm 4.4.2 takes between 3 and 16 iterations for the 12 instances tested, as indicated in the tables below.

We remark that we evaluated Algorithm 4.4.3 on the same 12 instances and produced solutions that were universally inferior to Algorithm 4.4.2. This is not unexpected, as the routing strings are larger and less flexible than the crewing strings, and folklore suggests making decisions on less flexible items first often produces better results. The results for Algorithm 4.4.3 are thus not reported.

The IP was always solved at the root node by column generation and did not require any further branching. As the network consisted of 54 flights, the master problem consisted of 54 set partitioning constraints for both the aircraft routing and crew pairing problems. Approximately 200 columns were generated in an aircraft routing iteration and approximately 120 in a crew pairing iteration.

We also solved (4.1) and (4.2) separately to minimise the individual propagated delay due to aircraft and crew, respectively. These values are tabulated below, along with their sum, which represents a rigorous lower bound. This lower bound is unlikely to be sharp as it completely ignores the additional delays due to the combination of aircraft and crew delay; in some instances this combined effect can be substantial. In most instances our IPD solution is close to this lower bound; given the lack of sharpness of this bound, the IPD solutions appear to be of high quality. When running the algorithms W and WI, we found that as our network consists of many restricted connections, we could not achieve a non-robustness measure (NRM) of zero; but rather terminated when the NRM could not be improved further, as stipulated in [98]. For each instance, there were 9 restricted aircraft changes in the final solution; 8 of these may be classified as "less severe", as the sit time exceeded the minimum sit time by more than 15 minutes.

Our numerical results for Algorithm 4.4.2 are tabulated below. Individual results are given for each instance, followed by a summary in Table 4.1, detailing the relative improvements in delay between the algorithms SSD, SPD, IPD, W, and WI. All experiments were done with CPLEX12.1 on a 2.4GHz PC with 4GB RAM.

Exponential Distribution With Mean $\lambda = 5$.

Instance 1:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement on B	Time (s)
B	214	316	530	–	9.17
SSD	155	229	384	27.55	21.53
SPD	146	229	375	29.25	28.41
IPD (3 iter.)	132	229	361	31.89	47.19
Lower Bound	106	210	316	–	–
W (10 iter.)	143	236	379	–	12.75
WI (8 iter.)	138	232	370	–	12.48

Instance 2:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement on B	Time (s)
B	367	395	762	–	9.10
SSD	326	394	720	5.51	23.56
SPD	326	379	705	7.48	31.20
IPD (8 iter.)	321	347	668	12.34	68.01
Lower Bound	177	335	512	–	–
W (10 iter.)	350	390	740	–	12.75
WI (9 iter.)	349	388	737	–	10.77

Instance 3:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement on B	Time (s)
B	158	316	474	–	10.15
SSD	164	295	459	3.16	23.41
SPD	160	297	457	3.59	28.55
IPD (7 iter.)	116	297	413	12.87	63.45
Lower Bound	104	275	379	–	–
W (10. iter)	141	316	457	–	12.75
WI (10 iter.)	126	315	441	–	15.02

Exponential Distribution With Mean $\lambda = 10$.

Instance 4:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement on B	Time (s)
B	341	544	885	–	8.12
SSD	312	501	813	8.14	24.21
SPD	267	478	745	15.82	29.50
IPD (4 iter.)	241	471	712	19.55	72.26
Lower Bound	185	468	653	–	–
W (10 iter.)	312	501	813	–	12.75
WI (10 iter.)	304	485	789	–	16.00

Instance 5:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement on B	Time (s)
B	999	1114	2113	–	15.53
SSD	826	1216	2042	3.36	26.22
SPD	856	1039	1895	10.32	28.16
IPD (16 iter.)	825	879	1704	19.36	214.19
Lower Bound	590	879	1469	–	–
W (10 iter.)	895	1042	1937	–	12.75
WI (8 iter.)	890	1027	1917	–	11.78

Instance 6:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement on B	Time (s)
B	1217	1846	3063	–	14.51
SSD	1117	1653	2770	9.57	22.34
SPD	1108	1516	2624	14.33	23.19
IPD (4 iter.)	1032	1500	2532	17.34	92.35
Lower Bound	994	1456	2450	–	–
W (10 iter.)	1070	1589	2659	–	12.75
WI (10 iter.)	1053	1573	2626	–	11.33

Truncated Normal Distribution With $\mu = 5$, $\sigma = 10$.

Instance 7:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement on B	Time (s)
B	438	665	1103	–	14.19
SSD	465	598	1063	3.63	22.10
SPD	441	598	1039	5.80	25.46
IPD (4 iter.)	387	573	960	12.96	39.44
Lower Bound	260	434	694	–	–
W (10 iter.)	425	591	1016	–	12.75
WI (8 iter.)	416	582	998	–	10.80

Instance 8:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement on B	Time (s)
B	536	650	1186	–	13.31
SSD	503	689	1192	−0.51	24.75
SPD	503	652	1155	2.61	25.91
IPD (7 iter.)	505	571	1076	9.27	168.74
Lower Bound	481	562	1043	–	–
W (10 iter.)	526	647	1173	–	12.75
WI (9 iter.)	524	645	1169	–	12.56

Instance 9:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement on B	Time (s)
B	274	562	836	–	15.11
SSD	269	408	677	19.02	27.89
SPD	260	434	694	16.99	28.61
IPD (6 iter.)	227	408	635	24.04	57.98
Lower Bound	168	401	569	–	–
W (10 iter.)	267	455	722	–	12.75
WI (10 iter.)	267	452	719	–	11.14

Truncated Normal Distribution With $\mu = 10$, $\sigma = 5$.

Instance 10:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement on B	Time (s)
B	482	799	1281	–	14.91
SSD	526	780	1306	−1.95	23.66
SPD	399	731	1130	11.79	25.11
IPD (4 iter.)	366	731	1097	14.36	53.35
Lower Bound	312	703	1015	–	–
W (10 iter.)	470	792	1262	–	12.75
WI (7 iter.)	470	788	1258	–	10.16

Instance 11:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement on B	Time (s)
B	895	1134	2029	–	14.70
SSD	804	1144	1948	3.99	24.18
SPD	825	1034	1859	8.38	25.67
IPD (5 iter.)	721	944	1665	17.94	61.72
Lower Bound	682	920	1602	–	–
W (10 iter.)	757	1010	1767	–	12.75
WI (8 iter)	746	976	1722	–	9.98

Instance 12:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement on B	Time (s)
B	446	616	1062	–	15.04
SSD	437	604	1041	1.98	19.58
SPD	442	551	993	6.50	20.01
IPD (8 iter.)	380	544	924	12.99	72.97
Lower Bound	347	532	879	–	–
W (10 iter.)	440	574	1014	–	12.75
WI (10 iter.)	440	562	1002	–	11.54

4.6 Discussion And Conclusions

Table 4.1: Relative improvements of the algorithm SPD over SSD, and IPD over SPD, SSD, W, and WI.

Instance	$\frac{(SSD-SPD)}{SSD} \times 100\%$	$\frac{(SPD-IPD)}{SPD} \times 100\%$	$\frac{(SSD-IPD)}{SSD} \times 100\%$	$\frac{(W-IPD)}{W} \times 100\%$	$\frac{(WI-IPD)}{WI} \times 100\%$
1	2.34	3.73	5.99	4.75	2.43
2	2.08	5.25	7.22	9.73	9.36
3	0.43	9.63	10.02	9.63	6.35
4	8.36	4.43	12.42	12.42	9.76
5	7.20	10.10	16.55	12.03	11.11
6	5.27	3.51	8.59	4.78	3.58
7	2.26	7.60	9.69	5.51	3.81
8	3.10	6.84	9.73	8.27	7.96
9	−2.51	8.50	6.20	12.05	11.68
10	13.48	2.92	16.00	13.07	12.80
11	4.57	10.44	14.53	5.77	3.31
12	4.61	6.95	11.24	8.88	7.78
Average	4.27	6.67	10.68	8.91	7.49

Our iterative integrated methodology for minimising propagated delay in a combined routing and crewing network has clear advantages over approaches that do not explicitly calculate propagated delay or fail to properly integrate routing and crewing.

- The value of integrating routing and crewing, rather than sequentially minimising propagated delay in routing strings, then minimising propagated delay in crew strings is clear from a comparison of IPD and SPD delays in our 12 instances. There is universal improvement over all instances; on average our IPD approach improves by 6.7% over the SPD approach.

- For the two sequential approaches tested, accurately calculating propagated delay is an improvement over using a simpler additive delay; 11 out of the 12 instances showed an improvement. On average over the 12 instances, the SPD approach improves over SSD by 4.3%.

- Finally, integrating routing and crew delays and accurately calculating the propagated delays (our IPD approach) is a clear and universal improvement over SSD with an average improvement of 10.7%.

When comparing our IPD approach with the methodology of [98] on average our approach produced schedules with 8.91% less total delay (IPD vs. W) and 7.49% less total delay (IPD vs. our "improved" version of [98] WI). The delay reductions over Algorithms W and WI are comparable to those observed by (i) the correct propagated delay was used in place of the simplified "summed" delay (SSD vs. SPD) and (ii) iteration was used in place of sequential optimisation (SPD vs. IPD).

In this proof of concept work, we have limited our study to minimising expected propagated delay, however, our methodology allows other extensions to mitigate delay related risk. For example, it is straightforward to limit the maximum expected propagated delay of any single flight. In Algorithm 4.4.4, one may disallow the creation of a path with an unacceptably high single flight delay cost in the same way that crew strings of duration greater than H hours are disallowed in Algorithm 4.4.5. Similarly, it is easy to limit the total delay cost of either a routing or crew string.

Our new integrated framework is in principle extendable to a third aspect, such as delays due to passengers and future work may explore this possibility.

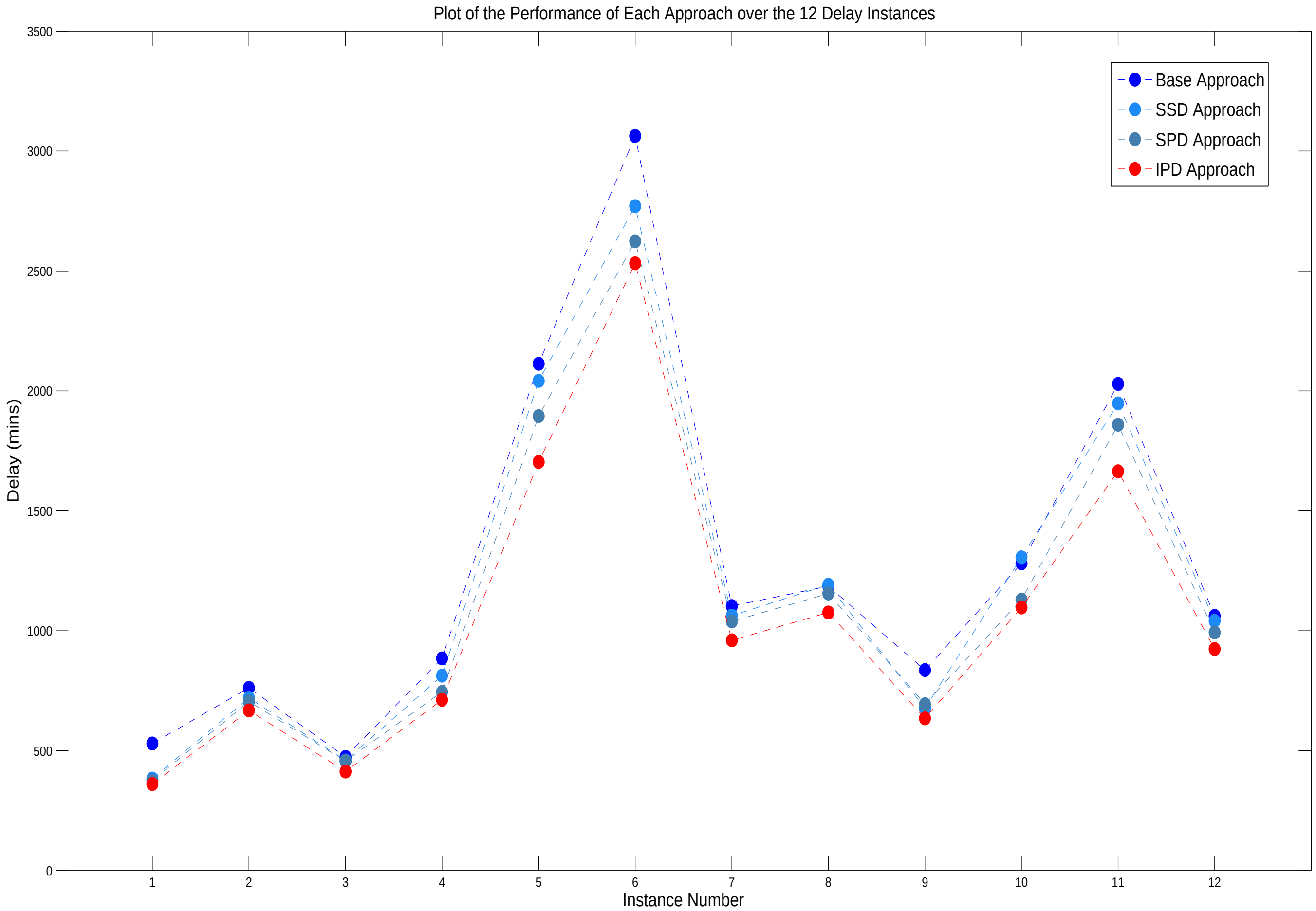
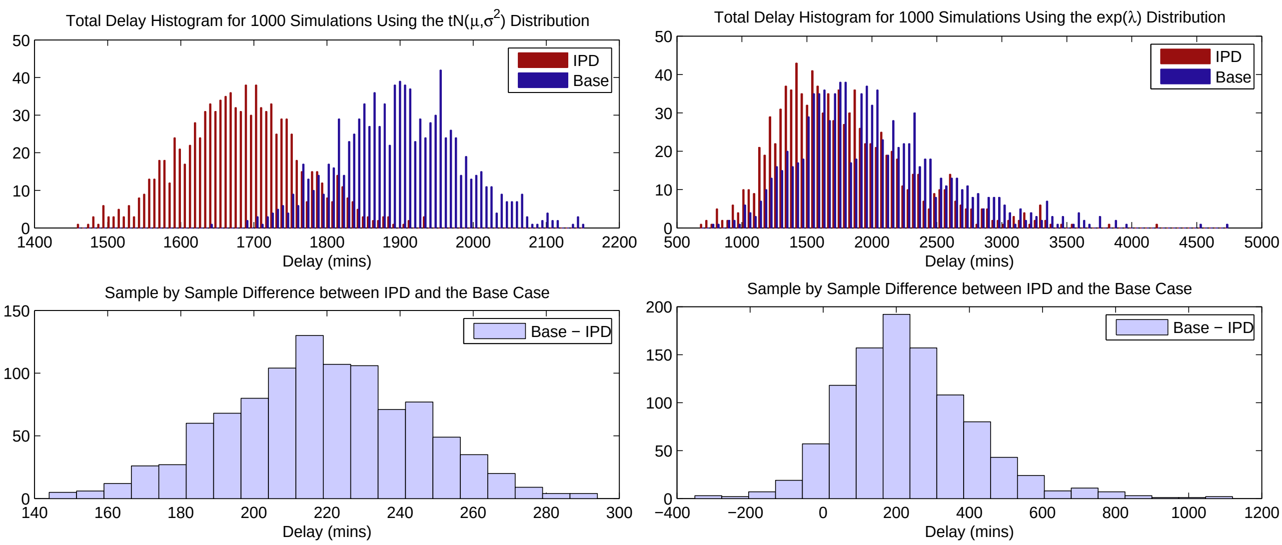


Figure 4.2: Performance of each approach on the 12 delay instances.

4.7 Performance On Different Delay Scenarios

The iterative scheme proposed in this Chapter works well for single delay instances (i.e. one (mean) primary delay for each connection in the network). However, as delays over certain connections may vary from day-to-day in operations, perhaps a natural question to ask is: How well will the solution obtained using such an approach perform on multiple sets of (different) delay scenarios? From a preliminary analysis, it appears as though the solution obtained using this approach performs particularly well (see Figure 4.3 below). It may be observed in Figure 4.3 (a) that the solution obtained using the IPD approach performs universally better than that of the Base Case. Similarly, in Figure 4.3 (b), with the exception of a few cases, the IPD approach also outperforms the Base Case solution.

Example 4.7.1. Performance of each solution on different delay scenarios.



(a) Truncated normal distribution with $\mu = 10$, $\sigma = 5$. (Instance 12).

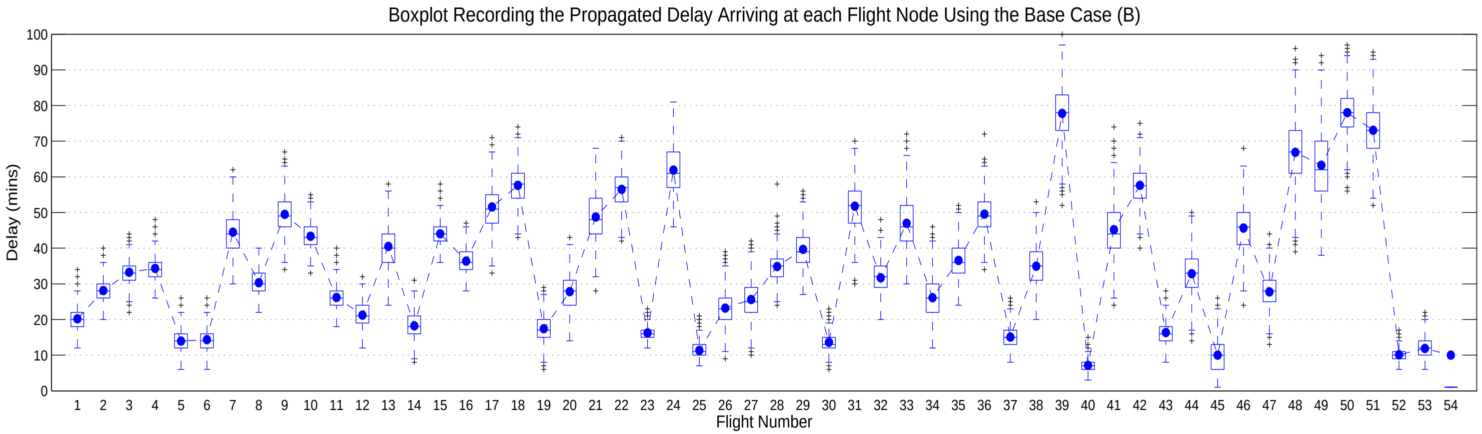
(b) Exponential distribution with $\lambda = 10$. (Instance 5).

Figure 4.3: Performance of each solution on different delay scenarios.

A similar analysis was performed to allow a comparison between the IPD approach and the other approaches (namely SSD, SPD, W and WI). However for the sake of brevity, we have not included this analysis here. However, it was observed (on average) that the IPD approach outperformed the these other approaches over 1000 different delay scenarios.

In Figure 4.4, we provide a box plot of the delay (in minutes) experienced at each flight node (over 1000 scenarios), where the flights have been sorted into topological order according to departure time. The dots correspond to the mean delay for each flight node, calculated over all the scenarios, the box represents the variation in delay for each flight node and the + symbols denote outliers. It may be seen that the IPD solution partially assists in dampening the delay propagation towards the end of the day, with all but 2 of the flights (after flight 34) arriving late less than 50% of the time, whereas approximately 1/3 of flights for the Base Case arrive late more than 50% of the time. Moreover, the IPD case allows for some flights, such as flights 39, 48 and 49 to receive a dramatic reduction in delay.

However, it would be preferable if it were somehow possible to include the delay scenarios within the aircraft routing and crew pairing subproblems of the iterative scheme to allow for a potentially more robust solution. We explore this possibility in more detail in Chapter 6.



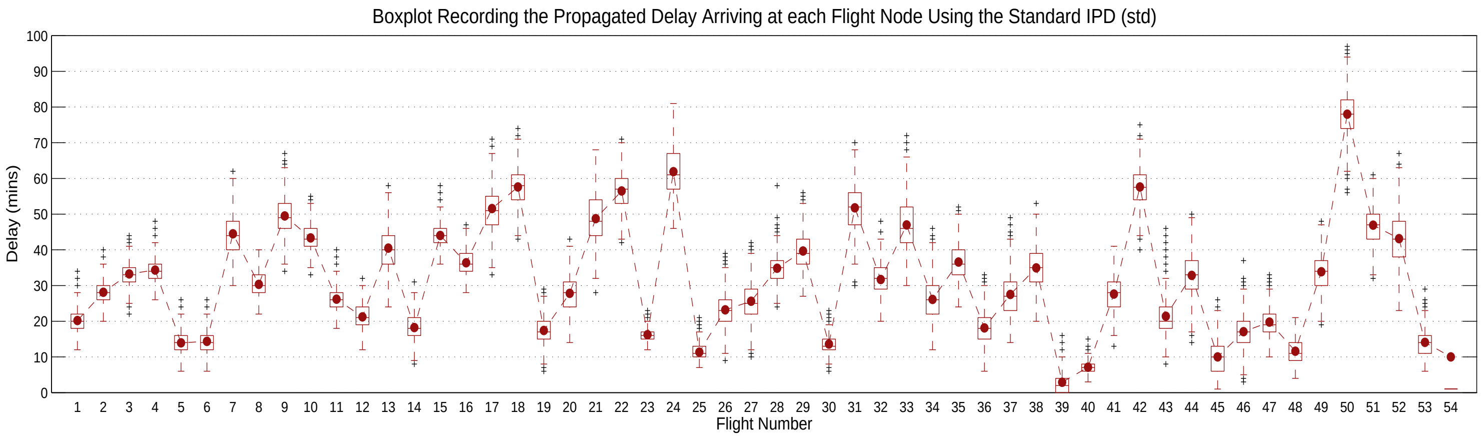


Figure 4.4: Comparison of the delay arriving at each flight node for the Base approach and the IPD approach.

CHAPTER

FIVE

The Re-timing Heuristic

5.1 Rationale Behind The Re-timing Heuristic

In Chapter 4 and in [35] we proposed an integrated aircraft routing and crew pairing model for obtaining a minimal (delay propagation) cost solution for the integrated aircraft routing and crew pairing problem. In this model, we assumed a given set of feasible connections, corresponding slack times and corresponding primary delays for both aircraft and crew over each of these connections. However, an optimal solution to this problem is the optimal assignment of aircraft and crew to flights, for the provided (*fixed*) set of departure times and list of feasible connections.

In Chapter 4 we observed that the integrated algorithm had the potential to significantly improve the delay propagation cost over that of existing models. A question naturally arises as to whether one could improve the solution further, via the adjustment of the flight departure times, so as to provide additional 'slack' over critical connections and re-allocate excess slack from the remaining connections. Such an adjustment must reflect a real-world network and consequently, although it may be easy to minimise delay propagation by indiscriminately padding every connection in the schedule with additional slack; such a solution is expensive in practice, as resources remaining idle represent lost revenue for the airline. Furthermore, in order to ensure that the fleet assignment (completed prior to the aircraft routing and crew pairing) remains feasible, it would be preferable if we could ensure such a re-timed schedule remains as close as possible to the original schedule.

In this chapter we address this question and propose a greedy heuristic algorithm capable of re-timing an incumbent aircraft and crew assignment in order to minimise the cost of delay propagation. The heuristic re-times flights without altering the aircraft and crew assignments of the incumbent solution. The algorithm is greedy in the sense that it moves forward in time throughout the day, starting with the very first flight in each route (or pairing) and then moving on to the second and so on. The algorithm proceeds forward in time, so that changes made later in the day do not affect earlier changes made. We ensure that the algorithm accepts only improvements in the overall delay and if no improvement can be made, we do not re-time the flight under consideration.

5.2 Assumptions

In order for the re-timing heuristic to obtain a solution that reflects real-world requirements, we are required to make several key assumptions; specifically, these restrictions are placed on the amount by which we can alter the departure times and therefore the corresponding slack over certain connections and in the network as whole.

Firstly, as noted above, although it may seem beneficial to pad the schedule with additional 'slack' to assist in the absorption of potential delays, this is not an optimal solution in practice; as monetary factors relating to aircraft and crew costs (eg. costs of obtaining new slot times or overtime costs for crew) make excess ground time expensive and undesirable as idle resources represent lost revenue for the airline. As a result, the associated costs of lost revenue may outweigh the benefits of the potential for delay absorption. We thus, in our algorithm, ensure that we only make use of the current slack present in the network. That is, the algorithm must only draw from the slack already present in the network and attempt to place it where it is required the most. Secondly, we seek to avoid significant changes to the publicised (incumbent) timetable so as to ensure both the aircraft routing and crew pairing solutions remain feasible and similarly the demand for each flight remains approximately the same, so as to ensure the fleet assignment remains optimal. Moreover, we wish to limit the potential for a flight re-timing to cause knock-on effects (such as the potential for passengers to miss connecting flights) and wish to avoid unnecessarily delaying the last flights of the day beyond curfew. Therefore, we borrow the idea of restricting the amount by which we can re-time each flight to within a time-window; an idea successfully employed by a number of authors namely, Rexing *et al.* (2000), [77], Lan *et al.* (2006) [57] and Mercier and Soumis (2007) [69] and Weide (2009) [99]. For simplicity, we assume a time window of $(t-10, t+10)$ around the originally scheduled departure time t, discretised into 5 minute intervals. For simplicity we assume that changes made to slack over a connection are the same for aircraft and crew. As a final note, we conclude by mentioning that by moving a flight earlier we are effectively reducing slack and moving a flight later increases the slack for that flight. Thus if we add slack to a certain flight connection, we subtract the same amount of slack from the next flight connection in the route (or pairing) as the first flight will arrive later - thus our algorithm involves only a localised effect. Furthermore, slack is only taken from within the same aircraft or crew route (or pairing). However, we note that this can be extended to allow for slack to be taken from elsewhere in the network.

5.3 Description Of Key Contributions

We now outline the key contributions of the heuristic proposed in this Chapter. Our proposed heuristic seeks to improve upon existing models in a number of key areas. Firstly, many existing models that incorporate an element of re-timing do so for either the aircraft or crew in isolation. For example, in the model of Klabjan *et al.* (2002) [56] and Lan *et al.* (2006) [57], the authors seek to allow re-timing possibilities for the aircraft routing solution (resp. crew pairing solution), but do not consider crew pairing (resp. aircraft routing). More recently the models proposed by Papadakos (2007) [74] and Mercier and Soumis (2007) [69] have sought to include re-timing possibilities for both aircraft and crew; however neither of these approaches consider the effects of delay propagation between aircraft and crew, and involve the addition of extra decision variables within the master problem, potentially leading to lengthy computation times.

Our proposed heuristic seeks to improve upon these approaches by *simultaneously* re-timing the aircraft and crew whilst *preserving* the aircraft and crew assignments of an incumbent solution. The simultaneous re-timing of aircraft and crew avoids complications arising from potential incompatibilities and allows for *accurate delay propagation calculation* between aircraft and crew. This provides an accurate assessment of the *effects of delay propagation* across certain connections and thus allows an effective re-timing to be chosen. This approach additionally overcomes the difficulties associated with the heuristic of AhmadBeygi (2008) [4] which both under-estimates and over-estimates the effects of delay propagation in certain cases.

The primary advantage of preserving the assignments of the incumbent solution is that it ensures that the re-timed solution for aircraft and crew does not differ significantly from the original aircraft and crew assignment. This may be beneficial for an airline that has aircraft routes or pairings that do not change substantially from day-to-day. Additionally, this also allows the heuristic to either be used in conjunction with a model such as the one proposed in Chapter 4 (and in Dunbar *et al.* (2011) [35]), or as an 'add-on' to an existing model, allowing for a qualitative assessment of potential areas for improvement within a given solution (e.g. bottlenecks at certain points), which may allow schedule planners an insight into where improvements can be made. The algorithm can be easily incorporated into the model of Dunbar *et al.* (2011) [35] to allow for an iterative process between the aircraft routing, crew pairing and re-timing *without increasing the overall complexity.* This possibility is discussed in further detail in Section 6.5. Finally, if an airline wishes to specify particular flights that it would prefer not to be re-scheduled (eg. First flights of the day), such a restriction is easily included within the algorithm; whereas this would more complicated to capture within an exact algorithm.

We propose two heuristics, the first in this Chapter and the second an extension of the first, and investigate the relative merits of both. Both heuristics allow for an airline to *easily customise the choice of flights* they wish to consider re-timing, as well as the extent by which each flight may be re-timed, to suit their specific needs. The first heuristic performs re-timing based on one set of primary (mean) delays for each instance, whilst the second makes use of multiple sets of possible primary delays for each instance to allow for a potentially more robust re-timing. The first heuristic is outlined in the following section, and the second heuristic is outlined in Chapter 6. Chapter 6 additionally includes two new approaches for including delay scenarios within the the aircraft routing and crew pairing subproblems of Chapter 4 and within the heuristic.

5.4 The Algorithm

Algorithm 5.4.1: The re-timing heuristic using mean delays: (Re-timed IPD)

Input: An incumbent aircraft routing and crew pairing solution and a set of primary delays with corresponding slack over each connection for the aircraft and crew.

Output: An improved choice of slack for each feasible connection in the network.

```
Start AR
Set l := 1 (the first flight in each string) and set slackOptions = [−10, −5, 0, 5, 10].
for i from 1 to numberOfAircraft do
    Pick flight string i.
    for j from 1 to numSlackOptions do
        - Set s_opt :=slackOptions[j].
        - Find the flight k, that precedes flight l in string i. (N.B: k = 0, if l = 1).
        - Find the flight m that follows flight l in string i.
        - Construct the temporary slack vectors ŝ^R := s^R and ŝ^P := s^P.
        if (ŝ^R_{k,l} + s_opt) ≥ 0 and (ŝ^P_{k,l} + s_opt) ≥ 0 and (ŝ^R_{l,m} − s_opt) ≥ 0 and
        (ŝ^P_{l,m} − s_opt) ≥ 0 then
            - Set ŝ^R_{k,l} := ŝ^R_{k,l} + s_opt.    (resp. ŝ^P_{k,l} := ŝ^P_{k,l} + s_opt)
            - Set ŝ^R_{l,m} := ŝ^R_{l,m} − s_opt.    (resp. ŝ^P_{l,m} := ŝ^P_{l,m} − s_opt)
            - Run the evaluation algorithm (4.4.1) using the updated slack vectors ŝ^R
            and ŝ^P and store the total delay.
        end
    end
    - Choose the delay/slack option (s*_opt) corresponding to the smallest propagated
    delay for the network and make the appropriate changes in the real slack vector.
    That is,
    - Set s^R_{k,l} := s^R_{k,l} + s*_opt.    (resp. s^P_{k,l} := s^P_{k,l} + s*_opt)
    - Set s^R_{l,m} := s^R_{l,m} − s*_opt.    (resp. s^P_{l,m} := s^P_{l,m} − s*_opt)
    - Set i := i + 1.
end
```

(Continued on next page..)

The re-timing heuristic using mean delays: Re-timed IPD (continued)

16 <u>StartCP</u>

17 **for** i *from* 1 **to** `numberOfCrew` **do**

18   Pick flight string i.

19   **for** j *from* 1 **to** `numSlackOptions` **do**

20   - Set s_{opt} :=slackOptions[j].

21   - Find the flight k that precedes flight l in string i.

22   - Find the flight m that follows flight l in string i.

23   - Construct the temporary slack vectors $\hat{s}^R := s^R$ and $\hat{s}^P := s^P$.

        if $(\hat{s}^R_{k,l} + s_{opt}) \geq 0$ **and** $(\hat{s}^P_{k,l} + s_{opt}) \geq 0$ **and** $(\hat{s}^R_{l,m} - s_{opt}) \geq 0$ **and** $(\hat{s}^P_{l,m} - s_{opt}) \geq 0$ **then**

24   - Set $\hat{s}^P_{k,l} := \hat{s}^P_{k,l} + s_{opt}$. (resp. $\hat{s}^R_{k,l} := \hat{s}^R_{k,l} + s_{opt}$)

25   - Set $\hat{s}^P_{l,m} := \hat{s}^P_{l,m} - s_{opt}$. (resp. $\hat{s}^R_{l,m} := \hat{s}^R_{l,m} - s_{opt}$)

26   - Run the evaluation algorithm (4.4.1) using the updated slack vectors $\hat{s}^R$ and $\hat{s}^P$ and store the total delay.

        end

    end

27   - Choose the delay/ slack option (s^*_{opt}) corresponding to the smallest propagated delay for the network and make the appropriate changes in the real slack vector. That is,

28   - Set $s^P_{k,l} := s^P_{k,l} + s^*_{opt}$ (resp. $s^R_{k,l} := s^R_{k,l} + s^*_{opt}$)

29   - Set $s^P_{l,m} := s^P_{l,m} - s^*_{opt}$ (resp. $s^R_{l,m} := s^R_{l,m} - s^*_{opt}$)

30   - Set $i := i + 1$

end

31 -Set $l := l + 1$. (Move on to the next flight for each string)

32 Return to 1.

5.4.1 Integration Of Aircraft Routing And Crew Pairing (Mean Delay), Followed By Heuristic Re-timing (Mean Delay)

In the following Section, we implement the first re-timing heuristic and evaluate its effectiveness on the 12 delay instances used in Chapter 4. Using this first re-timing heuristic, we wish to determine whether it is possible to improve the IPD solution further via the adjustment of slack and thus we analyse the improvements achieved by applying the heuristic to the incumbent IPD solutions obtained in Chapter 4. Algorithm 5.4.3 (below) outlines this approach.

Algorithm 5.4.3: Integrated AR, CP and re-timing (using mean delay values)

1 Solve the integrated aircraft routing and crew pairing problem using the method in algorithm (4.4.2).
2 Apply the re-timing heuristic (5.4.1) to the incumbent AR and CP solutions.
3 Apply the propagated delay algorithm (4.4.1) to the new solution to obtain the total delay.

5.5 Numerical Results

Our numerical results for Algorithm 5.4.1 are tabulated below. Aircraft routing and crew pairing solutions were obtained using the IPD case (for each instance) in Chapter 4 and were used as input for Algorithm 5.4.1. Additionally, we retain the same network and slack information from Chapter 4 and to allow for a fair comparison, utilise the same number of aircraft and crew as in Chapter 4. The algorithm was written in MATLAB (R2010b) using a 2.4GHz PC with 4GB RAM.

Individual results are provided for each instance and a summary is provided in Table 5.1, detailing the relative improvements of the Re-timed IPD over the Base Case, SSD, SPD and IPD.

Exponential Distribution With Mean $\lambda = 5$.

Instance 1:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
B	214	316	530	–	9.17
IPD (3 iter.)	132	229	361	31.89	47.19
Re-timed IPD	105	215	320	39.62	$47.19 + 2.51$

Instance 2:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
B	367	395	762	–	9.10
IPD (8 iter.)	321	347	668	12.34	68.01
Re-timed IPD	262	209	471	38.20	$68.01 + 2.60$

Instance 3:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
B	158	316	474	–	10.15
IPD (7 iter.)	116	297	413	12.87	63.45
Re-timed IPD	104	196	300	36.71	$63.45 + 3.09$

Exponential Distribution With Mean $\lambda = 10$.

Instance 4:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
B	341	544	885	–	8.12
IPD (4 iter.)	241	471	712	19.55	72.26
Re-timed IPD	185	360	545	38.42	$72.26 + 2.42$

Instance 5:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
B	999	1114	2113	–	15.53
IPD (16 iter.)	825	879	1704	19.36	214.19
Re-timed IPD	565	751	1316	37.72	$214.19 + 3.55$

Instance 6:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
B	1217	1846	3063	–	14.51
IPD (4 iter.)	1032	1500	2532	17.34	92.35
Re-timed IPD	998	1142	2140	30.13	$92.35 + 2.06$

Truncated Normal Distribution With $\mu = 5$, $\sigma = 10$.

Instance 7:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
B	438	665	1103	–	14.19
IPD (4 iter.)	387	573	960	12.96	39.44
Re-timed IPD	260	428	688	37.62	$39.44 + 2.55$

Instance 8:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
B	536	650	1186	–	13.31
IPD (7 iter.)	505	571	1076	9.27	168.74
Re-timed IPD	505	341	846	28.67	$168.74 + 2.87$

Instance 9:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
B	274	562	836	–	15.11
IPD (6 iter.)	227	408	635	24.04	57.98
Re-timed IPD	172	374	546	34.69	$57.98 + 2.23$

Truncated Normal Distribution With $\mu = 10$, $\sigma = 5$.

Instance 10:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
B	482	799	1281	–	14.91
IPD (4 iter.)	366	731	1097	14.36	53.35
Re-timed IPD	300	648	948	26.00	$53.35 + 2.40$

Instance 11:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
B	895	1134	2029	–	14.70
IPD (5 iter.)	721	944	1665	17.94	61.72
Re-timed IPD	517	698	1215	40.11	$61.72 + 2.59$

Instance 12:

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
B	446	616	1062	–	15.04
IPD (8 iter.)	380	544	924	12.99	72.97
Re-timed IPD	312	483	795	25.14	$72.97 + 3.41$

Table 5.1: Relative improvements of the Re-timed IPD over B, SSD, SPD and IPD.

Instance	$\frac{(B-\text{Re-timed IPD})}{B} \times 100\%$	$\frac{(SSD-\text{Re-timed IPD})}{SSD} \times 100\%$	$\frac{(SPD-\text{Re-timed IPD})}{SPD} \times 100\%$	$\frac{(IPD-\text{Re-timed IPD})}{IPD} \times 100\%$
1	39.62	16.67	14.67	11.35
2	38.20	34.58	33.19	29.49
3	36.71	34.64	34.35	27.36
4	38.42	32.96	26.84	23.46
5	37.72	35.55	30.55	22.76
6	30.13	22.74	18.44	15.48
7	37.62	35.28	33.78	28.33
8	28.67	29.02	26.75	21.38
9	34.69	19.35	21.32	14.02
10	26.00	27.41	16.11	13.58
11	40.11	37.63	34.64	27.03
12	25.14	23.63	19.94	13.96
Average	34.42	29.12	25.88	20.68

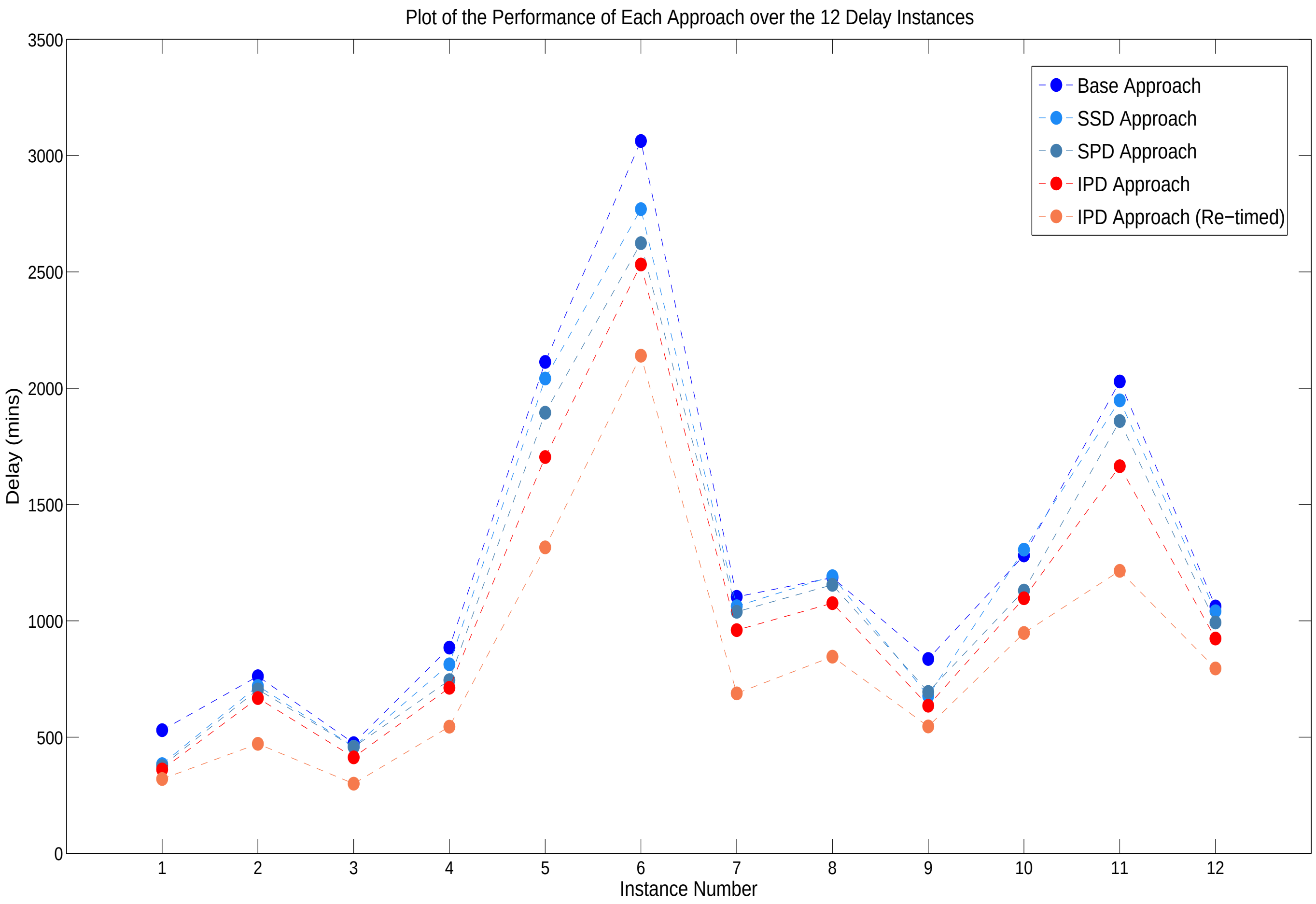


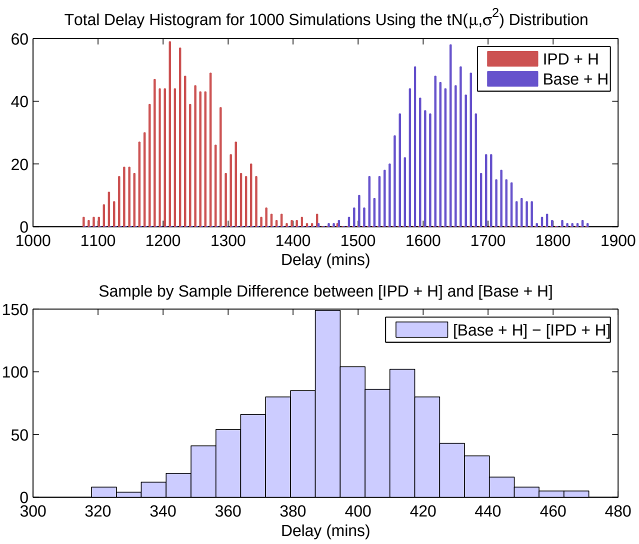
Figure 5.1: Plot of the performance of each approach over the 12 delay instances.

5.5.1 Example Solutions

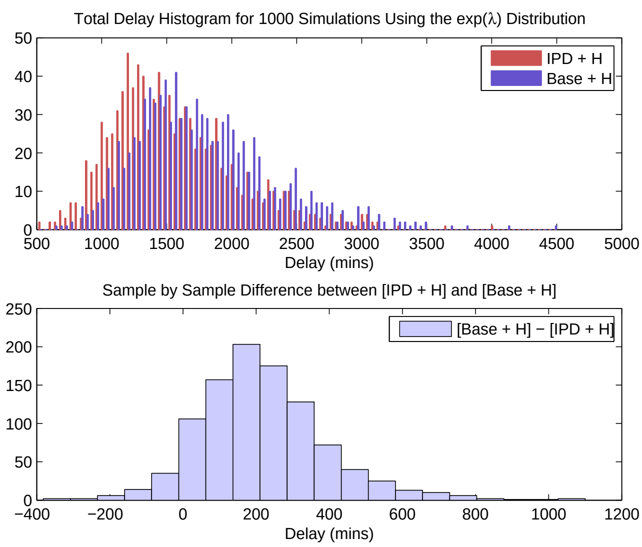
Below we present two example solutions. In both of these examples, we evaluated the performance of the re-timed aircraft and crew solution on 1000 delay scenarios drawn from a distribution with the same mean and standard deviation as the relevant delay instance. We compare the performance between the Re-timed IPD solution and the Base Case, evaluated on the same delay scenarios (see Figure 5.2 below).

We provide a sample-by-sample difference between the Base Case and the Re-timed IPD solutions. It may be observed that in a sample-by-sample comparison, the Re-timed IPD performs universally better than the Base Case for the Truncated normal distribution example, and with the exception of a few cases, performs in a similar manner on the exponential distribution example.

Example 5.5.1. Performance of each re-timed solution on different delay scenarios.



(a) Truncated normal distribution with $\mu = 10$, $\sigma = 5$. (Instance 12).



(b) Exponential distribution with $\lambda = 10$. (Instance 5).

Figure 5.2:

5.6 Discussion

From the analysis provided in the tables for each instance and further tabulated in Table 5.1, it may be observed that the re-timing heuristic, used in combination with each of the IPD solutions, out-performs the IPD approach and the alternative approaches of Chapter 4. This may be visualised most clearly within Figure 5.1, for which a significant improvement may be observed for a number of the delay instances; namely instance numbers 6, 7, 8 and 11.

In particular, the Re-timed IPD displays a significant improvement over the Base case, with an average improvement of 34.42%. An improvement of 29.12% is achieved over the SSD case and an improvement of 25.88% over the SPD case. Finally, the Re-timed IPD improves upon the IPD approach by an average of 20.68%. These results indicate, that for these test instances, there is significant potential for improvement in the reduction of delay propagation through the consideration of flight re-timing. Furthermore, as the re-timing heuristic re-times each flight within a small window of the originally scheduled departure time and retains the incumbent aircraft routing and crew pairing solution, it appears as though significant improvements may be achieved without drastic changes to the original schedule. The re-timed solutions also perform well on different sets of delay data, as highlighted in Figure 5.2 above.

In the following chapter, we seek to extend this idea and examine whether it is possible to improve upon the performance of these solutions further. We seek to improve in two key areas; namely, the *aircraft routing and crew pairing subproblems* and the *heuristic* via the inclusion of delay scenarios within the subproblems and within the re-timing heuristic.

CHAPTER

SIX

Integration Of Aircraft Routing And Crew Pairing With The Re-timing Heuristic: Including Scenarios Within The Subproblem And Re-timing Heuristic

6.1 Two Approaches For Including Scenarios Within The Subproblem

In Chapter 4 we proposed an Algorithm for minimising propagated delay in the integrated aircraft routing and crew pairing framework. In Chapter 5 we motivated the concept of re-timing within the context of airline scheduling and proposed a re-timing heuristic that may be used in conjunction with the model proposed in Chapter 4. While these two approaches delivered significant improvements in delay propagation reduction, one of the drawbacks involves the assumption that the expected delay over each connection is known prior to the assignment of aircraft and crew to flights, with delay propagation calculations utilising the expected delay over each connection. In practice, however, it may not be possible for an airline to possess this knowledge prior to aircraft routing and crew pairing, or indeed have the information to a necessary level of certainty. Moreover, given the number of sources of primary delay over a particular connection, an airline may prefer to model delay across individual connections using a distribution (as mentioned in Wu (2005, 2007) [101, 100]) – perhaps derived from data from previous years, to represent a more complete range of delays observed.

With this in mind, we wish to investigate whether it is possible to achieve further improvements on the results of both Chapter 4 and Chapter 5 by incorporating additional stochastic delay information within Algorithms 4.4.4 and 4.4.5, and additionally, within the re-timing heuristic of the previous Chapter.

We propose to incorporate this information by prescribing 1000 potential primary delay values for each connection in the network. These 1000 primary delay values (1000 for each aircraft connection and 1000 for each crew connection) will be drawn from an appropriate distribution (i.e. truncated normal or exponential distribution) possessing the same mean as the original primary delay prescribed in the previous Chapters. As mentioned in Chapter 4, the delay distribution across a particular connection is usually the combination of multiple distributions (e.g. Truncated Normal, log-normal, exponential distribution etc.), resulting from multiple interactions from multiple resources. However in general, as discussed in Chapter 4, a truncated normal, or exponential distribution are reasonable approximations to the distributions experienced in practice. Consequently, we will make use of both of these distribution types in this Chapter.

We now outline two alternative approaches for the inclusion of multiple delay scenarios within the subproblem. The first approach, referred to as simSub1, is an exact approach for which we enumerate every path in the subproblem and then calculate the average delay propagation along each string, over all delay scenarios. Our numerical results indicate that this approach, having the advantage of being exact, produces very good results on our network, but may have the disadvantage of becoming computationally expensive for larger networks. To overcome this, we propose a second approach, referred to as simSub2, in which we at each step of the label setting algorithm ((4.4.4) and (4.4.5)), calculate the average delay propagation arriving at each node and then use this to decide which label(s) to propagate. This has the advantage that potentially fewer labels and paths are produced, and we are not required to enumerate all possible paths. Both approaches perform well and outperform the case in which we simply use mean delays. We now formalise the above description and outline each approach in more mathematical detail.

6.1.1 An Exact Approach: simSub1

In this first approach, we enumerate every feasible path between the source node so, and sink node t, and determine the average delay propagation along each path over all scenarios. The algorithm then finds the path that minimises:

$$z^R = \min \left\{ \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \sum_{i \in \pi} \left(a_i^R d_{i,\omega}^R + w_i^R + a_i^P (\tilde{d}_{i,\omega;\pi^-(i)}^P - d_{i,\omega}^P) \right) \ : \ \pi \text{ is a path from } so \text{ to } t \right\}, \tag{6.1}$$

where $\omega \in \Omega$ and Ω is the set of all delay scenarios.

For the crew pricing problem, a completely analogous procedure is used to construct the reduced cost estimate $a_j^R(\tilde{d}_{j,\omega;\pi^-(j)}^R - d_{j,\omega}^R)$ for the routing delay, attributable to node j, from the crew string under construction. For the crew pairing pricing problem, we impose the additional upper limit H on the number of hours worked.

$$z^P = \min \left\{ \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \sum_{i \in \xi} \left(a_i^P d_{i,\omega}^P + w_i^P + a_i^R (\tilde{d}_{i,\omega;\xi^-(i)}^R - d_{i,\omega}^R) \right) \ : \ \begin{array}{c} \xi \text{ is a path from } so \text{ to } t, \\ \text{total hours worked} \leq H. \end{array} \right\}, \tag{6.2}$$

with the further restriction that the path ξ begins and ends at the same crew base.

Upon obtaining a solution to problem 6.1 (resp. problem 6.2), the minimising path forms a column A_j of the matrix A^R (resp. A^P). A routing string is assigned a cost of

$$\begin{aligned} c_j^R &= z^R - \sum_{i \in \pi} w_i^R, \\ &= \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \sum_{i \in \pi} \left(a_i^R d_{i,\omega}^R + a_i^P (\tilde{d}_{i,\omega;\pi^-(i)}^P - d_{i,\omega}^P) \right). \end{aligned} \tag{6.3}$$

and a crew pairing string is assigned a cost of

$$\begin{aligned} c_j^P &= z^P - \sum_{i \in \xi} w_i^P, \\ &= \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \sum_{i \in \xi} \left(a_i^P d_{i,\omega}^P + a_i^R (\tilde{d}_{i,\omega;\xi^-(i)}^R - d_{i,\omega}^R) \right). \end{aligned} \tag{6.4}$$

We now outline the algorithms for the aircraft routing and crew pairing subproblems used in the simSub1 case:

Algorithm 6.1.1: Algorithm for the Aircraft Routing Pricing Problem: simSub1

1 *Generate all paths:*

Consider the full directed graph $G = (\mathcal{N}, \mathcal{A})$. Generate all *distinct*, *directed* paths from the source node *so*, to the sink node t.

2 Record each path π in the set S of all paths.

3 For each path $\pi \in S$, calculate

$$\frac{1}{|\Omega|} \sum_{\omega \in \Omega} \sum_{i \in \pi} \left(a_i^R d_{i,\omega}^R + w_i^R + a_i^P (\tilde{d}_{i,\omega;\pi^-(i)}^P - d_{i,\omega}^P) \right),$$

and return the path that minimises (over all paths $\pi \in S$):

$$z^R = \min \left\{ \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \sum_{i \in \pi} \left(a_i^R d_{i,\omega}^R + w_i^R + a_i^P (\tilde{d}_{i,\omega;\pi^-(i)}^P - d_{i,\omega}^P) \right) \right\},$$

where π is a path from *so* to t.

Algorithm 6.1.2: Algorithm for the Crew Pairing Pricing Problem: simSub1

1 *Generate all paths:*

Consider the full directed graph $G = (\mathcal{N}, \mathcal{A})$. Generate all *distinct*, *directed* paths from the source node *so*, to the sink node t, that satisfy:

– The total number of hours worked $\leq H$ **and**,

– The last flight in the path returns to the crew-base at which the path began.

2 Record each path ξ in the set S of all paths.

3 For each path $\xi \in S$, calculate:

$$\frac{1}{|\Omega|} \sum_{\omega \in \Omega} \sum_{i \in \xi} \left(a_i^P d_{i,\omega}^P + w_i^P + a_i^R (\tilde{d}_{i,\omega;\xi^-(i)}^R - d_{i,\omega}^R) \right),$$

and return the path that minimises (over all paths $\pi \in S$):

$$z^P = \min \left\{ \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \sum_{i \in \xi} \left(a_i^P d_{i,\omega}^P + w_i^P + a_i^R (\tilde{d}_{i,\omega;\xi^-(i)}^R - d_{i,\omega}^R) \right) \right\},$$

where ξ is a path from *so* to t and total hours worked $\leq H$.

6.1.2 An Approximate Approach: simSub2

In this second approach, we make use of the label setting algorithm in Chapter 4. We refer the reader to Algorithms 4.4.4 and 4.4.5 above for the background detail. Unlike the modification used in the simSub1 pricing problem, we wish to incorporate delay information from the scenarios within Step 3 of the label setting algorithm used in the pricing problem. The motivation for including the scenario information within this step is to allow the average delay propagation cost at each node to be calculated, before the label is propagated further. Thus, through use of the dominance condition, we encourage the propagation of labels whose corresponding (partial) path experiences minimal delay propagation on average.

In this approach, we once again, make use of the scenarios consisting of 1000 primary delays, for each connection. We denote the primary delay for the aircraft across connection (i,j) under scenario $\omega \in \Omega$ by $p^R_{ij,\omega}$. Similarly, $p^P_{ij,\omega}$ denotes the primary delay for the crew across connection (i,j) under scenario ω. As flights from the source so, and to the sink t, are 'dummy' flights, we set $p^R_{jt,\omega} = 0$ and $p^P_{jt,\omega} = 0$ for all $(j,t) \in \mathcal{A}$ and $\omega \in \Omega$.

In this modification of the label setting Algorithms 4.4.4 and 4.4.5, we wish to retain Steps 1 and 2, but modify Step 3 to allow for the average delay propagation arriving at a particular node to be calculated. We thus use the notation $\widehat{d_j}$ to denote the average delay arriving at node j and more specifically, $\widehat{d^R_j}$ for the average propagated aircraft delay and $\widehat{d^P_j}$ for the average propagated crew delay. Since the source node so, is a 'dummy' node, we fix $\widehat{d^R_{so}} = 0$ and $\widehat{d^P_{so}} = 0$ and for a general flight node $j \neq 0$, we calculate the average propagated delay arriving at node j, denoted by $\widehat{d^R_j}$ as follows:

$$\begin{aligned}\widehat{d^R_j} &= \mathbb{E}_\omega\left(\max\{\widehat{d^R_i} + (s_{ij} - p^R_{ij,\omega})\ ,\ \widehat{d^P_k} + (s_{kj} - p^P_{kj,\omega})\}\right)\\ &= \frac{1}{|\Omega|}\sum_{\omega\in\Omega}\left(\max\{\widehat{d^R_i} + (s_{ij} - p^R_{ij,\omega})\ ,\ \widehat{d^P_k} + (s_{kj} - p^P_{kj,\omega})\}\right) \qquad (6.5)\end{aligned}$$

Similarly, we calculate the average propagated crew delay arriving at node j, denoted by $\widehat{d^P_j}$ as follows.

$$\begin{aligned}\widehat{d^P_j} &= \mathbb{E}_\omega\left(\max\{\widehat{d^R_i} + (s_{ij} - p^R_{ij,\omega})\ ,\ \widehat{d^P_k} + (s_{kj} - p^P_{kj,\omega})\}\right)\\ &= \frac{1}{|\Omega|}\sum_{\omega\in\Omega}\left(\max\{\widehat{d^R_i} + (s_{ij} - p^R_{ij,\omega})\ ,\ \widehat{d^P_k} + (s_{kj} - p^P_{kj,\omega})\}\right) \qquad (6.6)\end{aligned}$$

The modified aircraft and crew label setting algorithms for the simSub2 approach are outlined below. On the following page we specify how to include these label setting algorithms within the iterative aircraft routing and crew pairing framework.

Algorithm 6.1.3: Label Setting Algorithm for the Aircraft Routing Problem: `simSub2`

---

1. *Initialisation:*

   Set $I_{so} = \{so\}$ and $I_i = \emptyset$ for all $i \in \mathcal{N} \backslash \{so\}$.

   Set $M_i = \emptyset$ for each $i \in \mathcal{N}$.

2. *Selection of the label to be treated:*

   if $\bigcup_{i \in \mathcal{N}} (I_i \backslash M_i) = \emptyset$ **then go to** Step 4; all efficient labels have been generated.

   else choose $i \in \mathcal{N}$ and $\pi(i) \in I_i \backslash M_i$ so that $\widehat{A}^R_{\pi(i)} + W^R_{\pi(i)}$ is minimal.

3. *Treatment of label* $(\widehat{A}^R_{\pi(i)} + W^R_{\pi(i)}, \widehat{d}^R_{\pi(i)})$

   forall $(i, j) \in \mathcal{A}$

   Calculate the $\widehat{d}^R_j$ and $\widehat{d}^P_j$ using equations (6.5) and (6.6) over all scenarios.

   if $(\widehat{A}^R_{\{\pi(i),j\}} + W^R_{\{\pi(i),j\}}, \widehat{d}^R_{\{\pi(i),j\}})$ is not dominated by $(\widehat{A}^R_{\eta(j)} + W^R_{\eta(j)}, \widehat{d}^R_{\eta(j)})$ for any $\eta(j) \in I_j$ **then**

   set $I_j = I_j \cup \{\pi(i), j\}$

   end do

   Set $M_i := M_i \cup \{\pi(i)\}$. **Go to** Step 2.

4. Return $\arg\min_{\pi(t) \in I_t} \widehat{A}^R_{\pi(t)} + W^R_{\pi(t)}$.

Algorithm 6.1.4: Label Setting Algorithm for the Crew Pairing Problem: simSub2

As in Algorithm 6.1.3, replacing R superscripts by P superscripts throughout and replacing the **if** clause in Step 3 with:

if $T_{\{\pi(i),j\}} \leq H$ hours **and** $(\widehat{A}^P_{\{\pi(i),j\}} + W^P_{\{\pi(i),j\}}, \widehat{d}^P_{\{\pi(i),j\}}, T_{\{\pi(i),j\}})$ is not dominated by $(\widehat{A}^P_{\eta(j)} + W^P_{\eta(j)}, \widehat{d}^P_{\eta(j)}, T_{\eta(j)})$ for any $\eta(j) \in I_j$

then set $I_j = I_j \cup \{\pi(i), j\}$

We now outline how these modifications to the aircraft routing and crew pairing subproblems may be included within the iterative integrated aircraft routing and crew pairing framework of Chapter 4 to form the simSub1 and simSub2 approaches. The following Algorithms specify the necessary changes to be made to the iterative Algorithm 4.4.2.

Algorithm 6.1.5: Integrated AR and CP using simSub1 in each subproblem

---

1 Solve the integrated aircraft routing and crew pairing problem using the Iterative Algorithm 4.4.2 outlined in Chapter 4, making the following changes:

(i) Solve aircraft routing problem (4.1) via column generation, using the simSub1 aircraft routing Algorithm 6.1.1.

(i) Solve the crew pairing problem (4.2) via column generation, using the simSub1 crew pairing Algorithm 6.1.2.

Algorithm 6.1.6: Integrated AR and CP using simSub2 in each subproblem

---

1 Solve the integrated aircraft routing and crew pairing problem using the Iterative Algorithm 4.4.2 outlined in Chapter 4, making the following changes:

(i) Solve aircraft routing problem (4.1) via column generation, using the simSub2 Label Setting Algorithm 6.1.3.

(ii) Solve the crew pairing problem (4.2) via column generation, using the simSub2 Label Setting Algorithm 6.1.4.

6.2 Numerical Results: Using simSub1 And simSub2

In this section we tabulate the results for the three different approaches and investigate the significance of incorporating delay scenarios within the subproblem. As with the previous Chapters, we retain the same network and slack information from Chapter 4 and to allow for a fair comparison, utilise the same number of aircraft and crew as in the IPD solution. Average delays are calculated over 1000 different delay scenarios drawn from the same distribution as the relevant instance. The algorithm was written and solved in MATLAB (R2010b) using a 2.4GHz PC with 4GB RAM.

6.2.1 Exponential Distribution With $\lambda = 5$.

Instance 1

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improv. over B	Time(s)
Base (B)	515	439	954	–	9.17
IPD (3 iter.)	430	381	811	15.00	47.19
simSub1 (3 iter.)	416	347	763	20.02	56.19
simSub2 (3 iter.)	430	363	793	16.88	50.00

Instance 2

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improv. over B	Time(s)
Base (B)	473	478	951	–	9.10
IPD (8 iter.)	401	435	836	12.03	68.01
simSub1 (12 iter.)	403	391	794	16.51	75.58
simSub2 (10 iter.)	401	418	819	13.88	71.11

Instance 3

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improv. over B	Time(s)
Base (B)	519	550	1069	–	10.15
IPD (7 iter.)	477	532	1010	5.52	63.45
simSub1 (8 iter.)	459	464	923	13.66	69.12
simSub2 (5 iter.)	451	484	935	12.54	68.44

6.2.2 Exponential Distribution With $\lambda = 10$.

Instance 4

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improv. over B	Time(s)
Base (B)	1056	1075	2131	–	8.12
IPD (4 iter.)	985	904	1889	11.35	72.26
simSub1 (4 iter.)	982	873	1855	12.95	80.47
simSub2 (3 iter.)	969	891	1860	12.72	79.20

Instance 5

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improv. over B	Time(s)
Base (B)	954	1057	2011	–	15.53
IPD (16 iter.)	832	945	1776	11.69	214.19
simSub1 (14 iter.)	844	886	1730	13.97	228.89
simSub2 (12 iter.)	844	889	1733	13.82	220.29

Instance 6

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improv. over B	Time(s)
Base (B)	1871	1890	3761	–	14.51
IPD (4 iter.)	1704	1660	3364	10.56	92.35
simSub1 (5 iter.)	1651	1697	3348	10.98	104.39
simSub2 (3 iter.)	1683	1669	3352	10.87	96.49

6.2.3 Truncated Normal Distribution With $\mu = 5$, $\sigma = 10$.

Instance 7

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improv. over B	Time(s)
Base (B)	708	811	1519	–	14.19
IPD (4 iter.)	612	746	1358	10.60	39.44
simSub1 (4 iter.)	610	700	1310	13.76	48.18
simSub2 (4 iter.)	612	745	1358	10.60	40.86

Instance 8

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improv. over B	Time(s)
Base (B)	576	504	1080	–	13.31
IPD (10 iter.)	527	423	950	12.04	168.74
simSub1 (9 iter.)	533	384	918	15.00	175.22
simSub2 (9 iter.)	533	400	933	13.61	172.35

Instance 9

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improv. over B	Time(s)
Base (B)	723	803	1527	–	15.11
IPD (6 iter.)	667	641	1308	14.34	57.98
simSub1 (5 iter.)	656	619	1275	16.50	64.12
simSub2 (5 iter.)	656	619	1275	16.50	64.02

6.2.4 Truncated Normal Distribution With $\mu = 10$, $\sigma = 5$.

Instance 10

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improv. over B	Time(s)
Base (B)	781	990	1771	–	14.91
IPD (4 iter.)	667	904	1571	11.29	53.35
simSub1 (4 iter.)	683	850	1532	13.50	61.64
simSub2 (6 iter.)	688	866	1554	12.25	54.12

Instance 11

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improv. over B	Time(s)
Base (B)	827	1171	1999	–	14.70
IPD (5 iter.)	734	1042	1776	11.16	61.72
simSub1 (5 iter.)	713	1015	1728	13.57	65.35
simSub2 (6 iter.)	711	1030	1741	12.90	54.12

Instance 12

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improv. over B	Time(s)
Base (B)	894	1006	1901	–	15.04
IPD (8 iter.)	768	914	1682	11.52	72.97
simSub1 (8 iter.)	769	842	1611	15.26	82.66
simSub2 (4 iter.)	778	851	1629	14.31	79.53

Table 6.1: Relative improvements of the simSub1 and simSub2 approach over the IPD approach.

Instance	$\frac{(\text{IPD}-\text{simSub1})}{\text{IPD}} \times 100\%$	$\frac{(\text{IPD}-\text{simSub2})}{\text{IPD}} \times 100\%$
1	5.98	2.22
2	5.02	2.03
3	8.61	7.43
4	1.80	1.53
5	2.59	2.42
6	0.48	0.36
7	3.53	0.00
8	3.37	1.79
9	2.52	2.52
10	2.48	1.08
11	2.70	1.97
12	4.22	3.15
Average	3.61	2.30

6.2.5 Discussion

From the data recorded in the tables for each instance and summarised in Table 6.1, it may be observed that both the simSub1 and simSub2 approaches out-perform the IPD approach when embedded within the iterative aircraft routing and crew pairing approach proposed in Chapter 4. The simSub1 approach provides a 3.61% improvement (on average) over the IPD approach and the simSub2 approach provides a 2.30% improvement (on average) over the IPD approach. Both approaches universally improve over the IPD approach, across all test instances, for which the simSub1 approach consistently out-performs the simSub2 approach.

Figure 6.1 provides a plot of the percentage of time that each flight $1, 2, \ldots, 54$ would have been classified as "late" (i.e. arriving more than 15 minutes after the scheduled arrival time) over all scenarios. N.B. The flights have been sorted into topological order, according to the scheduled departure time. It may be observed that the Base Case and the IPD case perform in a similar manner until around flight number 33, after which the IPD case results in fewer flights classified as late.

Similarly, Figure 6.1 on the following page compares the simSub1 and simSub2 approaches which improve upon both the Base and IPD approaches further, with even fewer flights classified as late. Therefore, although the number of flights arriving late is fairly similar at the beginning of the day, from the middle of the day onwards our proposed approaches appear to assist in providing a "dampening effect"; as opposed to the Base Case, for which there are a greater number of flights arriving late $> 90\%$ of the time at the end of the day.

In this Chapter we have so far incorporated the 1000 primary delay scenarios within each pricing subproblem using two different approaches; namely the subproblems simSub1 and simSub2. We observed that both of these subproblems have the potential to further reduce delay propagation and thus provide an improved solution. With this in mind, a natural question to ask is whether it is possible incorporate or embed these delays within the re-timing heuristic (Algorithm 5.4.1) and improve the solution further. In the section that follows, we investigate such an approach by first specifying an improved heuristic in Algorithm 6.3.1 that chooses a re-timing based on the average delay performance over all scenarios. This improved algorithm is used extensively in the remainder of the thesis and is also combined with simSub1 and simSub2. These algorithms are outlined in detail in the following section.

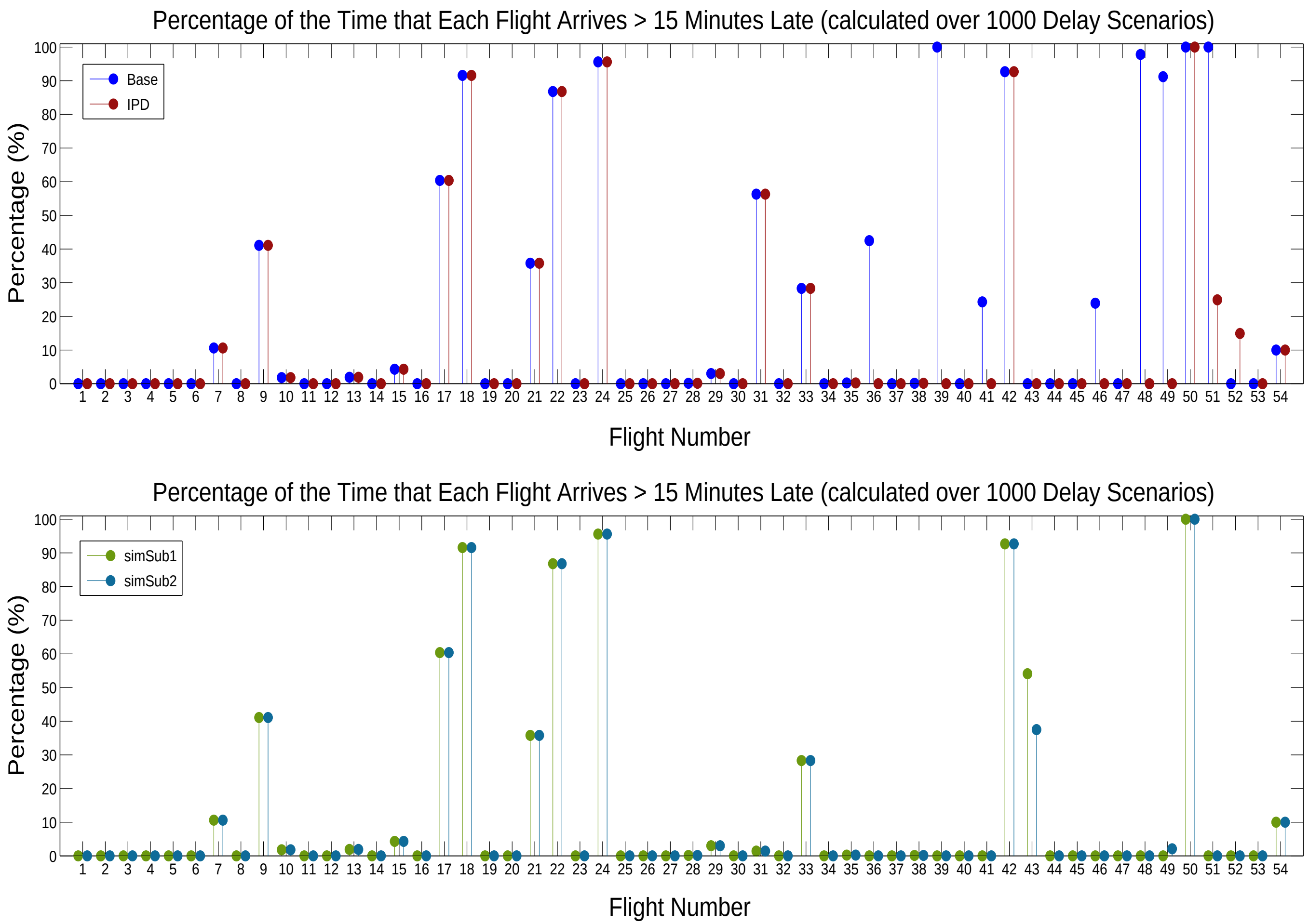


Figure 6.1: Percentage of the time that each flight arrives more than 15 minutes late (over all scenarios).

6.3 Incorporation Of Delay Scenarios Within The Heuristic

Algorithm 6.3.1: The re-timing heuristic (H) using delay scenarios

Input: An incumbent aircraft routing and crew pairing solution and a set of primary delays $\{p_{ij}\}_{ij\in\mathcal{A}}$, with corresponding slack $\{s_{ij}\}_{ij\in\mathcal{A}}$, over each connection for the aircraft and crew.

Output: An improved choice of slack for each feasible connection in the network.

Start AR

Set $l := 1$ (the first flight in each string) and set slackOptions $= [-10, -5, 0, 5, 10]$.

for i *from* 1 **to** `numberOfAircraft` **do**

  Pick flight string i.

  for j *from 1* **to** `numSlackOptions` **do**

    - Set s_{opt} :=slackOptions[j]
    - Find the flight k, that precedes flight l in string i. (N.B: $k = 0$, if $l = 1$)
    - Find the flight m, that follows flight l in string i.

    if $(\hat{s}^R_{k,l} + s_{opt}) \geq 0$ **and** $(\hat{s}^P_{k,l} + s_{opt}) \geq 0$ **and** $(\hat{s}^R_{l,m} - s_{opt}) \geq 0$ **and** $(\hat{s}^P_{l,m} - s_{opt}) \geq 0$ **then**

      - Set $\hat{s}^R_{k,l} := \hat{s}^R_{k,l} + s_{opt}$. (resp. $\hat{s}^P_{k,l} := \hat{s}^P_{k,l} + s_{opt}$)
      - Set $\hat{s}^R_{l,m} := \hat{s}^R_{l,m} - s_{opt}$. (resp. $\hat{s}^P_{l,m} := \hat{s}^P_{l,m} - s_{opt}$)

      forall the $\omega \in \Omega$ **do**

        Run the Evaluation Algorithm (4.4.1) using the updated slack vector $\hat{s}$ and store the corresponding total delay.

      end

      - Calculate the average delay over all the simulations and record this average delay with its corresponding slack option.

    end

  end

  - Choose the best delay/slack option (s^*_{opt}) and make the appropriate changes in the real slack vector. That is,
  - Set $s^R_{k,l} := s^R_{k,l} + s^*_{opt}$. (resp. $s^P_{k,l} := s^P_{k,l} + s^*_{opt}$)
  - Set $s^R_{l,m} := s^R_{l,m} - s^*_{opt}$. (resp. $s^P_{l,m} := s^P_{l,m} - s^*_{opt}$)
  - Set $i := i + 1$.

end

(continued on next page..)

The re-timing heuristic (H) using scenarios (continued)

StartCP

17 **for** i *from* 1 **to** `numberOfCrew` **do**

18 Pick flight string i.

19 **for** j *from* 1 **to** `numSlackOptions` **do**

20 - Set s_{opt} :=slackOptions[j].

21 - Find the flight k that precedes flight l in string i.

22 - Find the flight m that follows flight l in string i.

if $(\hat{s}^R_{k,l} + s_{opt}) \geq 0$ **and** $(\hat{s}^P_{k,l} + s_{opt}) \geq 0$ **and** $(\hat{s}^R_{l,m} - s_{opt}) \geq 0$ **and** $(\hat{s}^P_{l,m} - s_{opt}) \geq 0$ **then**

23 - Set $\hat{s}^P_{k,l} := \hat{s}^P_{k,l} + s_{opt}$. (resp. $\hat{s}^R_{k,l} := \hat{s}^R_{k,l} + s_{opt}$)

24 - Set $\hat{s}^P_{l,m} := \hat{s}^P_{l,m} - s_{opt}$. (resp. $\hat{s}^R_{l,m} := \hat{s}^R_{l,m} - s_{opt}$)

25 **forall the** $\omega \in \Omega$ **do**

Run the Evaluation Algorithm (4.4.1) using the updated slack vector $\hat{s}$ and store the corresponding total delay.

end

26 - Calculate the average delay over all the simulations and record this average delay with its corresponding slack option.

end

end

27 - Choose the best delay/ slack option (s^*_{opt}) and make the appropriate changes in the real slack vector. That is,

28 -$s^P_{k,l} := s^P_{k,l} + s^*_{opt}$ (resp. $s^R_{k,l} := s^R_{k,l} + s^*_{opt}$)

29 -$s^P_{l,m} := s^P_{l,m} - s^*_{opt}$ (resp. $s^R_{l,m} := s^R_{l,m} - s^*_{opt}$)

30 -Set $i := i + 1$

end

31 -Set $l := l + 1$. (Move on to the next flight for each string)

32 Return to 1.

6.4 Iterative Algorithms Utilising simSub1, simSub2 And The Improved Heuristic (H)

We now outline the three algorithmic approaches that will be compared in this Chapter: the IPD approach (i.e. the approach outlined in Chapter 4) that uses mean delays combined with re-timing, the integrated aircraft routing and crew pairing simSub1 approach combined with re-timing and the integrated aircraft routing and crew pairing simSub2 approach combined with re-timing as outlined below. Note that we will use the improved re-timing algorithm referred to as (H) in all three cases.

Algorithm 6.4.1: Integrated AR, CP and re-timing: (IPD + H).

1. Solve the integrated aircraft routing and crew pairing problem using the approach outlined in Algorithm 4.4.2.
2. Apply the re-timing heuristic (H), algorithm (6.3.1) to the incumbent AR and CP solutions.
3. Apply the propagated delay algorithm (4.4.1) to the new solution to obtain the total delay.

Algorithm 6.4.2: Integrated AR, CP and re-timing: (simSub1 + H).

1. Solve the integrated aircraft routing and crew pairing problem using the simSub1 approach in each subproblem; as outlined in Algorithm 6.1.5.
2. Apply the re-timing heuristic (H), algorithm (6.3.1) to the incumbent AR and CP solutions.
3. Apply the propagated delay algorithm (4.4.1) to the new solution to obtain the total delay.

Algorithm 6.4.3: Integrated AR, CP and re-timing: (simSub2 + H).

1. Solve the integrated aircraft routing and crew pairing problem using the simSub1 approach in each subproblem; as outlined in Algorithm 6.1.6.
2. Apply the re-timing heuristic (H), algorithm (6.3.1) to the incumbent AR and CP solutions.
3. Apply the propagated delay algorithm (4.4.1) to the new solution to obtain the total delay.

6.4.1 Example Solutions

In Section 6.5 we outline a further improvement on the algorithms presented on the previous page and for ease of comparison, tabulate the results together with this further improvement in Section 6.6. Below we present a few example solutions that illustrate the effectiveness of Algorithms 6.4.1, 6.4.2 and 6.4.3. In each figure, we evaluated the performance of each re-timing approach over 1000 different delay scenarios drawn from the same distribution as the particular instance. It may be observed that the re-timed solution not only experiences less total delay, but additionally appears to possess less variability in the solution. This is particularly evident in Figure 6.2 below. In Section 6.6 we will confirm that on the test instances used, the re-timed solution appears to result in less variability over the 1000 delay scenarios used.

Truncated Normal Distribution With $\mu = 10$, $\sigma = 5$ (Instance 12).

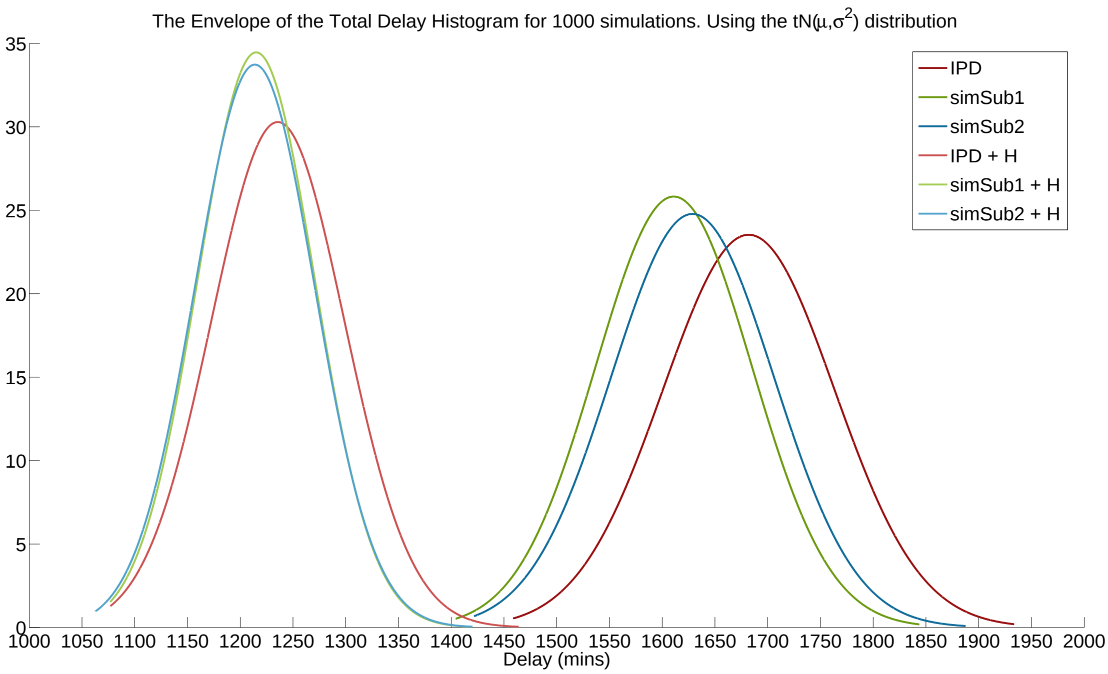
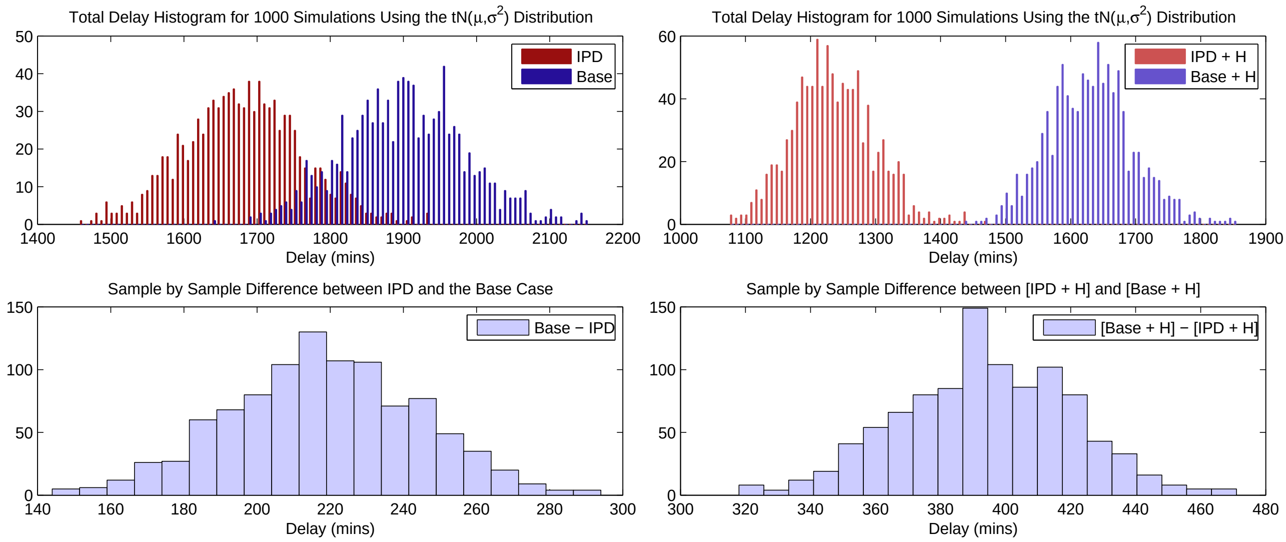
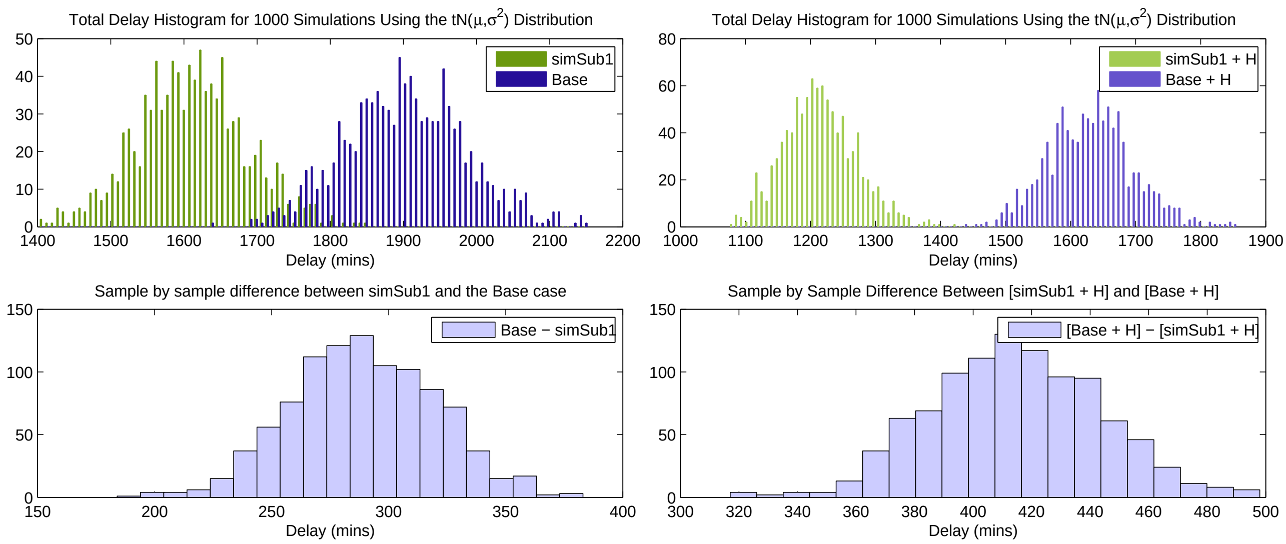


Figure 6.2: The envelope of the total delay histogram for each approach over the 1000 delay scenarios.



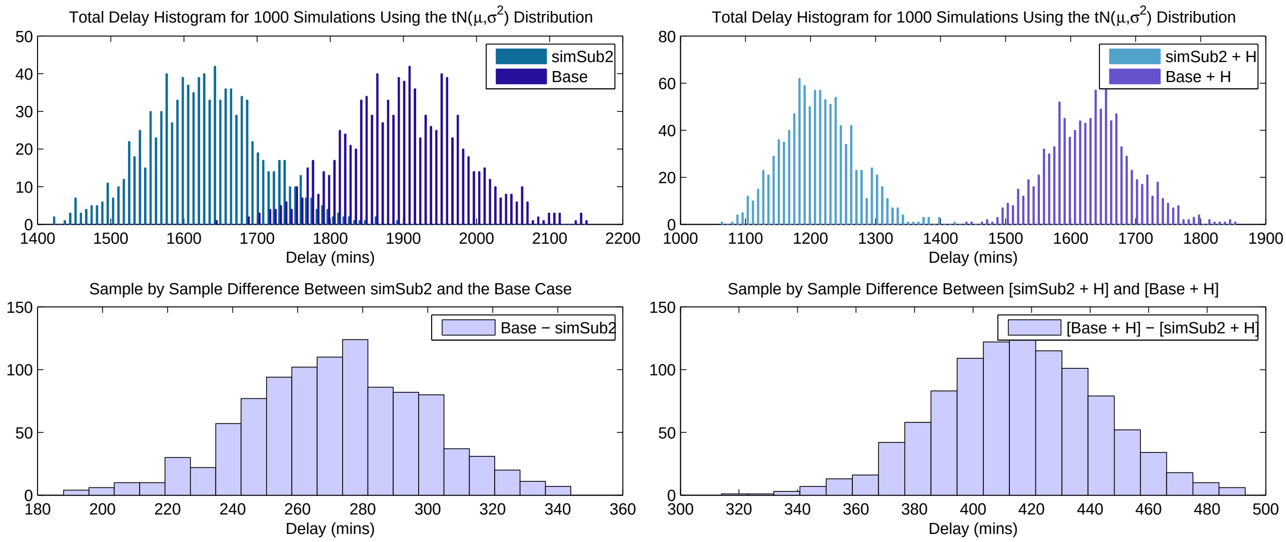
(a) Prior to Re-timing (b) After Re-timing

Figure 6.3: The Base Case and the IPD case, prior to re-timing and after re-timing.



(a) Prior to Re-timing (b) After Re-timing

Figure 6.4: The Base Case and the simSub1 case, prior to re-timing and after re-timing.



(a) Prior to Re-timing (b) After Re-timing

Figure 6.5: The Base Case and the simSub2 case, prior to re-timing and after re-timing.

Exponential Distribution With $\lambda = 10$ (Instance 5).

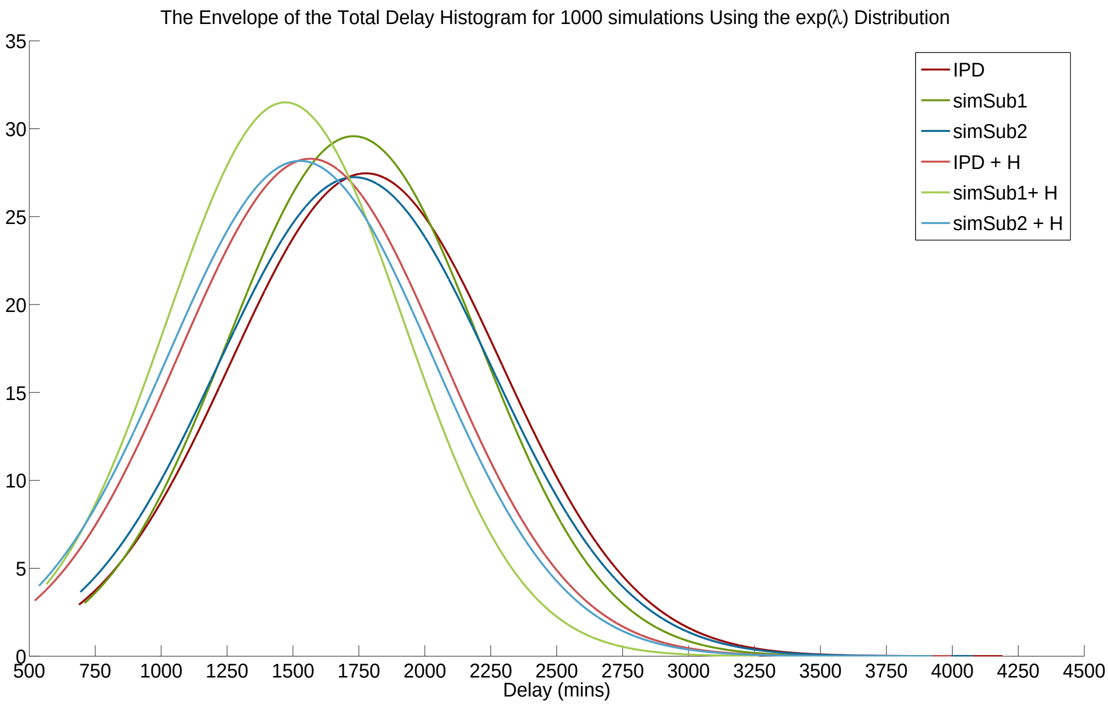
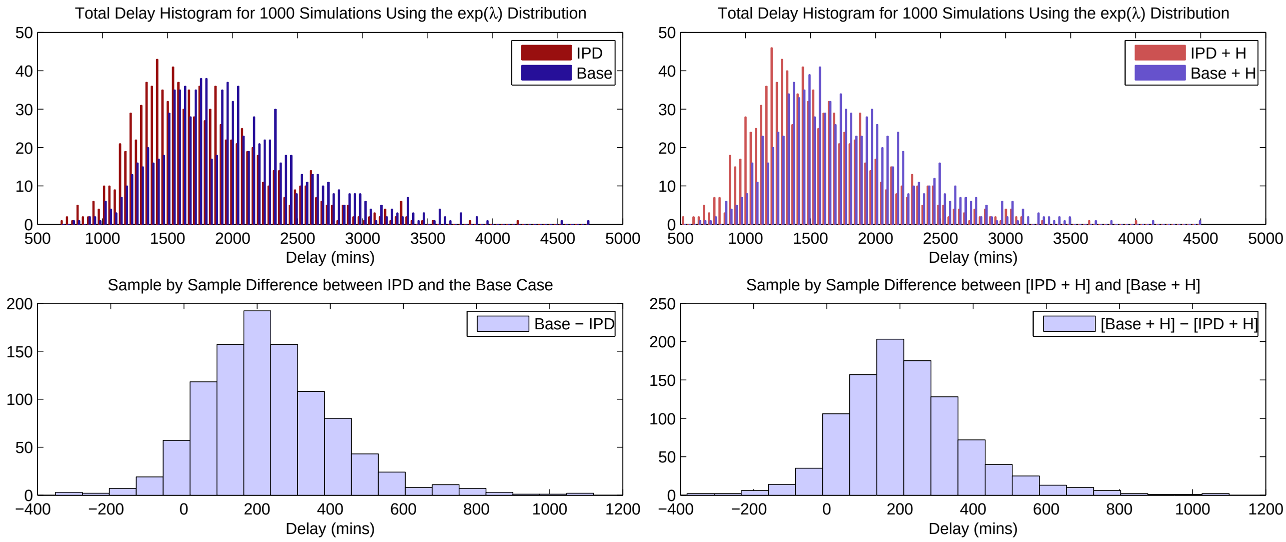
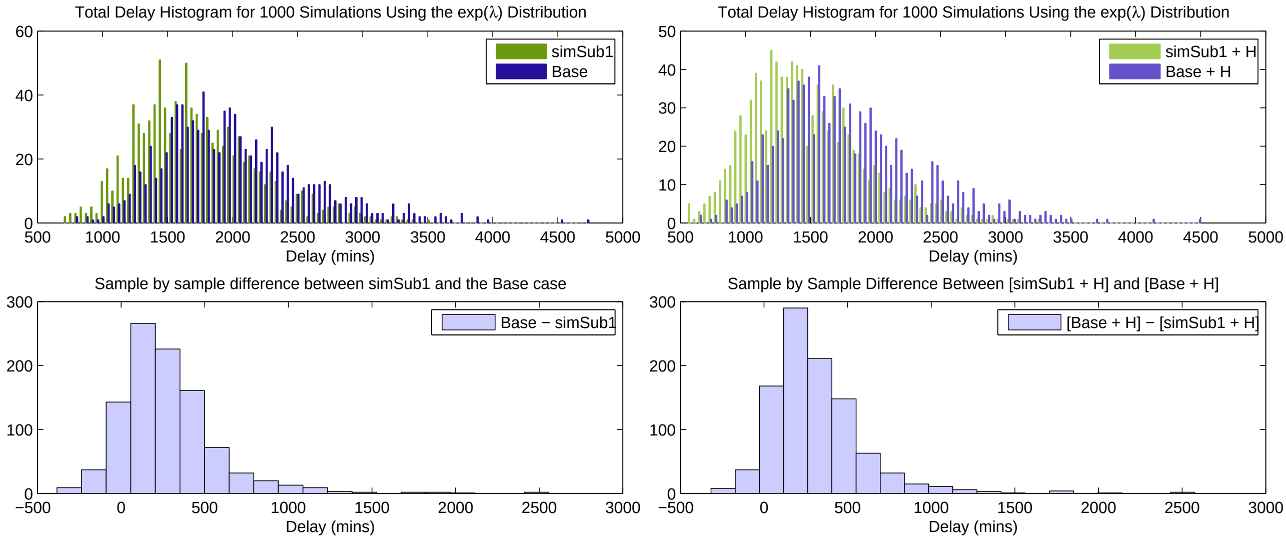


Figure 6.6: The envelope of the total delay histogram for each approach over the 1000 delay scenarios



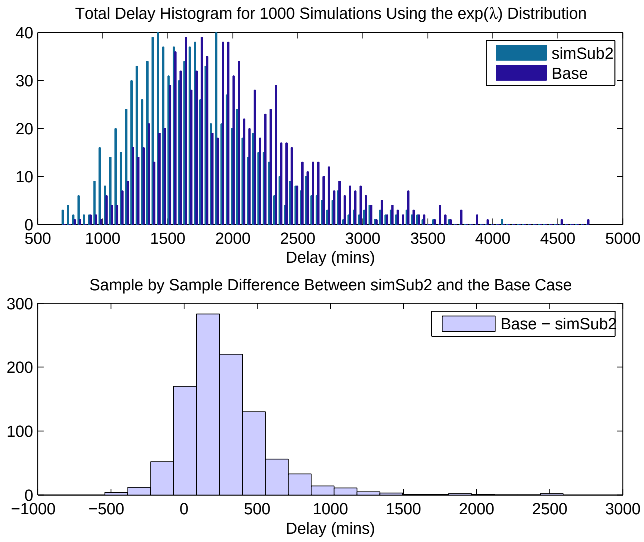
(a) Prior to Re-timing (b) After Re-timing

Figure 6.7: The Base Case and the IPD case, prior to re-timing and after re-timing

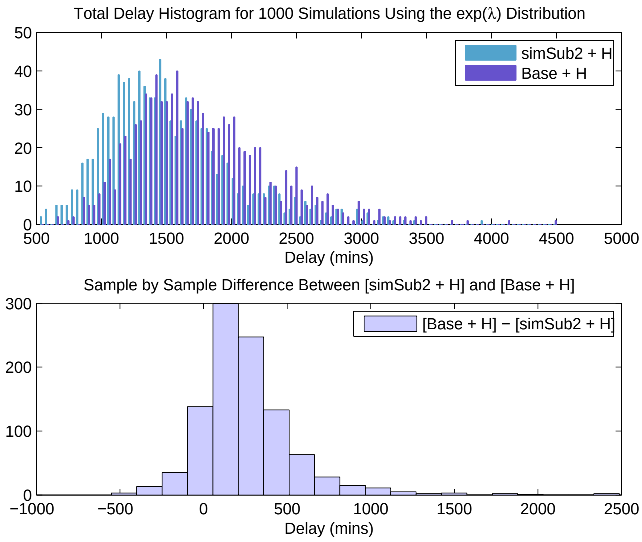


(a) Prior to Re-timing (b) After Re-timing

Figure 6.8: The Base Case and the simSub1 case, prior to re-timing and after re-timing



(a) Prior to Re-timing



(b) After Re-timing

Figure 6.9: The simSub2 and the IPD case, prior to re-timing and after re-timing

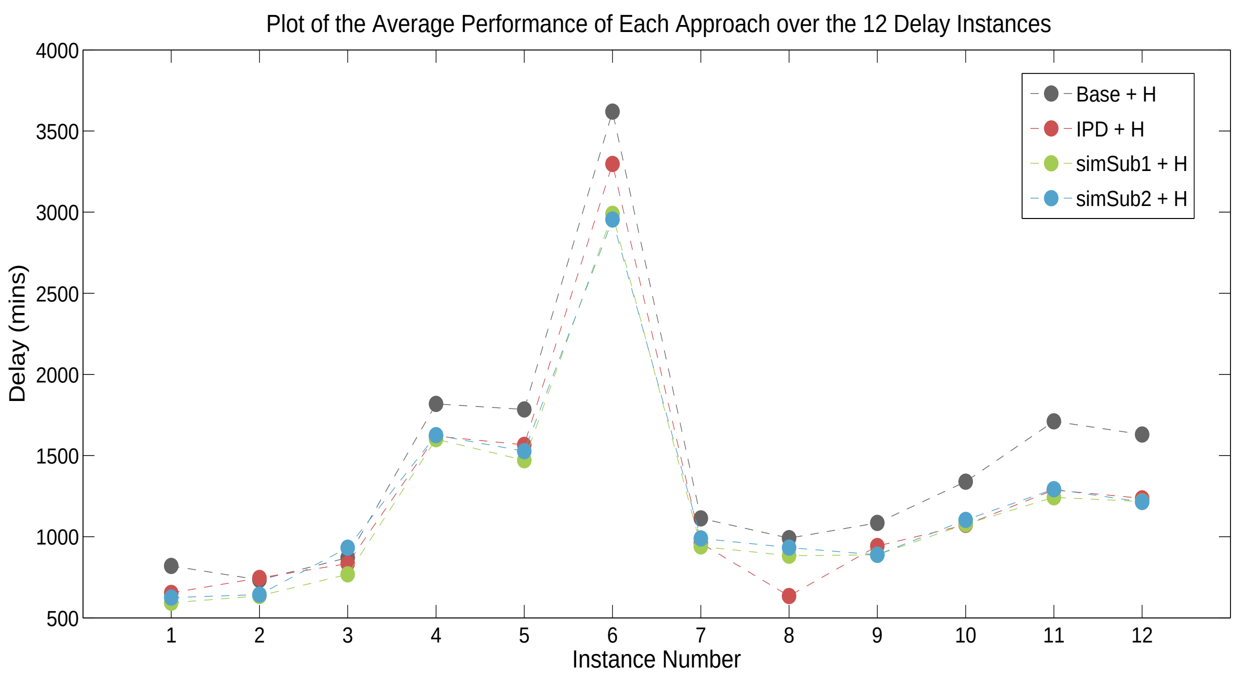


Figure 6.10: Plot of the Average Performance of Each Approach over the 12 Delay Instances.

6.5 Further Improvements: Re-solving The Integrated Aircraft Routing And Crew Pairing Problem Using The New Departure Times

In this Chapter we have observed that it is possible to improve upon the performance of both the integrated aircraft routing and crew pairing problems and the heuristic via the inclusion of multiple delay scenarios. As was mentioned in Section 6.1, one advantage of the heuristic is that it does not introduce extra complexity within the problem and may be easily embedded within an iterative scheme; in which one iterates between the integrated aircraft routing and crew pairing model of Chapter 4, re-times using the heuristic and then re-solves the integrated aircraft routing and crew pairing problem using the new departure times (as determined by the heuristic). If desired, this process may be extended further, until a specified level of improvement is obtained.

In this Section we outline a simple iterative scheme that incorporates the integrated aircraft routing and crew pairing problem outlined in Chapter 4 and utilises the improved heuristic outlined in Section 6.3 above, in an attempt to achieve further improvements using these two solution techniques. Since the results above indicate that the simSub1 approach and simSub2 approach have the capacity to outperform the IPD approach, we will additionally investigate the performance of these approaches when used within such an iterative scheme. Thus we investigate three alternative iterative schemes.

We outline the iterative schemes and list the results for each delay instance below. We will denote the use of the improved heuristic by (H) and the re-solving of the integrated aircraft routing and crew pairing model, with modified departure times, by (R). Thus the notation:

Approach $+ H$:	Denotes the specified approach, re-timed using the improved heuristic.
Approach $+ H + R$:	Denotes the specified approach, re-timed using the improved heuristic and re-solved using the modified departure times from the heuristic.

Algorithm 6.5.1: Integrated AR, CP and re-timing: (IPD + H + R).

1. Solve the integrated aircraft routing and crew pairing problem using the IPD approach as outlined in Algorithm 4.4.2.
2. Apply the re-timing heuristic (H) (6.3.1) to the incumbent AR and CP solutions.
3. Apply the Propagated delay algorithm (4.4.1) to the new solution to obtain the total delay.
4. Re-solve (R) the integrated aircraft routing and crew pairing problem with the new re-timing for each flight. Use the standard approach in each subproblem; as outlined in Algorithm 4.4.2.
5. Apply the propagated delay algorithm (4.4.1) to the new solution to obtain the total delay.

Algorithm 6.5.2: Integrated AR, CP and re-timing: (simSub1 + H + R).

1. Solve the integrated aircraft routing and crew pairing problem using the simSub1 approach in each subproblem; as outlined in Algorithm 6.1.5.
2. Apply the re-timing heuristic (H) (6.3.1) to the incumbent AR and CP solutions.
3. Apply the Propagated delay algorithm (4.4.1) to the new solution to obtain the total delay.
4. Re-solve (R) the integrated aircraft routing and crew pairing problem with the new re-timing for each flight. Use the simSub1 approach in each subproblem as outlined in Algorithm 6.1.5.
5. Apply the propagated delay algorithm (4.4.1) to the new solution to obtain the total delay.

Algorithm 6.5.3: Integrated AR, CP and re-timing: (simSub2 + H + R).

1. Solve the integrated aircraft routing and crew pairing problem using the simSub2 approach in each subproblem; as outlined in Algorithm 6.1.6.
2. Apply the re-timing heuristic (H) (6.3.1) to the incumbent AR and CP solutions.
3. Apply the propagated delay algorithm (4.4.1) to the new solution to obtain the total delay.
4. Re-solve (R) the integrated aircraft routing and crew pairing problem with the new re-timing for each flight. Use the simSub2 approach in each subproblem as outlined in Algorithm 6.1.6.
5. Apply the Propagated delay algorithm (4.4.1) to the new solution to obtain the total delay.

6.6 Numerical Results

In this section we tabulate the results for each of the approaches discussed thus far. Each table compares the Base Case with each Approach + H and Approach + H + R. Additionally, for ease of comparison, we provide results for our final approach outlined in the final chapter (Chapter 7). Our final approach involves the integration of all three components solved simultaneously, (i.e. aircraft routing, crew pairing and re-timing simultaneously performed within one model) and we use the acronym ARCPR to refer to the Aircraft Routing, Crew Pairing and Retiming approach. In each table we provide the time taken to solve each approach (in seconds) and for approach, we specify the breakdown of time for each component explicitly.

6.6.1 Exponential Distribution With $\lambda = 5$.

Instance 1

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	515	439	954	–	9.17
Base + H.	466	354	820	14.05	$9.17 + 15.11$
IPD + H.	370	283	653	31.55	$47.19 + 17.06$
simSub1 + H.	350	243	593	37.84	$56.19 + 17.05$
simSub2 + H.	365	260	625	34.50	$50.00 + 17.05$
IPD + H + R.	329	276	605	36.58	$47.19 + 17.06 + 44.22$
simSub1 + H + R.	350	243	593	37.84	$56.19 + 17.05 + 53.25$
simSub2 + H + R.	365	260	625	32.50	$50.00 + 17.05 + 50.12$
ARCPR (3 iter.)	218	206	424	55.56	7280

Instance 2

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	473	478	951	–	9.10
Base + H.	353	381	734	22.82	$9.10 + 16.19$
IPD + H.	336	410	746	21.56	$68.01 + 19.54$
simSub1 + H.	312	323	635	33.23	$75.58 + 20.00$
simSub2 + H.	308	335	643	32.39	$71.11 + 19.00$
IPD + H + R.	317	336	653	31.33	$68.01 + 19.54 + 70.00$
simSub1 + H + R.	312	323	635	33.23	$75.58 + 20.00 + 76.22$
simSub2 + H + R.	308	335	643	32.39	$71.11 + 19.00 + 70.34$
ARCPR (3 iter.)	321	243	564	40.69	7560

Instance 3

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	519	550	1069	–	10.15
Base + H.	444	427	871	18.52	$10.15 + 16.01$
IPD + H.	411	423	834	21.98	$63.45 + 17.55$
simSub1 + H.	398	370	768	28.16	$69.12 + 18.39$
simSub2 + H.	460	472	932	12.82	$68.44 + 17.20$
IPD + H + R.	287	322	609	43.03	$63.45 + 17.55 + 66.37$
simSub1 + H + R.	331	324	654	38.82	$69.12 + 18.39 + 72.23$
simSub2 + H + R.	287	322	609	43.03	$68.44 + 17.20 + 70.09$
ARCPR (2 iter.)	280	241	521	51.26	7200

6.6.2 Exponential Distribution With $\lambda = 10$

Instance 4

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	1056	1075	2131	–	8.12
Base + H.	929	889	1818	14.70	$8.12 + 15.48$
IPD + H.	868	752	1620	23.98	$72.26 + 16.58$
simSub1 + H.	874	727	1601	24.87	$80.47 + 18.03$
simSub2 + H.	872	754	1626	23.70	$79.20 + 18.00$
IPD + H + R.	775	685	1460	31.49	$72.26 + 16.58 + 73.09$
simSub1 + H + R.	806	719	1525	28.44	$80.47 + 18.03 + 82.27$
simSub2 + H + R.	811	767	1579	25.90	$79.20 + 18.00 + 81.36$
ARCPR (4 iter.)	529	563	1092	48.76	8460

Instance 5

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	954	1057	2011	–	15.53
Base + H.	867	918	1784	11.29	15.53 + 15.03
IPD + H.	755	811	1566	22.13	214.19 + 16.24
simSub1 + H.	746	724	1471	26.85	228.65 + 18.41
simSub2 + H.	766	761	1527	24.07	220.29 + 17.54
IPD + H + R.	712	768	1480	26.40	214.19 + 16 + 182.17
simSub1 + H + R.	746	724	1471	26.85	228.65 + 18.41 + 210.00
simSub2 + H + R.	722	777	1498	25.51	220.29 + 17.54 + 188.63
ARCPR (2 iter.)	494	575	1069	48.84	7380

Instance 6

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	1871	1890	3761	–	14.51
Base + H.	1836	1784	3620	3.75	14.51 + 16.04
IPD + H.	1710	1587	3297	12.34	92.35 + 18.25
simSub1 + H.	1578	1412	2990	20.50	104.39 + 18.50
simSub2 + H.	1538	1417	2955	21.43	96.49 + 18.25
IPD + H + R.	1605	1602	3208	14.70	92.35 + 18.25 + 90.89
simSub1 + H + R.	1512	1458	2970	21.03	104.39 + 18.50 + 102.06
simSub2 + H + R.	1538	1417	2955	21.43	96.49 + 18.25 + 95.88
ARCPR (3 iter.)	1107	1206	2313	38.50	9720

6.6.3 Truncated Normal Distribution With $\mu = 5$, $\sigma = 10$

Instance 7

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	708	811	1519	–	14.19
Base + H.	593	520	1113	26.73	14.19 + 16.48
IPD + H.	509	451	960	36.80	39.44 + 17.78
simSub1 + H.	514	426	940	38.12	48.18 + 18.34
simSub2 + H.	524	465	989	34.89	40.86 + 18.12
IPD + H + R.	398	444	843	44.50	39.44 + 17.78 + 38.30
simSub1 + H + R.	431	456	886	41.67	48.18 + 18.34 + 50.01
simSub2 + H + R.	413	457	870	42.73	40.86 + 18.12 + 40.00
ARCPR (2 iter.)	412	308	720	52.60	8424

Instance 8

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	576	504	1080	–	13.31
Base + H.	569	422	991	8.24	13.31 + 15.87
IPD + H.	416	218	634	41.30	168.74 + 18.23
simSub1 + H.	548	335	883	18.24	175.22 + 18.82
simSub2 + H.	571	377	948	13.61	172.35 + 17.43
IPD + H + R.	416	218	634	41.30	168.74 + 18.23 + 168
simSub1 + H + R.	414	331	746	30.93	172.35 + 17.43 + 173
simSub2 + H + R.	413	347	760	29.63	172.35 + 17.43 + 170
ARCPR (3 iter.)	356	265	621	42.50	9396

Instance 9

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	723	803	1527	–	15.11
Base + H.	574	511	1085	28.95	15.11 + 16.29
IPD + H.	549	393	942	38.31	57.98 + 16.00
simSub1 + H.	523	365	888	41.85	64.12 + 17.53
simSub2 + H.	523	365	888	41.85	64.02 + 17.07
IPD + H + R.	443	446	889	41.78	57.98 + 16.00 + 52.83
simSub1 + H + R.	523	365	888	41.85	64.12 + 17.53 + 60.46
simSub2 + H + R.	523	365	888	41.85	64.02 + 17.07 + 62.68
ARCPR (3 iter.)	370	321	691	54.75	10296

6.6.4 Truncated Normal Distribution With $\mu = 10$, $\sigma = 5$

Instance 10

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	781	990	1771	–	14.91
Base + H.	622	717	1339	24.39	14.91 + 16.45
IPD + H.	468	605	1073	39.41	53.35 + 18.26
simSub1 + H.	516	561	1077	39.19	61.64 + 18.02
simSub2 + H.	525	578	1103	37.72	54.12 + 18.45
IPD + H + R.	475	550	1025	42.12	53.35 + 18.26 + 55.28
simSub1 + H + R.	451	542	993	43.93	61.64 + 18.02 + 60.69
simSub2 + H + R.	425	447	873	50.71	54.12 + 18.45 + 53.06
ARCPR (3 iter.)	486	515	1001	43.48	6840

Instance 11

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	827	1171	1999	–	14.70
Base + H.	751	960	1711	14.40	14.70 + 16.71
IPD + H.	572	716	1288	35.57	61.72 + 18.22
simSub1 + H.	532	711	1243	37.82	65.35 + 17.48
simSub2 + H.	558	735	1293	35.32	62.23 + 17.30
IPD + H + R.	520	647	1167	41.62	61.72 + 18.22 + 66.19
simSub1 + H + R.	532	711	1243	37.82	65.35 + 17.48 + 65.00
simSub2 + H + R.	534	674	1208	39.57	62.23 + 17.30 + 60.64
ARCPR (2 iter.)	535	660	1195	40.22	4320

Instance 12

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	894	1006	1901	–	15.04
Base + H.	814	816	1630	14.26	15.04 + 17.37
IPD + H.	622	614	1236	34.98	72.97 + 16.54
simSub1 + H.	639	577	1216	36.03	82.66 + 17.30
simSub2 + H.	646	568	1214	36.13	74.22 + 16.54
IPD + H + R.	541	576	1117	41.24	72.97 + 16.54 + 76.80
simSub1 + H + R.	573	553	1126	40.77	82.66 + 17.30 + 85.77
simSub2 + H + R.	542	541	1083	43.03	74.22 + 16.54 + 70.32
ARCPR (4 iter.)	547	662	1209	36.40	10440

Table 6.2: Relative improvements of each [Approach + H] over Approach.

Instance	$\frac{(B-[B+H])}{B}\times 100\%$	$\frac{(IPD-[IPD+H])}{IPD}\times 100\%$	$\frac{(simSub1-[simSub1+H])}{simSub1}\times 100\%$	$\frac{(simSub2-[simSub2+H])}{simSub2}\times 100\%$
1	14.04	19.48	22.28	21.18
2	22.82	10.77	20.03	21.49
3	18.52	17.43	16.79	0.32
4	14.69	14.24	13.69	12.58
5	11.29	11.82	14.97	11.88
6	3.75	1.99	10.69	11.84
7	26.72	29.73	28.24	27.17
8	8.24	33.26	3.81	0.00
9	28.95	27.98	30.35	30.35
10	24.39	31.70	29.70	29.02
11	14.41	27.48	28.07	25.73
12	14.26	26.52	24.52	25.48
Average	16.84	21.03	20.26	18.09

Table 6.3: Relative improvements of each [Approach + H + R] over [Approach + H].

Instance	$\frac{([IPD+H]-[IPD+H+R])}{IPD+H}\times 100\%$	$\frac{([simSub1+H]-[simSub1+H+R])}{simSub1+H}\times 100\%$	$\frac{([simSub2+H]-[simSub2+H+R])}{simSub2+H}\times 100\%$
1	7.35	0.00	0.00
2	12.47	0.00	0.00
3	26.98	14.84	34.65
4	8.64	8.11	7.87
5	5.49	0.00	1.90
6	2.70	0.67	0.00
7	12.19	5.74	12.03
8	0.00	15.52	19.83
9	5.62	0.00	0.00
10	4.47	7.80	20.85
11	9.39	0.00	6.57
12	9.63	7.40	10.79
Average	8.74	5.01	9.54

Table 6.4: Relative improvements in standard deviation for each [Approach + H] over [Approach].

Instance	$\frac{([\mathsf{IPD+H}]-[\mathsf{IPD+H+R}])}{\mathsf{IPD+H}}\times 100\%$	$\frac{([\mathsf{simSub1+H}]-[\mathsf{simSub1+H+R}])}{\mathsf{simSub1+H}}\times 100\%$	$\frac{([\mathsf{simSub2+H}]-[\mathsf{simSub2+H+R}])}{\mathsf{simSub2+H}}\times 100\%$
1	12.06	13.38	12.19
2	28.05	14.36	23.40
3	11.78	9.90	11.14
4	7.06	5.44	5.57
5	5.25	8.18	5.98
6	4.15	4.09	4.66
7	24.56	23.64	25.00
8	27.08	28.26	33.33
9	24.50	27.03	27.08
10	27.84	26.25	25.61
11	20.55	21.62	22.67
12	23.36	24.32	25.97
Average	18.02	17.21	18.55

6.6.5 Discussion

In Table 6.2 we record the relative improvement of each Approach + H over each corresponding Approach to gain an insight into the effectiveness of the improved re-timing heuristic. The results indicate that over the 12 test instances, each of the approaches: B, IPD, simSub1 and simSub2 benefit significantly from the improved heuristic re-timing algorithm. The Base + H approach achieves a 16.84% improvement over the Base approach and the simSub2 + H approach the next largest improvement of 18.09% over the simSub2 approach.

The IPD + H and simSub1 + H approaches appear to perform (almost) equally well, with the IPD + H approach achieving a 21.03% improvement over the IPD approach and the simSub1 + H slightly lower, with a 20.26% improvement over the simSub1 approach. However, this is not a universal improvement over all test instances, as it may be seen from the table that for different instances the simSub1 + H approach out-performs the IPD + H approach and vice versa. This may be due to the fact that the heuristic preserves the incumbent aircraft and crew assignments, and therefore individual aircraft routes or crew pairings in the IPD approach may allow for greater improvements once the flight departure times are allowed to be adjusted.

It may be observed in Table 6.4 that using the improved heuristic in conjunction with each of the three approaches appears to produce solutions with not only smaller average total delay, but with less variance when exposed to 1000 different delay scenarios. This may be seen particularly clearly in Figure 6.2 of section 6.4.1. The re-timing produces an average improvement of 18.02% for the IPD approach, and improvement of 17.21% for the simSub1 approach and an improvement of 18.55% for the simSub2 approach.

Table 6.3 records the relative improvement of each Approach + H + R over Approach + H to determine the effectiveness of re-solving after heuristic re-timing. From the table, it may be observed that re-solving has the potential to be very effective. In particular, for instance 3, the IPD + H, simSub1 + H and simSub2 + H experience an improvement of 26.98%, 14.84% and 34.65% respectively. The simSub2 + H approach appears to benefit most from the re-timing with an average improvement of 9.54%, closely followed by the IPD + H approach, with an improvement of 8.74%. The simSub1 + H approach received the least benefit, with an average improvement of 5.01%. This stems from the fact that it had the most number of solutions that could not be improved further through re-solving; however, in 4/5 cases for which this occurred, the solution using simSub1 + H either equalled or out-performed the solution obtained using simSub2 + H + R.

It is interesting to note that for a few delay instances, the Approach + H + R only improves slightly upon, or equals the Approach + H results. This may be perhaps due to the particular delay instance(s) chosen. An additional set of experiments were run using 20 “batches” of 1000 delay scenarios for each instance, to provide a more statistically significant set of results. From these experiments (See Appendix), it may be be observed that the Approach + H + R consistently outperforms Approach + H. Thus for larger sample sizes, we would expect the Approach+H+R approach to out-perform the Approach + H approach on average.

A natural question is: How well do these methods scale as the network size increases? We explore this question briefly in the Appendix. We provide results for one delay instance for a network consisting of 320 flights with 818 feasible connections and assume a one-day schedule. We determine that the minimum number of aircraft and crew pairs required to cover this network are 64 and 102 crew, respectively. It may be observed that although each the algorithms take quite a bit longer to solve, the solutions are nonetheless consistent with the results achieved in this Chapter. Each of the Approach + H methods achieve approximately $23 - 28\%$ improvement over the Base Case and the Approach + H + R methods improving further still with approximately a $42 - 44\%$ improvement over the Base Case. Finally, The ARCPR approach outperforms all others with an improvement of 48.69%. However, the ARCPR method was also the slowest - with around 12.5 hours required to solve this problem; whereas the Approach + H + R methods solved in just under an hour. Thus, if it is preferable to obtain a solution quickly, these approaches provide a relatively favourable trade-off between solution quality and time required for solution.

CHAPTER

SEVEN

Integration Of Aircraft Routing, Crew Pairing And Re-timing: Using Scenarios In The Subproblem

7.1 Introduction And Motivation

In the previous Chapters, we outlined two different approaches for including scheduling decisions in the planning process. We observed in Chapters 4, 5 and 6, that these algorithms deliver a marked improvement over the standard sequential approaches, and their primary advantage, computationally, is that they do not introduce any additional complexity to the problem. That is, the integrated problem proposed above, scales as a standard aircraft routing or crew pairing problem. One of the drawbacks however, is that the algorithm only re-times a fixed aircraft and crew assignment, and so it may be possible to obtain better quality solutions by integrating this scheduling process directly with the integrated aircraft routing and crew pairing problem. That is, include all three decisions – aircraft routing, crew pairing and re-timing – within one problem and solve them simultaneously.

With this idea in mind, it would be preferable if it were somehow possible to embed the scheduling decisions within the integrated aircraft routing and crew pairing problem of Dunbar *et al.* (2011) [35], without increasing the complexity unnecessarily.

In this Chapter, we outline a new model capable of embedding the scheduling decisions within the iterative, integrated framework of [35] whilst retaining the simple form of the aircraft routing and crew pairing formulations. We propose to do this via an expanded network consisting of flight copies that represent different flight departure time choices. This expanded network exists only within the subproblem, and so does not increase the dimension (or complexity) of either the aircraft routing or crew pairing master problems. A more detailed description on how this is achieved can be found in the Sections that follow. We now outline a few key assumptions.

7.2 Notes And Assumptions

Firstly, to ensure that our model reflects real world restrictions on slot times and that the original fleet assignment remains feasible, we only attempt to re-time flights to within a 10 minute window $[t-10, t+10]$ either side of the scheduled departure time, t. We create for each flight i, five duplicate flight nodes i_1, i_2, i_3, i_4, i_5 that lie within this discretised minute window. For example, as we have 54 flights we will now have $54 \times 5 = 270$ (+2 corresponding to the source and sink) flight nodes in our subproblem.

As mentioned in the introduction, this process of duplicating flight nodes is performed only within the subproblem, so as to ensure that we do not increase the complexity of the integrated iterative aircraft routing and crew pairing master problems. The process of duplication is done within a pre-processing step, and does not need to be repeated with each call to the subproblem. Moreover, we only perform this duplication within the *aircraft* routing subproblem, and assume that the schedule (departure times) chosen is also followed by the crew; thus eliminating potential conflicts between the aircraft and crew schedules. If one was to additionally allow for this duplication in the crew subproblem, an additional constraint would need to be introduced into the master problem to check for potential conflicts. To ensure a fair comparison with the model in [35], we do not wish to introduce any “new” arcs into the network. Rather, we only allow connections between flights nodes that correspond to connections in the original network. More specifically, if connection (i, j) was a feasible connection in the original network, we must allow all connection possible pairs (i_m, j_n) for $i = 1, \ldots, 5$, $j = 1, \ldots, 5$ in the new network (provided the slack across the connection is non-negative). To construct all the possible connections in the expanded network from the original set of connections, we make use of the following pre-processing algorithm:

7.3 The Pre-processing Algorithm

Algorithm 7.3.1: Pre-processing algorithm for the aircraft routing subproblem

1 Label the flight copies for flight i as i_1,i_2,i_3,i_4,i_5 where i_3 corresponds to the original flight i.

2 Label the flight copies for flight j as j_1,j_2,j_3,j_4,j_5 where j_3 corresponds to the original flight j.

Suppose connection (i, j) in the original network has primary delay p_{ij} and slack s_{ij}.
Let $\mathcal{A}$ be the original arc set and $\mathcal{A}^*$ be the new (expanded) arc set.

3 Set $\mathcal{A}^* = \emptyset$ initially.

forall the $(i, j) \in \mathcal{A}$ **do**
    for m *from* 1 *to* `numFlightCopies` **do**
        for n *from* 1 *to* `numFlightCopies` **do**
            if $(i, j) \in \mathcal{A}$ **and** (i_m, j_n) *has non-negative slack* **then**
                $\mathcal{A}^* := \mathcal{A}^* \cup (i_m, j_n)$.
            end
        end
    end
end

As an illustration, consider the original connection (i, j) below:

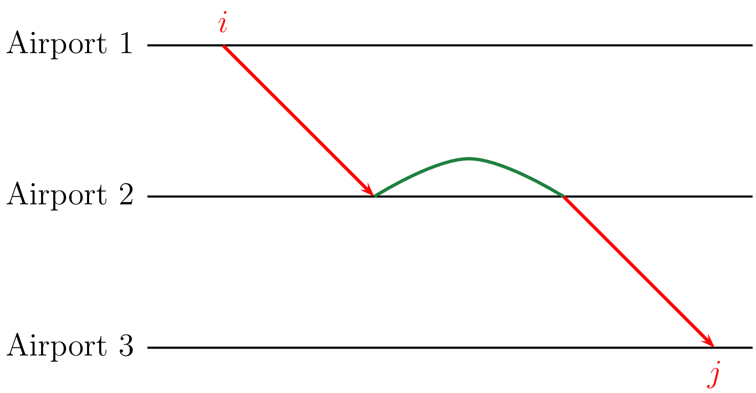


Figure 7.1: The connection (i, j) denoted by the green arc.

For each flight copy $i_1, i_2, \ldots, i_5$, we check whether we can connect to each of the flight copies $j_1, j_2, \ldots, j_5$. For example, it may be feasible to connect flight copy i_3 to any of $j_1, j_2, \ldots, j_5$, in which case we include the feasible connections $(i_3, j_1), (i_3, j_2), \ldots, (i_3, j_5)$ in our network.

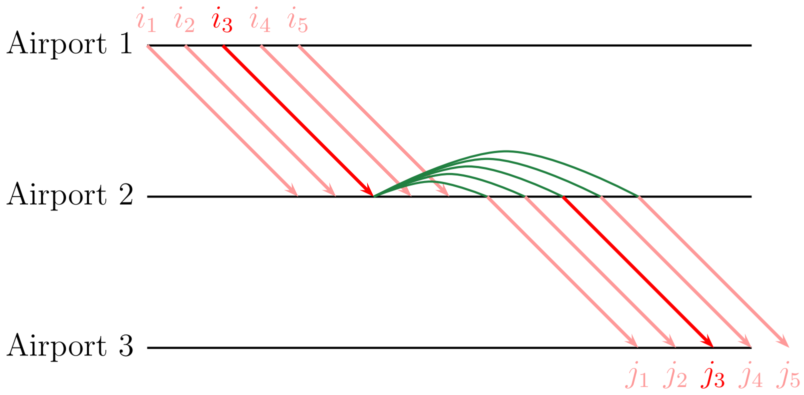


Figure 7.2: Possible outbound connections from flight i_3.

In some cases, it may be possible to form the complete, directed, bi-partite graph in Figure 7.3.

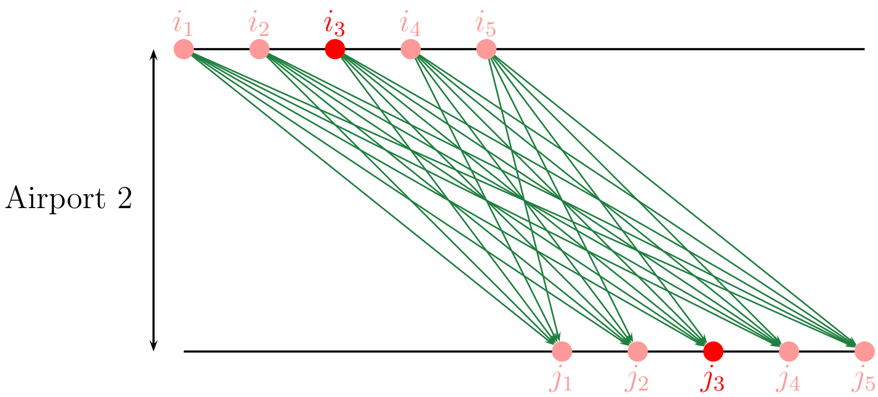


Figure 7.3: Possible connections between the flight copies for i and j at an airport.

For each of these new connections (i_m, j_n), we assume the primary delay p_{ij}, is the same as for the original connection (i, j). Similarly, we assume that each flight copy $i_1, i_2, \ldots, i_5$ inherits the dual value w_i as for flight i in the original network. Once the above pre-processing algorithm has been performed, along with the modifications to the primary delays and dual values, the aircraft routing subproblem is solved over this larger network using the label setting algorithm discussed in the previous sections.

The solution to the subproblem (an aircraft string) is of length $54 \times 5 = 270$, which is converted back into the appropriate length of 54 for the master problem using Algorithm 7.3.2 below. Note that in the conversion process, we label the choice of departure time and store this in a vector for future reference. We use the following labels $[1, 2, 3, 4, 5]$ to refer to the choices $T := [-10, -5, 0, 5, 10]$ respectively. We then keep track of the chosen departure times and use this to update the slack over the relevant connections. This new slack information, along with the chosen departure times, is then fed into the crew pairing problem. This is achieved as follows:

Algorithm 7.3.2: Subproblem solution conversion

Input: The expanded subproblem solution $\mathbf{x}_{old}$ and the departure time choices vector, T.

Output: The aircraft string for the master problem and a corresponding string containing the choice of flight departure time

Set `numFlightsExpanded` := `numFlights` × `numFlightCopies`

Create the vectors $\mathbf{x}_{new}$ and $\mathbf{z}$, each of length `numFlights`, corresponding to the subproblem solution (aircraft string) and the chosen departure times.

for *i from* 1 *to* `numFlightsExpanded` **do**

  $j = \lfloor \frac{(i-1)}{\texttt{numFlightCopies}} \rfloor + 1.$

  if $(x_{old}[i] = 1)$ **then**

    $x_{new}[j] = 1.$ `/* Store solution */`

    if $i \equiv 0 \bmod (\texttt{numFlightCopies})$ **then**

      $r_{pos} = \texttt{numFlightCopies}.$

      $z[j] = r_{pos}.$ `/* Label is equal to numFlightCopies */`

      forall the (i, j) **and** $(j, k) \in \mathcal{A}$ **do**

        $s^R_{i,j} := s^R_{i,j} + 10$ and $s^P_{i,j} := s^P_{i,j} + 10.$

        $s^R_{j,k} := s^R_{j,k} - 10$ and $s^P_{j,k} := s^P_{j,k} - 10.$

      end

    else

      $r_{pos} = i \bmod (\texttt{numFlightCopies}).$

      $z[j] = r_{pos}.$ `/* Labels: 1 to numFlightCopies-1 */`

      forall the (i, j) **and** $(j, k) \in \mathcal{A}$ **do**

        $s^R_{i,j} := s^R_{i,j} + T[r_{pos}]$ and $s^P_{i,j} := s^P_{i,j} + T[r_{pos}].$

        $s^R_{j,k} := s^R_{j,k} - T[r_{pos}]$ and $s^P_{j,k} := s^P_{j,k} - T[r_{pos}].$

      end

    end

  end

end

Example 7.3.1. *For example, the solution:*

$$(0,1,0,0,0,0,0,0,0,1,\cdots,1,0,0,0,0)^T_{1\times 270}$$

becomes:

$$\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}_{54\times 1} \quad \textit{with chosen departure times:} \quad \begin{pmatrix} 2 \\ 5 \\ \vdots \\ 1 \end{pmatrix}_{54\times 1}$$

□

Additionally, we wish to keep the chosen schedule as close as possible to the original schedule, so as to remain attractive to the airline – as extra costs may be associated with changes to the original schedule. We make the assumption that greater penalties would be associated with greater changes to the original schedule. We keep track of the changes to the schedule along each path in the label setting algorithm and penalise each change as follows:

If 1 is the chosen departure time: `penalty` = 5
If 2 is the chosen departure time: `penalty` = 1
If 3 is the chosen departure time: `penalty` = 0
If 4 is the chosen departure time: `penalty` = 1
If 5 is the chosen departure time: `penalty` = 5

We keep track of the penalty sum along each path (in a similar manner to the way in which we kept track of the weights w_i). We set an upper bound on the penalty sum of `maxPenaltySum`, and use this to help reduce the number of labels produced at each node in the label setting algorithm (in the same way as for the maximum crew time limits). We experimented with three different values for `maxPenaltySum`, namely 15, 25 and 30, with the time taken to solve the problem increasing with each value. The best results were achieved by setting the upper bound to be 30 and results for this choice of setting are listed in the Results Section below.

7.4 The Integrated Aircraft Routing, Crew Pairing And Re-timing Algorithm

In Section 6.3 we discovered that it was possible to improve the solution quality of our aircraft and crew solutions by incorporating simulations within the subproblem. In particular, `simSub1` and `simSub2` were two such proposed approaches. However, as our aircraft routing subproblem is now significantly larger than the original and we wish to solve the problem as quickly as possible, we make use of the `simSub2` approach in both the aircraft routing and crew pairing subproblems to avoid generating every possible path. Furthermore, we observed in Chapter 6 that the percentage improvement obtained by the simSub2 approach was not significantly smaller than that of the simSub1 approach.

Algorithm 7.4.1: Integrated AR and CP using simSub2 in each subproblem (ARCPR)

---

1 Solve the integrated aircraft routing and crew pairing problem using the Iterative Algorithm 4.4.2 outlined in Chapter 4, making the following changes:

(i) Solve the aircraft routing problem (4.1) on the *expanded flight network* via column generation, using the simSub2 Label Setting Algorithm 6.1.3.

(ii) Use Algorithm 7.3.2 to convert the solution to the aircraft routing subproblem into a form compatible with the master problem and update the slack information and departure times according to the choices made in the aircraft routing subproblem.

(iii) Solve the crew pairing problem (4.2) on the *original flight network* via column generation, using the simSub2 Label Setting Algorithm 6.1.4.

7.5 Numerical Results

In this section we tabulate the results for the ARCPR approach, calculating the improvement this approach provides over all other approaches discussed in this thesis. For the specific delay improvements, the reader is referred to Table 6.2 and 6.3 in Section 6.6.

Table 7.1: Relative improvements of the ARCPR approach over each Approach + H.

Instance	$\frac{([\text{IPD+H}]-[\text{ARCPR}])}{\text{IPD+H}}\times 100\%$	$\frac{([\text{simSub1+H}]-[\text{ARCPR}])}{\text{simSub1+H}}\times 100\%$	$\frac{([\text{simSub2+H}]-[\text{ARCPR}])}{\text{simSub2+H}}\times 100\%$
1	35.07	28.50	32.16
2	24.40	11.18	12.29
3	37.53	32.16	44.10
4	32.59	31.79	32.84
5	31.74	27.33	30.00
6	29.85	22.64	21.73
7	25.00	23.40	27.20
8	2.05	29.67	34.49
9	26.65	22.18	22.18
10	6.71	7.06	9.25
11	7.22	3.86	7.58
12	2.18	0.58	0.42
Average	21.75	20.03	22.85

Table 7.2: Relative improvements of the ARCPR approach over each Approach + H + R.

Instance	$\frac{([\text{IPD+H+R}]-[\text{ARCPR}])}{\text{IPD+H+R}}\times 100\%$	$\frac{([\text{simSub1+H+R}]-[\text{ARCPR}])}{\text{simSub1+H+R}}\times 100\%$	$\frac{([\text{simSub2+H+R}]-[\text{ARCPR}])}{\text{simSub2+H+R}}\times 100\%$
1	29.91	28.50	32.16
2	13.63	11.18	12.29
3	14.45	20.34	14.45
4	25.21	28.39	30.84
5	27.77	27.33	28.64
6	29.85	22.64	21.73
7	14.59	18.74	17.24
8	2.05	16.76	18.30
9	22.27	22.18	22.18
10	2.34	−0.80	−14.66
11	−2.40	3.86	1.08
12	−8.23	−7.37	−11.63
Average	14.31	15.98	14.39

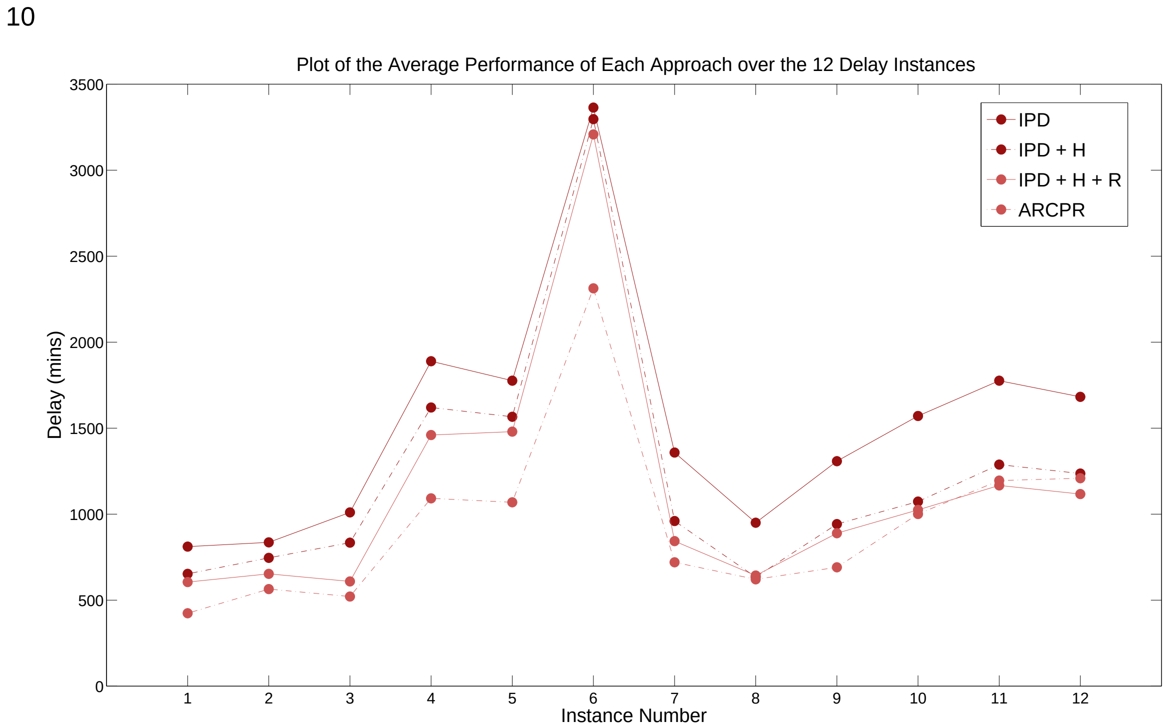


Figure 7.4: Plot of the average performance of each approach over the 12 delay instances.

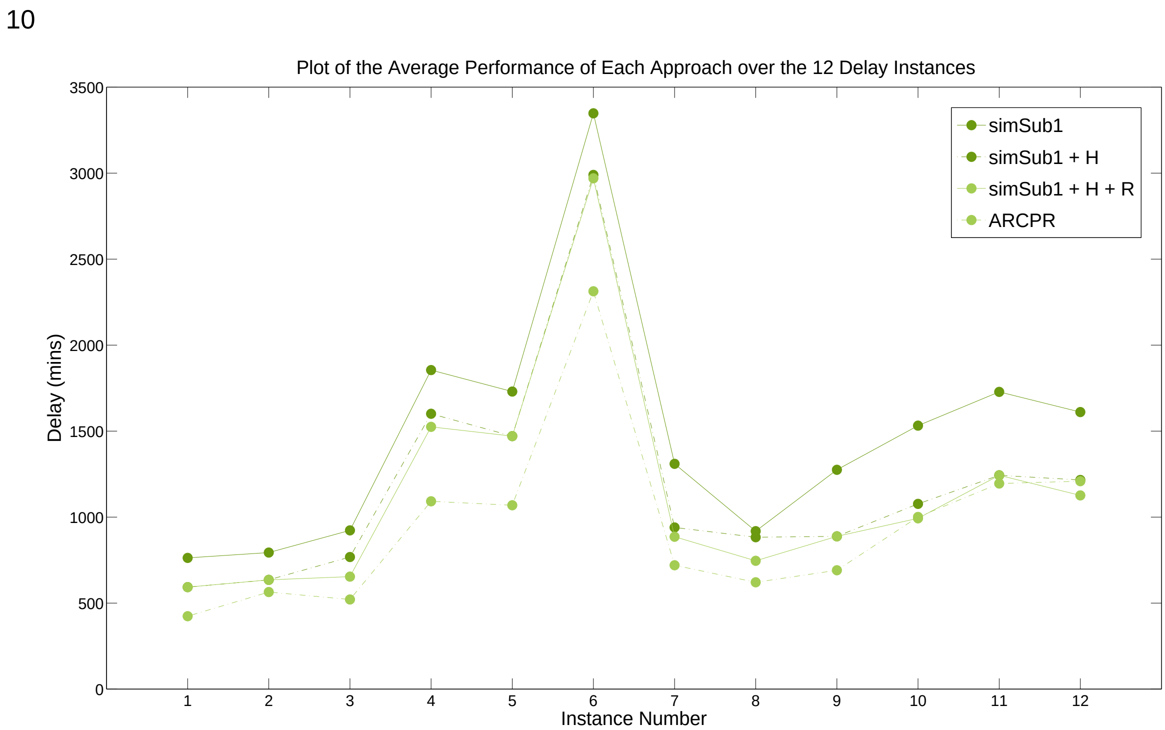


Figure 7.5: Plot of the average performance of each approach over the 12 delay instances.

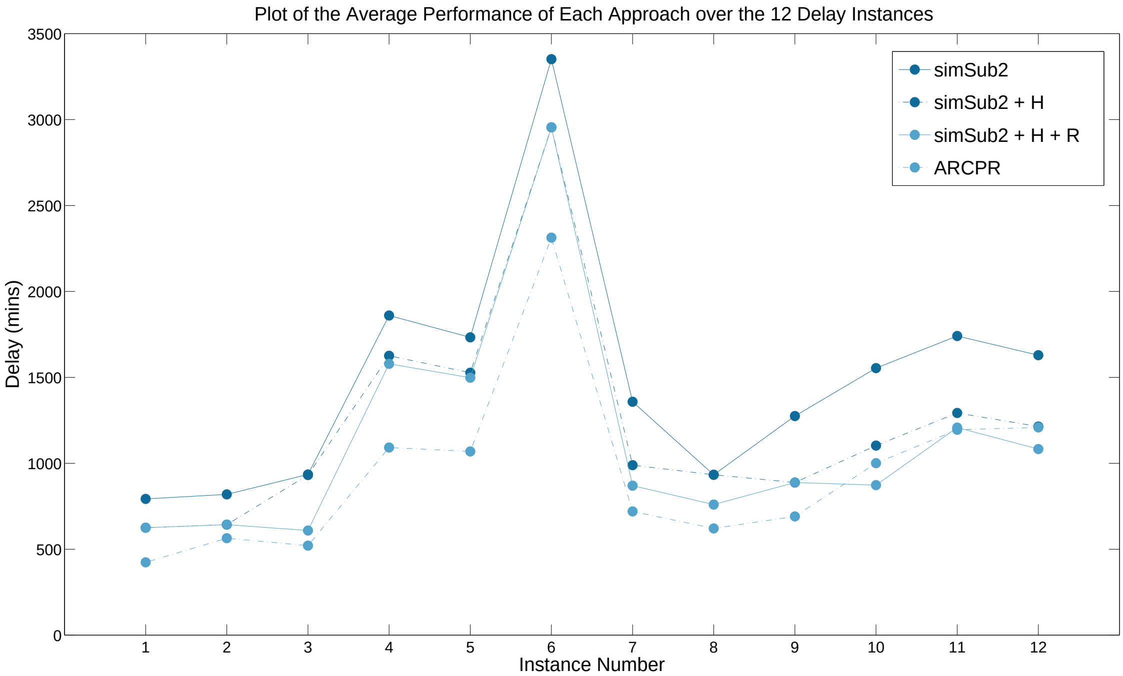


Figure 7.6: Plot of the average performance of each approach over the 12 delay instances.

7.6 Discussion

In Table 7.1 it may be observed that the proposed integrated aircraft routing, crew pairing and re-timing model achieves significant improvements over all other approaches discussed thus far. The ARCPR approach yields an impressive improvement of 34.83% over the B+H approach, with an improvement of approximately $20-23\%$ for the IPD+H, simSub1+H and simSub2+H approaches. These results, across all test instances are captured more clearly in the figures above, with the performance of the ARCPR approach also improving upon the results for the IPD+H+R, simSub1+H+R approaches and almost all instances in the simSub2+H+R approach. One of the drawbacks however, is the long computation time required for the ARCPR approach when compared with the iterative aircraft routing and crew pairing model and the re-timing heuristic. However, one advantage of all the approaches discussed thus far, is that the user may examine the trade-off in solution quality vs time, and use the approach best suited to their needs.

CHAPTER

EIGHT

Summary Of Contributions

In this thesis, we have proposed four different approaches for minimising propagated delay in an integrated aircraft routing and crew pairing framework. We tested our approaches on data from a real airline network and discovered that all of our approaches had the capacity to significantly minimise delay propagation in the network.

In Chapter 4 we proposed our first approach for minimising propagated delay in an integrated aircraft routing and crew pairing framework. We discussed the mathematical formulation behind our model and described how this work improved upon existing models in the field. We then obtained results for our model using data from a real airline network and demonstrated that our model outperformed existing models in a number of areas. Specifically, this model (i) accurately calculates *the combined effects* of delay propagation between aircraft and crew and (ii) uses this information for both the calculation of the *cost* of columns and the dynamic *selection* of optimal columns. By solving the problem in an iterative manner we avoided introducing extra complexity within the problem and achieved fast solution times over all test instances.

In Chapter 5 we proposed a re-timing heuristic that may be used in conjunction with the model proposed in Chapter 4; allowing for the possibility of obtaining a potentially more operationally robust solution. The proposed heuristic *simultaneously* re-times aircraft and crew whilst attempting to minimise overall delay propagation in the network. This heuristic *preserves the aircraft and crew assignments* whilst *simultaneously* re-timing aircraft and crew; allowing for accurate assessment of delay propagation between aircraft and crew and an effective re-timing to be chosen. We demonstrated that, despite its simplicity, the heuristic performed very well on a number of test instances.

In Chapter 6, we proposed two new methods of embedding delay scenarios within the aircraft routing and crew pairing subproblems of the model proposed in Chapter 4. Additionally, we extended the re-timing heuristic of Chapter 5 to include delay scenarios. We demonstrated that embedding each of these methods within the subproblem is beneficial and when further combined with the improved re-timing heuristic, the solution obtained out-performs the solution obtained using the heuristic (based on mean values) of Chapter 5. Additionally we investigated whether it was possible to improve upon this re-timed solution by re-solving each approach using the new departure times obtained from the heuristic. The results obtained indicated that for certain delay instances this approach was highly effective and able to achieve significant improvements with relatively short computation times.

Finally, In Chapter 7 we proposed a fully integrated model that captured the aspects of aircraft routing, crew pairing and re-timing without introducing any extra complexity within the master problem. All three decisions were made simultaneously by allowing re-timing decisions to be made within the aircraft routing subproblem. The proposed method improved the solution further and we demonstrated that this approach had the capacity to obtain significant delay improvements; out-performing all others used in the thesis.

Possible areas for research may include an extension of the model in Chapter 4 to a third aspect, such as passengers. Such an extension may investigate the effects of delay propagation on passenger spill and may also allow for the integration of fleet assignment, to assist in minimising passenger spill. The model may also be extended to a planning period of several days (or a week) and incorporate more complex crew rules for overnights etc. Finally, further computational investigation may allow for a reduction in the time needed to solve the ARCPR approach proposed in Chapter 7.

Appendix

In this Appendix we include computational results for an additional set of experiments. In the first section we list the results for the algorithms outlined in Chapters 4, 5, 6 and 7; in which we for each instance, record the results for 20 different "batches" of 1000 delay scenarios. In the second section we list the computational results for one delay instance (one set of 1000 delay scenarios) using a larger flight network. A more detailed discussion and analysis of each set of tests may be found in Chapter 6 and 7.

8.1 Additional Delay Scenarios

8.1.1 Exponential Distribution With $\lambda = 5$.

Instance 1

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	688	504	1192	–	9.17
Base + H.	536	472	1008	15.44	9.17 + 21.16
IPD + H.	428	377	805	32.47	47.19 + 24.22
simSub1 + H.	415	343	758	36.41	61.00 + 24.61
simSub2 + H.	428	359	787	33.98	58.98 + 24.33
IPD + H + R.	368	240	648	45.64	47.19 + 24.22 + 44.22
simSub1 + H + R.	352	241	593	50.25	61.00 + 24.61 + 60.21
simSub2 + H + R.	363	257	621	47.90	58.98 + 24.61 + 56.10
ARCPR (3 iter.)	310	226	536	55.03	9464

Instance 2

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	501	612	1113	–	9.10
Base + H.	422	436	858	22.91	9.10 + 16.19
IPD + H.	402	471	873	21.56	71.02 + 22.33
simSub1 + H.	404	395	799	28.21	83.12 + 27.42
simSub2 + H.	402	423	825	25.88	76.04 + 26.40
IPD + H + R.	334	409	743	33.24	71.02 + 22.33 + 70.00
simSub1 + H + R.	310	322	632	43.22	83.12 + 27.42 + 83.10
simSub2 + H + R.	305	334	639	42.58	76.04 + 26.40 + 77.05
ARCPR (3 iter.)	288	312	600	46.09	9828

Instance 3

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	639	751	1390	–	10.15
Base + H.	512	620	1132	18.56	10.15 + 20.45
IPD + H.	313	333	646	53.53	63.45 + 20.55
simSub1 + H.	314	297	611	56.04	78.31 + 24.12
simSub2 + H.	305	305	610	56.11	72.44 + 23.00
IPD + H + R.	255	236	491	64.68	63.45 + 20.55 + 67.43
simSub1 + H + R.	251	205	456	67.19	78.31 + 24.12 + 80.31
simSub2 + H + R.	326	203	529	61.94	72.44 + 23.00 + 70.10
ARCPR (2 iter.)	310	200	510	63.31	8640

8.1.2 Exponential Distribution With $\lambda = 10$

Instance 4

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	1381	1112	2493	–	8.12
Base + H.	993	902	1895	23.98	8.12 + 20.48
IPD + H.	991	912	1909	23.42	72.26 + 20.58
simSub1 + H.	986	889	1874	24.83	91.27 + 23.03
simSub2 + H.	972	908	1880	24.58	93.46 + 23.00
IPD + H + R.	878	770	1648	33.89	72.26 + 20.58 + 73.09
simSub1 + H + R.	884	747	1631	34.58	91.27 + 20.58 + 92.36
simSub2 + H + R.	874	768	1642	34.14	93.46 + 23.00 + 97.11
ARCPR (4 iter.)	836	646	1482	40.55	9475

Instance 5

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	1006	1105	2111	–	15.53
Base + H.	980	1036	2016	4.50	15.53 + 15.03
IPD + H.	824	948	1772	15.06	214.19 + 16.24
simSub1 + H.	835	881	1716	18.71	228.65 + 18.41
simSub2 + H.	835	886	1721	18.47	220.29 + 17.54
IPD + H + R.	745	811	1556	26.29	214.19 + 16 + 182.17
simSub1 + H + R.	739	720	1459	30.89	228.65 + 18.41 + 210.00
simSub2 + H + R.	759	755	1514	28.28	220.29 + 17.54 + 188.63
ARCPR (2 iter.)	620	706	1326	37.18	8118

Instance 6

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	2002	2110	4112	–	15.21
Base + H.	2006	2012	4018	2.29	15.21 + 19.04
IPD + H.	1694	1666	3360	18.28	92.35 + 21.31
simSub1 + H.	1681	1661	3341	18.75	112.91 + 22.05
simSub2 + H.	1653	1685	3338	18.82	99.21 + 20.15
IPD + H + R.	1690	1595	3285	20.11	92.35 + 21.31 + 90.89
simSub1 + H + R.	1568	1413	2981	27.50	112.91 + 22.05 + 109.61
simSub2 + H + R.	1527	1421	2948	28.31	99.21 + 20.15 + 95.88
ARCPR (3 iter.)	1180	1322	2502	39.15	10011

8.1.3 Truncated Normal Distribution With $\mu = 5$, $\sigma = 10$

Instance 7

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	946	952	1898	–	14.19
Base + H.	612	780	1392	26.65	14.19 + 18.22
IPD + H.	614	746	1360	28.35	168.74 + 19.41
simSub1 + H.	612	701	1313	30.82	217.43 + 27.61
simSub2 + H.	614	746	1360	28.35	201.22 + 25.87
IPD + H + R.	510	451	961	49.37	168.74 + 19.41 + 168.00
simSub1 + H + R.	515	426	941	50.42	217.43 + 27.61 + 200.17
simSub2 + H + R.	525	465	990	47.84	201.22 + 25.87 + 197.33
ARCPR (3 iter.)	388	412	800	57.85	12215

Instance 8

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	548	673	1221	–	13.31
Base + H.	545	564	1109	9.17	13.31 + 16.17
IPD + H.	528	422	950	22.19	39.44 + 19.82
simSub1 + H.	533	384	917	24.89	61.10 + 25.00
simSub2 + H.	533	400	933	23.58	51.14 + 24.33
IPD + H + R.	420	254	674	44.80	39.44 + 19.82 + 39.00
simSub1 + H + R.	549	335	884	27.60	61.10 + 25.00 + 60.18
simSub2 + H + R.	571	278	849	30.47	51.14 + 24.33 + 49.27
ARCPR (2 iter.)	480	235	715	41.44	9967

Instance 9

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	837	842	1679	–	15.11
Base + H.	675	637	1312	21.85	15.11 + 16.29
IPD + H.	667	643	1310	21.97	62.32 + 17.21
simSub1 + H.	657	620	1277	23.94	71.12 + 17.98
simSub2 + H.	657	620	1277	23.94	70.77 + 18.65
IPD + H + R.	550	393	943	43.84	62.32 + 17.21 + 58.81
simSub1 + H + R.	523	364	888	47.11	71.12 + 17.98 + 69.14
simSub2 + H + R.	523	364	888	47.11	70.77 + 18.65 + 67.00
ARCPR (3 iter.)	418	295	713	57.53	10708

8.1.4 Truncated Normal Distribution With $\mu = 10$, $\sigma = 5$

Instance 10

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	781	990	1948	–	14.91
Base + H.	796	811	1607	17.50	14.91 + 16.45
IPD + H.	667	905	1572	19.30	53.35 + 18.26
simSub1 + H.	682	851	1533	21.30	67.45 + 21.44
simSub2 + H.	688	867	1555	20.17	59.16 + 19.08
IPD + H + R.	468	605	1077	44.71	53.35 + 18.26 + 55.28
simSub1 + H + R.	516	561	1073	44.91	67.45 + 21.44 + 62.91
simSub2 + H + R.	525	578	1103	43.38	59.16 + 19.08 + 57.12
ARCPR (3 iter.)	465	530	995	48.92	7114

Instance 11

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	967	1131	2098	–	14.70
Base + H.	899	1017	1916	8.67	14.70 + 16.71
IPD + H.	735	1044	1779	15.20	61.72 + 18.22
simSub1 + H.	714	1017	1730	17.54	74.38 + 20.54
simSub2 + H.	711	1032	1744	16.87	69.27 + 21.55
IPD + H + R.	572	716	1288	38.61	61.72 + 18.22 + 66.19
simSub1 + H + R.	533	711	1244	40.71	74.38 + 20.54 + 69.00
simSub2 + H + R.	560	735	1294	38.32	69.27 + 21.55 + 62.14
ARCPR (2 iter.)	527	653	1180	43.76	4752

Instance 12

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	1020	1071	2091	–	15.04
Base + H.	901	973	1874	10.38	15.04 + 17.37
IPD + H.	769	914	1683	19.51	72.97 + 16.54
simSub1 + H.	769	842	1611	22.95	82.66 + 17.30
simSub2 + H.	778	851	1629	22.09	74.22 + 16.54
IPD + H + R.	622	616	1238	40.79	72.97 + 16.54 + 76.80
simSub1 + H + R.	638	578	1216	41.84	82.66 + 17.30 + 85.77
simSub2 + H + R.	646	569	1215	41.89	74.22 + 16.54 + 70.32
ARCPR (4 iter.)	512	614	1126	46.15	10440

8.2 Results for a Larger Network (One Instance)

In this section we include results for a larger flight network. Our results are obtained for a network consisting of 320 flights with 818 feasible connections, assuming a one-day schedule. We determine that the minimum number of aircraft and crew pairs required to cover this network are 64 and 102 crew, respectively.

8.2.1 Exponential Distribution With $\lambda = 10$

Instance 4

Approach	Aircraft Delay	Crew Delay	Total Delay	% Improvement over B	Time (s)
Base (B)	2698	4166	6864	–	78.08
Base + H.	2219	3012	5231	23.79	78.08 + 146.00
IPD + H.	2004	3006	5010	27.01	498.32 + 151.06
simSub1 + H.	2132	2756	4888	28.78	746.11 + 187.45
simSub2 + H.	2007	2947	4954	27.83	640.04 + 179.10
IPD + H + R.	1856	2010	3866	43.68	498.32 + 151.06 + 500.00
simSub1 + H + R.	1726	2074	3800	44.64	746.11 + 187.45 + 815.32
simSub2 + H + R.	1986	2000	3986	41.93	640.04 + 179.10 + 700.50
ARCPR (3 iter.)	1572	1950	3522	48.69	45360

Bibliography

[1] J. Abara. Applying Integer Linear Programming to the Fleet Assignment Problem. *Interfaces*, 19(4):20–28, 1989.

[2] K.F. Abdelghany, S.S. Shah, S. Raina, and A.F. Abdelghany. A Model for Projecting Flight Delays During Irregular Operation Conditions. *Air Transport Management*, 10:385–394, 2004.

[3] Y. Ageeva. Approaches to Incorporating Robustness into Airline Scheduling. Master's thesis, 2000.

[4] S. AhmadBeygi, A. Cohn, and M. Lapp. Decreasing Airline Delay Propagation by Re-Allocating Scheduled Slack. In *Proceedings of the Industry Studies Conference*, 2008.

[5] S. AhmadBeygi, A. Cohn, and M. Weir. An Integer Programming Approach to Generating Airline Crew Pairings. *Computers and Operations Research*, 26:1284–1298, 2009.

[6] A. H. Ahmed and C. A. Poojari. An Overview Of the Issues in the Airline Industry and the Role of Optimization Models and Algorithms. *Journal of Operational Research Society*, 59:267–277, 2008.

[7] M. Ball, C. Barnhart, G. Nemhauser, and A. Odoni. *Air Transportation : Irregular Operations and Control*, volume 14, chapter 1, pages 1–67. Elsevier, 2007.

[8] C. Barnhart, P. Belobaba, and A. R. Odoni. Applications of Operations Research in the Air Transport Industry. *Transportation Science*, 37(4):368–391, November 2003.

[9] C. Barnhart, A. M. Cohn, D. Klabjan, G. L. Nemhauser, and P. Vance. Airline crew scheduling. In R. W. Hall, editor, *Handbook of Transportation Science*, pages 517–560. Kluwer, 2003.

[10] C. Barnhart, A. Farahat, and M. Lohatepanont. Airline Fleet Assignment: An Alternative Model and Solution Approach Based on Composite Variables, 2001.

[11] C. Barnhart, A. Farahat, and M. Lohatepanont. Airline Fleet Assignment with Enhanced Revenue Modelling. *Operations Research*, 57(1):231–244, 2009.

[12] C. Barnhart, T. S. Kniker, and M. Lohatepanont. Itinerary-Based Airline Fleet Assignment. *Transportation Science*, 36(2):199–217, May 2002.

[13] C. Barnhart, F. Lu, and R. Shenoi. Integrated Airline Scheduling. In *Operations Research in the Air Industry, International Series in Operations Research and Management Science*, pages 384–403. Kluwer Academic Publishers, 1997.

[14] Cynthia Barnhart, Natashia L. Boland, Lloyd W. Clarke, Ellis L. Johnson, George L. Nemhauser, and Rajesh G. Shenoi. Flight String Models for Aircraft Fleeting and Routing. *Transportation Science*, 32(3):208–220, 1998.

[15] Cynthia Barnhart, Ellis L. Johnson, George L. Nemhauser, Martin W. P. Savelsbergh, and Pamela H. Vance. Branch-and-Price: Column Generation for Solving Huge Integer Programs. *Operations Research*, 46(3):316–329, May–June 1998.

[16] N. Bélanger, G. Desaulniers, F. Soumis, J. Desroisers, and J. Lavigne. Weekly Airline Fleet Assignment with Homogeneity. *Transportation Research Part B: Methodological*, 40(4):306–318, 2006.

[17] N. Bélanger, G. Desaulniers, F. Soumis, and J. Desrosiers. Periodic Fleet Assignment with Time Windows, Spacing Constraints and Time Dependent Revenues. *European Journal of Operational Research*, 175:1754–1766, 2006.

[18] A. Ben-Tal, S. Boyd, and A. Nemirovski. Extending the Scope of Robust Optimization: Comprehensive Robust Counterparts of Uncertain Problems. *Mathematical Programming, Series B*, 107:63–89, 2006.

[19] M. Berge and C. A. Hopperstad. Demand Driven Dispatch: A Method for Dynamic Aircraft Capacity Assignment, Models and Algorithms. *Operations Research*, 41(1):153–168, 1993.

[20] D. Bertsimas, D. B. Brown, and C. Caramanis. Theory and Applications of Robust Optimization. *SIAM Revue*, 53(3):464–501, 2011.

[21] D. Bertsimas and M. Sim. Robust Discrete Optimization and Network Flows. *Mathematical Programming Series B*, 98(1):49–71, 2003.

[22] H. G. Beyer and B. Sendhoff. Robust Optimization - A Comprehensive Survey. *Computational Methods in Applied Mechanics and Engineering*, 196:3190–3218, 2007.

[23] Ş. İlker Birbil, J. B. G. Frenk, Joaquim A. S. Gromicho, and Shuzhong Zhang. The Role of Robust Optimization in Single-Leg Airline Revenue Management. *Management Science*, 55:148–163, January 2009.

[24] A. Charnes and W. Cooper. Chance Constrained Programming. *Management Science*, 6(1):73–79, 1959.

[25] L. Clarke, C. Hane, E. Johnson, and G. Nemhauser. Maintenance and Crew Considerations in Fleet Assignment. *Transportation Science*, 30:249–260, 1994.

[26] L. Clarke, E. Johnson, G. Nemhauser, and Z. Zhu. The Aircraft Rotation Problem. *Annals of Operations Research*, 69:33–46, 1997.

[27] A. Cohn and C. Barnhart. Improving Crew Scheduling by Incorporating Key Maintenance Routing Decisions. *Operations Research*, 51(3):387–396, 2003.

[28] J. F. Cordeau, G. Laporte, M. Savelsbergh, and D. Vigo. *Short-Haul Routing.* Centre for Research on Transportation, 2005.

[29] J. F. Cordeau, G. Stojkovic, F. Soumis, and J. Desrosiers. Benders Decomposition for Simultaneous Aircraft Routing and Crew Scheduling. *Transportation Science*, 35(4):375–388, 2001.

[30] G. Desaulniers, J. Desrosiers, Y. Dumas, M.M. Soloman, and F. Soumis. Daily Aircraft Routing and Scheduling. *Management Science*, 43(6):841–855, June 1997.

[31] M. Desrochers and F. Soumis. A Generalized Permanent Labelling Algorithm for the Shortest Path Problem with Time Windows. *Informs*, 26(3):191–212, 1988.

[32] J. Dumas, F. Aithnard, and F. Soumis. Improving the Objective Function of the Fleet Assignment Problem. *Transportation Research Part B*, 43(4), 2008.

[33] J. Dumas and F. Soumis. Passenger Flow Model for Airline Networks. *Transportation Science*, 42(2):197–207, May 2008.

[34] I. Dumitrescu and N. Boland. Improved preprocessing, labeling and scaling algorithms for the weight-constrained shortest path problem. *Networks*, 42(3):135–153, 2003.

[35] M. Dunbar, G. Froyland, and C. L. Wu. Robust Airline Schedule Planning: Minimizing Propagated Delay in an Integrated Routing and Crewing Framework. *Transportation Science*, 46(2):204–216, 2012.

[36] M. Ehrgott and D. M. Ryan. Constructing Robust Crew Schedules with Bi-Criteria Optimization. *Journal of Multi-Criteria Decision Analysis*, 11:1–12, 2002.

[37] M. Ehrgott and J. Tind. Column Generation in Integer Programming with Applications in Multi-criteria Optimization. Technical report, University of Aukland, 2007. University of Auckland School of Engineering 651.

[38] M. Etschmaier and D. Mathaisel. Airline Scheduling: An Overview. *Transportation Science*, 19(2):127–138, 1985.

[39] T. A. Feo and J. F. Bard. Flight Scheduling and Maintenance Base Planning. *Management Science*, 35(12):1415–1432, 1989.

[40] R. Freling, D. Huisman, and A. P. M. Wagelmans. Models and Algorithms for Integration of Vehicle and Crew Scheduling. *Journal of Scheduling*, 6(1):63–85, 2003.

[41] B. Golden, S. Raghavan, and E. Wasil, editors. *The Vehicle Routing Problem: Latest Advances and New Challenges*, volume 43 of *Operations Research/Computer Science Interfaces Series*. Springer, 2008.

[42] R. Gopalan and K. T. Talluri. The Aircraft Maintenance Routing Problem. *Operations Research*, 46(2):260–271, 1998.

[43] R. Gopalan and K.T Talluri. Mathematical Models In Airline Scheduling: A Survey. *Annals of Operation Research*, 76:155–185, 1998.

[44] M. Grönkvist. Accelerating Column Generation for Aircraft Scheduling Using Constraint Propagation. *Computers and Operations Research*, 33(10):2918–2934, 2006.

[45] Y. Guo, T. Mellouli, L. Suhl, and M. P. Thiel. A Partially Integrated Airline Crew Scheduling Approach with Time-Dependent Crew Capacities and Multiple Home Bases. *European Journal of Operational Research*, 17:1169–1181, 2006.

[46] C. A. Hane, C. Barnhart, E. L. Johnson, R. E. Marsten, G. L. Nemhauser, and G. Sigismondi. The Fleet Assignment Problem: Solving a Large-Scale Integer Program. *Mathematical Programming*, 70(2):211–232, 1995.

[47] M. Haouari, N. Aissaoui, and F.Z. Mansour. Network Flow-Based Approaches for Integrated Aircraft Fleeting and Routing. *European Journal of Operations Research*, 193(2):591–599, 2007.

[48] IATA. Airport Handling Manual. 2010.

[49] International Air Transport Association (IATA). Fact sheet: Economic and social benefits of air transport. http://www.iata.org/pressroom/facts_figures. Date Accessed: 3rd January, 2012.

[50] T. L. Jacobs, B. C. Smith, and E. L. Johnson. Incorporating Network Flow Effects into the Airline Fleet Assignment Process. *Transportation Science*, 45(4):514–529, November 2008.

[51] S. L. Janak, X. Lin, and X. A. Floudas. A New Robust Optimization Approach for Scheduling Under Uncertainty ii. Uncertainty with Known Probability Distribution. *Computers and Chemical Engineering*, 31:171–195, 2007.

[52] L. Kang. *Degradable Airline Scheduling : An Approach to Improve Operational Robustness and Differentiate Service Quality.* PhD thesis, 2004. http://hdl.handle.net/1721.1/17659.

[53] M. Kaut and S. Wallace. Evaluation of Scenario Generation Methods for Stochastic Programming. *Pacific Journal of Optimization*, 3(2):257–271, 2007.

[54] D. Klabjan. Large-Scale Models in the Airline Industry. In M. M. Solomon, G. Desaulniers, and J. Desrosiers, editors, *Column Generation*, pages 163–195. Springer, 2005.

[55] D. Klabjan, E. L. Johnson, and G. L. Nemhauser. Solving Large Airline Crew Scheduling Problems: Random Pairing Generation and Strong Branching. *Computational Optimization and Applications*, 20:73–91, 2001.

[56] D. Klabjan, E. L. Johnson, G. L. Nemhauser, E. Gelman, and S. Ramaswamy. Airline Crew Scheduling with Time Windows and Plane-Count Constraints. *Transportation Science*, 36(3):337–348, 2002.

[57] S. Lan, J. Clarke, and C. Barnhart. Planning for Robust Airline Operations Optimizing Aircraft Routings and Flight Departure Times to Minimize Passenger Disruptions. *Transportation Science*, 40(1):15–28, 2006.

[58] M. Lapp and A. Cohn. Modifying Lines-Of-Flight in the Planning Process for Improved Maintenance Robustness. *Computers and Operations Research*, 39(9):2051–2062, 2012.

[59] Z. Li and M. Ierapetritou. Process Scheduling Under Uncertainty: Review and Challenges. *Computers and Chemical Engineering*, 32:715–727, 2008.

[60] D. Liang and W. E. Wilhelm. A Generalization of Column Generation to Accelerate Convergence. *Mathematical Programming*, 122(2):349–378, 2008.

[61] G.F. List, B. Wood, L. K. Nozick, M. A. Turnquist, D. A. Jones, E. A. Kjeldgaard, and C. R. Lawton. Robust Optimization for Fleet Planning Under Uncertainty. *Transportation Research Part E: Logistics and Transportation Review*, 39(3):209 – 227, 2003.

[62] O. Listes and R. Dekker. A Scenario Aggregation-Based Approach for Determining a Robust Airline Fleet Composition for Dynamic Capacity Allocation. *Transportation Science*, 39(3):367–382, 2005.

[63] M. Lohatepanont and C. Barnhart. Airline Schedule Planning: Integrated Models and Algorithms for Schedule Design and Fleet Assignment. *Transportation Science*, 38(1):19–32, February 2004.

[64] F. Lu. *Integrated Airline Scheduling: Models and Algorithms*. PhD thesis, Department of Civil and Environmental Engineering, 1997.

[65] M. E. Lübbecke and J. Desrosiers. Selected Topics in Column Generation. *Operations Research*, 53(6):1007–1023, 2005.

[66] A. Makri and D. Klabjan. A New Pricing Scheme for Airline Crew Scheduling. *INFORMS Journal on Computing*, 16(1):56–67, Winter 2004.

[67] L. Marla and C. Barnhart. Robust Optimization: Lessons Learned from Aircraft Routing. *Transportation Science*, IN PRESS, 2011.

[68] A. Mercier, J.F. Cordeau, and F. Soumis. A Computational Study of Benders Decomposition for the Integrated Aircraft Routing and Crew Scheduling Problem. *Computational Operations Research*, 32(6):1451–1476, 2005.

[69] A. Mercier and F. Soumis. An Integrated Aircraft Routing, Crew Scheduling and Flight Retiming Model. *Computers and Operations Research*, 34:2251–2265, 2007.

[70] A. Mingozzi and A. Valleta. The Multi-Depot and Periodic Vehicle Routing Problem. volume 3607 of *Lecture Notes in Computer Science*. Springer, 2005. Presented at the 2005 Symposium on Abstraction, Reformulation and Approximation, Edinburgh.

[71] L. Moccia, J. F. Cordeau, G. Laporte, S. Røpke, and M. P. Valentini. Modelling and Solving a Multi-Modal Routing Problem with Timetables and Time Windows. January 2009.

[72] M. Mourgaya and F. Vanderbeck. Column Generation Based Heuristic for Tactical Planning in Multi Period Vehicle Routing. *European Journal of Operational Research*, 183(3):1028–1041, 2007.

[73] S. Mudchanatongsuk, F. Ordóñez, and J. Liu. Robust Solutions for Network Design under Transportation Cost and Demand Uncertainty. *Journal of the Operational Research Society*, 59:652–662, 2008.

[74] N. Papadakos. Integrated Airline Scheduling. *Computers and Operations Research*, 36:176–195, 2009.

[75] Yi Qin, Jin fu Zhu, and Wei Zhang. The robust optimization of the capacitated hub and spoke airline network design base on ant colony algorithms. In *Intelligent Computing and Integrated Systems (ICISS), 2010 International Conference on*, pages 601–604, October 2010.

[76] Research and Innovative Technology Administration. Bureau of Transportation Statistics (BTS). http://www.transtats.bts.gov/. Date Accessed: August 10, 2010.

[77] B. Rexing, C. Barnhart, T. Kniker, A. Jarrah, and N. Krishnamurthy. Airline Fleet Assignment with Time Windows. *Transportation Science*, 34(1):1–20, February 2000.

[78] J. Rosenberger, E. Johnson, and G. Nemhauser. A Robust Fleet Assignment Model with Hub Isolation and Short Cycles. *Transportation Science*, 38(3):357–368, August 2004.

[79] J.M. Rosenberger, A.J. Schaefer, D. Goldsman, E.L. Johnson, A.J. Kleywegt, and G.L. Nemhauser. A Stochastic Model of Airline Operations. *Transportation Science*, 36(4):357–377, November 2002.

[80] K. Roy and C.J. Tomlin. Solving the Aircraft Routing Problem Using Network Flow Algorithms. In *American Control Conference*, July 2007.

[81] R. A. Rushmeier, K. L. Hoffman, and M. Padberg. Recent Advances in Exact Optimization of Airline Scheduling Problems, 1995. http://iris.gmu.edu/∼khoffman/papers/fleetsched.ps.

[82] S. Ruther, N. Boland, F. Engineer, and I. Evans. A Column Generation Formulation for the Integrated Aircraft Routing, Crew Pairing and Tail Number Assignment Problem. Presentation at IFORS Conference, Melbourne 2011.

[83] D. M. Ryan. The Solution of Massive Generalized Set Partitioning Problems in Aircrew Rostering. *Journal of the Operational Research Society*, (43):459–467, 1992.

[84] R. Sandhu and D. Klabjan. Fleeting with Passenger and Cargo Origin-Destination Booking Control. *Transportation Science*, 40(4):517–528, November 2006.

[85] R. Sandhu and D. Klabjan. Integrated Airline Fleeting and Crew-Pairing Decisions. *Operations Research*, 55(3):439–456, May 2007.

[86] S. Sarmadi. Minimizing Airline Passenger Delay through Integrated Flight Scheduling and Aircraft Routing. Master's thesis, 2004. http://hdl.handle.net/1721.1/29401.

[87] A. Schaefer, E. L. Johnson, A. J. Kleywegt, and G. L. Nemhauser. Airline Crew Scheduling Under Uncertainty. *Transportation Science*, 39(3):340–348, August 2005.

[88] S. Shebalov and D. Klabjan. Robust Airline Crew Pairing: Move Up Crews. *Transportation Science*, 40(3):300–312, August 2006.

[89] H. D. Sherali, E. K. Bish, and X. Zhu. Invited Tutorial: Airline Fleet Assignment Concepts, Models, and Algorithms. *European Journal of Operational Research*, 172:1–30, 2006.

[90] H. D. Sherali and X. Zhu. Two-Stage Fleet Assignment Model Considering Stochastic Passenger Demands. *Operations Research*, 56:383–399, March 2008.

[91] H. D. Sherali and X. Zhu. Two-Stage Fleet Assignment Model Considering Stochastic Passenger Demands. *Operations Research*, 56(2):383–399, March–April 2008.

[92] R. W. Simpson and P. P. Belobaba. The Supply of Scheduled Air Transportation Services, 1992b. Lecture notes for 16.74 Air Transportation Economics, MIT, Cambridge, MA.

[93] B. Smith and E. L. Johnson. Robust Airline Fleet Assignment: Imposing Station Purity Using Station Decomposition. *Transportation Science*, 40(4):497–516, November 2006.

[94] I. Sungur, F. Ordóñez, and M. Dessouky. A Robust Optimization Approach for the Capacitated Vehicle Routing Problem with Demand Uncertainty. *IEEE Transactions*, 40:509–523, 2008.

[95] B. Tam, M. Ehrgott, D. Ryan, and G. Zakeri. A Comparison of Stochastic Programming and Bi-objective Optimization Approaches to Robust airline Crew Scheduling. volume 33, pages 49–75. 2011.

[96] Y. Tu, M. O. Ball, and W. S. Jank. Estimating Flight Departure Delay Distributions: A Statistical Approach With Long-Term Trend and Short-Term Pattern. *Journal of the American Statistical Association*, 103:112–125, 2008.

[97] P. Vance, C. Barnhart, E. L. Johnson, and G. L. Nemhauser. Airline Crew Scheduling: A New Formulation and Decomposition Algorithm. *Operations Research*, 45(2):188–200, March–April 1997.

[98] O. Weide, D. Ryan, and M. Ehrgott. Iterative Airline Scheduling. School of Engineering, University of Auckland, New Zealand, 2007. http://hdl.handle.net/2292/2669.

[99] Oliver Weide. *Robust and Integrated Airline Scheduling*. PhD thesis, 2009. http://hdl.handle.net/2292/4237.

[100] C. L. Wu. Inherent Delays and Operational Reliability of Airline Schedules. *Journal of Air Transport Management*, 11(4):273–282, 2005.

[101] C. L. Wu. Delay Propagation Modelling and the Implications In Robust Airline Scheduling. *Transportation Research Board 87th Annual Meeting*, 2008.

[102] C. L. Wu. *Airline Operations and Delay Management - Insights from Airline Economics, Networks and Strategic Schedule Planning*. Ashgate, 2010.

[103] C.L. Wu. Improving Airline Robustness and Operational Reliability by Sequential Optimisation Algorithms. *Network and Spatial Economics*, 6(3):235–251, 2006.

[104] J.W Yen and J.R Birge. A Stochastic Programming Approach to the Airline Crew Scheduling Problem. *Transportation Science*, 40(1):3–14, 2006.

[105] C. S. Yu and H. L. Li. A Robust Optimization Model For Stochastic Logistic Problems. *International Journal of Production Economics*, 64:385–397, 2000.