CHALMERS

A Simulation and Optimization Framework for Evaluating Airline Schedule Robustness

Master of Science Thesis

CARL ENGSTRÖM

Department of Applied Information Technology
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2012
Report No. 2012:015
ISSN: 1651-4769

A Simulation and Optimization Framework for Evaluating Airline Schedule Robustness
CARL ENGSTRÖM

Technical report no 2012:015
Department of Applied Information Technology
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone + 46 (0)31-772 1000

Göteborg, Sweden 2012

A Simulation and Optimization Framework for Evaluating Airline Schedule Robustness
CARL ENGSTRÖM
Department of Applied Information Technology
Chalmers University of Technology

Abstract

The airline industry is constantly attempting to find new ways to combat delays and cancellations. Much of the work is done in two fields, the areas of robust airline scheduling and the area of airline recovery. Robust scheduling strives to make schedules that can be executed even though the airline is faced with minor disruptions such as delays. Recovery is focused more towards returning to the original schedule when faced with severe disruptions.

A problem with these two areas is that they cannot fully be studied without first having to execute the schedule. This means that airlines must take risks when changing the way they schedule as they can only hope that the effects of their changes are good. For this purpose simulation of airline operations has been utilized in order to estimate these results beforehand.

Several tools for simulation are available for use today, but the majority of them only simulate isolated areas or use greatly simplified models. Nevertheless, these have proven to be of great use, not only to airlines but also to other areas of aviation such as airport and airspace congestion control.

This thesis shows how such existing simulation tools along with custom-written components can be combined with existing tools for recovery to create a simulation framework that can simulate the day-to-day operations of an airline in a realistic manner.

This framework is then used to evaluate performance of airline schedules both from a cost-efficiency and robustness point of view. The obtained results are then compared to historical results and show how interesting performance measures can be obtained from the simulated data.

A Simulation and Optimization Framework for Evaluating Airline Schedule Robustness
CARL ENGSTRÖM
Institutionen för tillämpad informationsteknologi
Chalmers tekniska högskola

Sammandrag

Flygindustrin lägger idag mycket kraft på att hitta nya sätt att minska antalet förseningsminuter och inställda flygningar. Två mycket intressanta områden har på senare tid fått mycket uppmärksamhet: robust schemaläggning, samt återställande av scheman som utsatts av yttre störningar. Robust schemaläggning syftar i huvudsak till att skapa scheman som inte påverkas nämnvärt av liten yttre påverkan. Konsten att återställa scheman fokuserar istället på hur man på bästa vis kan hantera effekterna av stora yttre störningar.

Ett stort problem för dessa områden är att dess effekter inte kan studeras fullt ut utan att schemat först genomförs. Detta innebär stora risker för flygbolagen då de ändrar sina strategier för schemaläggning och då endast kan hoppas att resultaten är goda. För detta syfte har datorsimulering använts till att estimera resultaten av sådana förändringar på förhand.

Idag finns flera simuleringsverktyg tillgängliga, men de allra flesta begränsar sig till enskilda delar av flygbolags verksamhet, eller förlitar sig på kraftigt förenklade modeller. Dessa verktyg har trots detta visat sig vara mycket användbara inte bara för flygbolag utan även för flygledning.

Denna rapport visar hur redan existerande simuleringsverktyg kan kombineras med specialskriven mjukvara och mjukvara för återställning av flygscheman för att skapa ett simuleringsramverk som på ett realistiskt sätt kan simulera den dagliga verksamheten hos flygbolag.

Detta ramverk används sedan för att utvärdera prestandan hos flygscheman både ur ett robusthets- och kostnadsperspektiv. Resultatet av simuleringen jämförs sedan mot historisk data och visar hur flera intresanta nyckelvärden kan estimeras genom simulering.

Rapporten är skriven på engelska.

Contents

1. Introduction 1
  1.1 Background 1
  1.2 Purpose 2
  1.3 Method 2
  1.4 Key findings 3
  1.5 Thesis outline 3
2. Theory and Literature 4
  2.1 The global airline industry in numbers 5
  2.2 Delays and cancellations 7
  2.3 Airline scheduling 9
  2.4 Airline recovery 11
  2.5 Robust scheduling 14
  2.6 Simulation 15
  2.7 Related work 16
3. Framework Design 18
  3.1 Simulation model 19
  3.2 Data set 20
  3.3 Simulation controller 21
  3.4 Disruption management 22
  3.5 Airport and airspace simulation 23
  3.6 Crew event generation 24
  3.7 Output 24
4. Framework Implementation 25
  4.1 Data set 26
  4.2 ISC 26
  4.3 MEANS 27
  4.4 Crew event generator 27
    4.4.1 Crew simulation model 29
    4.4.2 Customization 29
    4.4.3 Operation 31
    4.4.4 Mexicana event model 33
5. Results 35
  5.1 Outcome 36
  5.2 Comparison between simulated and real data 36
    5.2.1 Flight simulation 37
    5.2.2 Simulated events 38
    5.2.3 On- time performance 39
  5.3 Simulation of poor conditions 42
    5.3.1 Simulation results 43

5.4 Simulation stability ... 45
5.5 Simulation performance ... 45
6. Discussion ... 47
6.1 Results ... 48
6.2 Shortcomings ... 49
6.2.1 Passenger model ... 49
6.2.2 Airline operations ... 50
6.3 Performance and usage ... 50
6.4 Further applications ... 51
6.4.1 Planner training ... 51
6.4.2 Strategic decisions ... 52
6.4.3 Optimization ... 52
6.5 Summary ... 52
7. Future work ... 53
7.1 Passenger model ... 54
7.2 Airline modeling ... 54
7.3 Parameters for robust scheduling ... 54
7.4 Usability ... 55
8. Works Cited ... 56

1. Introduction

1.1 Background

The airline industry is, in spite of almost constantly growing passenger numbers, often struggling to turn a profit. The increase in revenue is often matched, or even surpassed, by increased running costs as well as costs related to delays and cancellations. Changes to running costs can be somewhat predicted but are often to a certain extent out of the airline's control. Costs due to delays and cancellation have almost opposite attributes; they are hard to predict but something that the airline can partially control.

Robustness in an airline schedule can be described as a schedule's inherent ability to be executed even though affected by external changes. In other words, a robust schedule can still be executed even if events, such as poor weather, occur. This is a much desired attribute of schedules but only serves a purpose if these kinds of events actually occur. Adding too much robustness might very well eliminate the costs related to delay and cancellations, but the cost of executing such a schedule would be extreme.

When studying the status quo of the airline industry one major problem appears: Airlines today can only assess how their current schedule performed given the conditions it was executed in. This performance is usually evaluated by comparing the sum of all costs during the scheduled period to delay minutes and cancelled flights.

In order to find possible improvements to the current schedule, a changed schedule must be executed in similar conditions or the same schedule executed in different conditions. Consider a pilot standing by to cover for other pilots in case of sickness. If no pilots call in sick then this is surely a useless expense, but if the schedule is changed so that there is no pilot standing by and a pilot *does* call in sick then there is no-one to fly the plane. The standby may or may not be a key component in making the schedule work. Airlines must therefore attempt to evaluate the how likely the standby is to be used and compare that to the cost of the standby and the possible effects of not having it. Even for such a trivial example the problem of evaluating what will happen if there is no pilot to fly the plane can be quite hard as the consequences might be far reaching. In reality whether a pilot can fly a plane depends on several things, such as previous flights and union rules.

Situations such as the one described above make it very hard to determine how the current schedule can be improved as even small changes to the schedule can have large effects. Robustness is very much desirable but it also comes at a cost. Having two pilots standing by obviously makes the schedule even more robust but may not be useful if it is not probable that two pilots call in sick on the same day. Airlines are faced with such choices both during planning and when executing schedules, known as *the day of operations*, where the schedule is usually adjusted as unexpected events occur. Due to the fact that the effects of schedule changes are hard to predict airlines are usually reluctant to changing the way they work and instead rely on taking decisions based on previous experience.

In short, airlines could benefit greatly from being able to estimate the results of “what-if” scenarios for certain situations. This would allow the testing of several parameters such as different scheduling methods, performance under stressful times, such as poor weather conditions, crew shortage, etc. Comparing the outcomes of different results for a given schedule provides understanding of the robustness of the schedule and a more informed decision can thereby be made.

1.2 Purpose

The purpose of this thesis is to construct a framework that can be used to simulate airline operations given circumstances that the user can control. By running the simulation the user can see how the current schedule performs under the given circumstances and produce estimates of key values regarding profitability and robustness of the schedule that are normally only known when the schedule has been executed. This can aid planners in choosing the schedule that is expected to perform best on the day of operations. Additionally, the simulation framework can also be used when airlines need to recover from disruptions by showing the expected effects of a recovery action.

The framework can also be utilized to compare various strategies for recovery to each other to determine the best course of action in a given set of circumstances. Such strategies include policies on when to cancel rather than delay, or when to make use of limited resources, such as making changes to a crew members schedule when only a certain amount of changes are allowed in a calendar month.

1.3 Method

Much work has already been done in the field of airline simulation, but mostly in isolated areas such as airport and airspace simulation. The framework will make use of previous work where possible, although it is not expected that all needed tools will exist and some will therefore need to be created. The area of crew simulation is known to be less investigated in previous work and is therefore expected to receive most of the attention. The final deliverable, the framework, is created by combining these tools and making them work together as one unit.

This thesis will be carried out using well tested simulation concepts and methods in order to create a framework that is reliable and that can be used to produce realistic data. Airlines and their workings will be studied to make sure that the framework can be applied to a multitude of situations to be of as much use as possible to an airline, but also to assert the quality of the framework.

1.4 Key findings

This thesis demonstrates how a simulation framework can be created that allows for realistic simulation of airline operations. The framework is mainly built using existing components along with a custom-built tool for crew simulation. The framework is then used to evaluate airline schedules and produce estimates for several of the values used to assess airline schedules.

Test data from a real airline is then used to test and evaluate the framework and to demonstrate the results of such a simulation and why such results are of use to an airline.

1.5 Thesis outline

In this, the first chapter, the reasons for this thesis has been explained. In the second chapter the scheduling problem and the process of schedule recovery are described. This chapter is mainly aimed at readers with no previous knowledge of airline operations and aims to give them a better understanding of the need for a simulation tool. Finally chapter 2 also describes some of the hurdles that need to be overcome when creating a good simulation and gives a description of the status quo of simulation tools at the time of writing.

Chapter 3, Framework Design, describes the theory behind the simulation framework and the components that are included in the final solution as well as the choices that led to this configuration. Chapter 4, Framework Implementation, describes how those components were realized and included in the framework. Chapter 5 describes the final framework and aims to evaluate its performance and usefulness. Chapter 6 discusses the results found in the previous chapter as well as some alternate uses for the framework. Finally, chapter 7 describes some proposed future work that can be performed both to improve the framework but also to use it as a tool for research of new areas.

2. Theory and Literature

This chapter serves as an introduction to the airline industry and how airlines operate. Readers with an understanding of airlines and how they perform scheduling and recovery may skip most of this chapter and continue at 2.6, Simulation, on page 15. What is discussed before that will mainly be an introduction to the aforementioned areas.

The outline for the chapter is as follows: Firstly statistics and figures for airlines are presented in order to give the reader a feel for the airline business climate today and what events that impact its operations. After this the major sources of delays are investigated to give the reader a better understanding of what causes them and how they can be battled. In parts 2.3 and 2.4 the scheduling and operations of an airline are presented to give the reader an idea of how airlines perform this task. In part 2.5 the concept of a robust schedule is investigated to show what is actually meant when robustness is discussed and how robustness can be created. In parts 2.6 and 2.7 difficulties in creating a simulation and how previous studies have dealt with them are discussed.

2.1 The global airline industry in numbers

On a global scale the airline industry has been growing almost every year during the first decade of the twenty-first century. It seems to be an industry where only major events such as the attack on World Trade Center or a global economic depression can cause a decline in emplaned passengers. Furthermore, the industry seems to be able to recover quickly from such major events. The fear of flying introduced by the attacks on September 11th 2001 seems to have vanished in 2003 where passenger rates grow once more and by 2004 the passenger growth was almost 15%. Similarly when the global depression struck in 2009 the number of passengers declined only to increase greatly the following year (International Air Transport Association (IATA), 2012). A summary of passenger growth can be found in Figure 1 below, where 2002 is used as index.

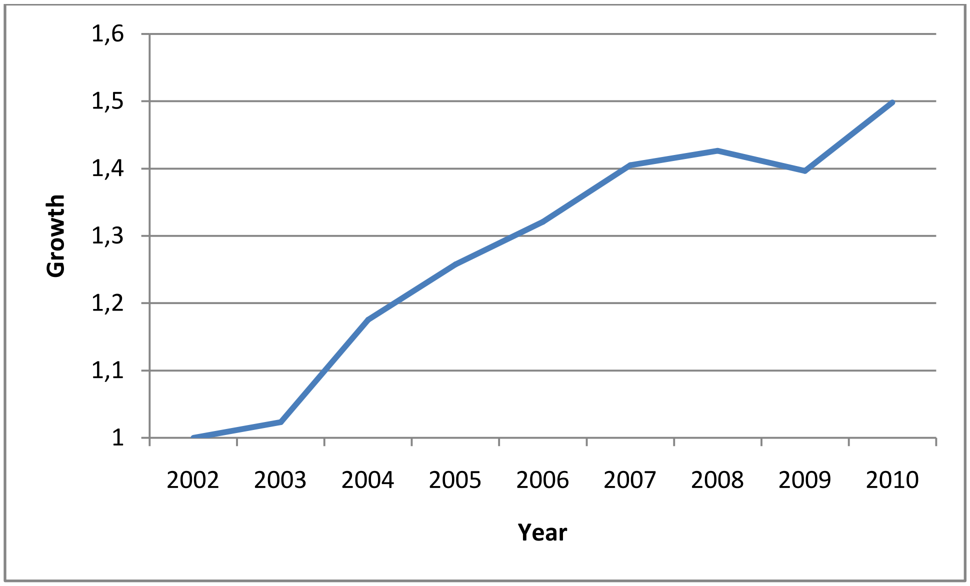

Figure 1. Global passenger growth using 2002 as index. (International Air Transport Association (IATA), 2012)

In spite of this the airline industry is not the money-making machine it's sometimes portrayed to be. As Belobaba wrote in "The Global Airline Industry": *"On a global scale and especially in the USA, the airline industry has been in a financial crisis for much of the twenty-first century"* (Belobaba, et al., 2009). This is hardly an exaggeration since US airlines reported losses for seven of the ten first years of 2000 and a total loss of over 55 billion dollars over the decade (Air Traffic Association of America Inc., 2010).

The surge in passengers have pushed major airports and frequently travelled airspaces to their maximum capacity. Airlines are therefore constantly battling congested airports and airspaces and even though passenger numbers have gone up, the number of flights departure has not changed in the same fashion. Figure 2 is a summary of the number of US flights departed and US Airline profit.

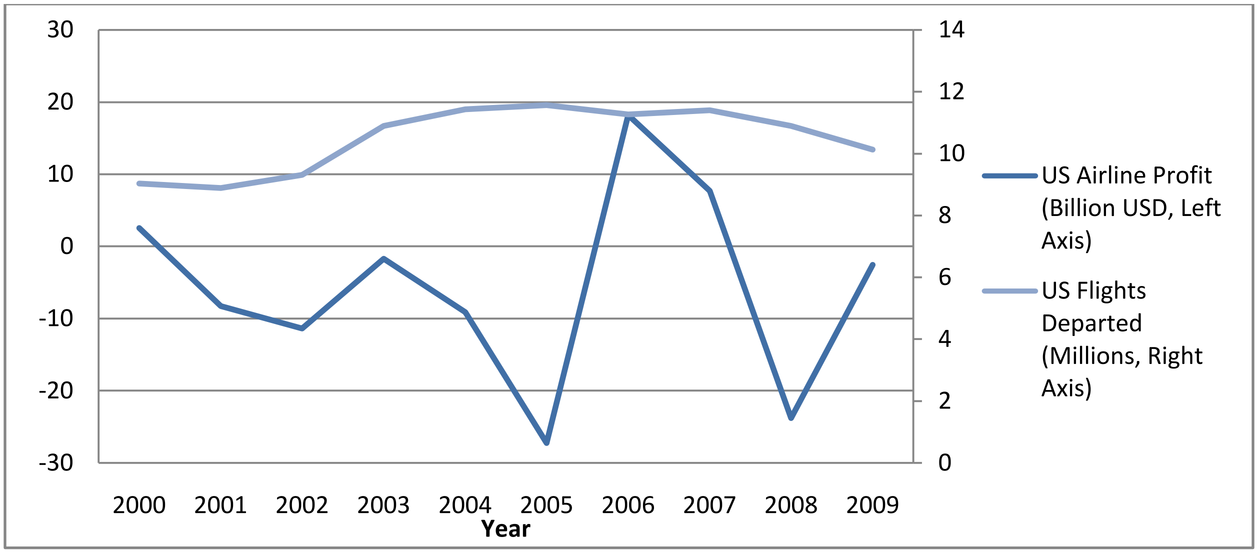

Figure 2. Net profit versus departed flights for US airlines. (Air Traffic Association of America Inc., 2010)

Clearly, more passengers and more flights are not enough to maintain profitability in an airline. In fact, more flights can actually be devastating since it adds further strain to an already congested airspace. In later years airlines have sought new solutions to these problems, by using larger airplanes, alternate airports, etc.

The reasons for the huge losses can however not be described simply as due to congestion. One factor that is often highlighted alongside the congested airports and airspaces is the dramatic rise in fuel prices. The price for one barrel of jet kerosene rose from $34.7 in 2003 to $91.4 in 2010 (International Air Transport Association (IATA), 2012). Crew-related costs also contribute greatly to the total costs of airlines and much research has been done on the subject of crew schedule optimization in order to minimize crew costs starting with (Stojković, et al., 1998).

Delayed and cancelled flights also play a significant role in the lack of revenue. In 2007 the estimated total cost to US airlines for passenger delay was 41 billion USD (Schumer & Maloney, 2008). This was caused by flight delays of a total of 4.3 million hours during this year. The total time wasted for US passengers during the same period was estimated to 320 million hours (over 37 thousand years).

The estimate only represents the direct costs to the airline and there are many other factors that may be considered when estimating the cost. An in-depth study on this matter was performed by Cook et.al. and provides detailed calculations for the costs of various delays (Cook, et al., 2004). While the total cost differs greatly with delay time, aircraft type and other factors, a cost of 72€ per delayed flight minute is suggested as an average estimate. This cost includes both direct costs to the airline as well as other losses such as a "passenger opportunity cost" which represents the cost for wasting the passenger's time. The costs suggested by Cook will not be discussed further in this thesis but deserves mentioning as it gives a more holistic view on the issue of delays and is a valuable source for the interested reader.

As can be understood from these figures, a growing market is not enough to sustain airline profitability and airlines must therefore focus on cutting costs. Unfortunately many of the costs, such as the price of jet fuel, remain largely outside the control of the airline. Whilst airlines have the option to negotiate fuel contracts with different airports and then fuel their planes accordingly, the global price of oil is still out of their control.

2.2 Delays and cancellations

As flights increase in numbers so does the percentage of delayed and cancelled flights, but the number of departed flights are not the only cause. In 2003 the percentage of US flights delayed more than 15 minutes on departure was 12% and 1, 5% of all flights were cancelled. In 2008, a bad year for delays, those figures had changed to 21% and 2, 9% respectively, yet the number of flights was similar. An explanation to this may be the fact that even though the same number of flights was performed, they were generally longer and more crowded, as both flight hours and enplaned passengers were higher in 2008 (Air Traffic Association of America Inc., 2010). In the years following 2008, the delays and cancellation percentages have decreased but still remain significantly higher than the values of 2003.

It is not only the US that is experiencing a surge in delays; in 2010 European Organisation for the Safety of Air Navigation (EUROCONTROL) estimated that 44.8% of all flights in the European region were delayed five minutes or more. The average delay time for a delayed flight was as much as 33 minutes (EUROCONTROL, 2010). It is worth noting that while the volcano eruption of Mt. Eyjafjallajökull did happen during April 2010 this did not have a significant impact on delays. On the other hand, the most critical days of the eruption (15^{th} to 22^{nd} of April) contributed greatly to the number of cancellations in 2010 and it is estimated that more than 100,000 flights were cancelled during this time.

During the same year EUROCONTROL estimated that 46.3% of delayed flights were delayed due to so-called *reactionary delay* which is delay caused by late-arriving crew and/or aircraft. This delay is sometimes also referred to as *propagated delay* since the true cause of the delay, known as *primary delay* has been propagated to this flight. The causes of primary delay for the same period can be seen in Figure 3.

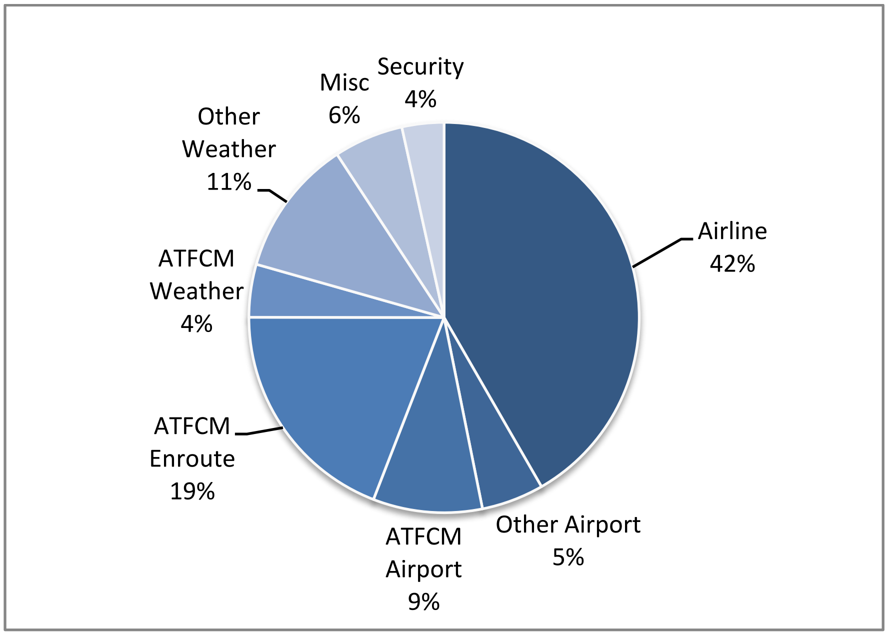

Figure 3. Sources of primary delay in Europe 2010 (EUROCONTROL, 2010).

The prefix "ATFCM" is an abbreviation for "Air Traffic Flow and Control Management" and can be signs of congested airports or airspaces as well as airports operating below maximum capacity due to poor weather. "Airline" indicates a delay that originated within the airline and includes faulty equipment, crew shortage etc

Propagated delay occurs when connections from a flight delayed by primary delay can no longer be made in the given time. An example of this is illustrated in Figure 4. Such delay can be partially or entirely accommodated for in the scheduling process by including some extra time, simply known as *slack*, between connections that are believed to be hard to make. Such connections do however come at a cost since resources are not used in an optimal manner if no delay occurs.

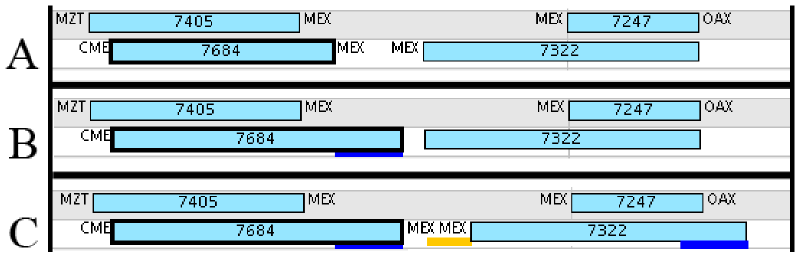

Figure 4. A. Original schedule for two aircraft. B. Flight 7684 is delayed on arrival. C. Delay propagates to flight 7322 which is delayed both on departure and on arrival.

As can be seen from Figure 3 the major sources of delay are caused by the airlines themselves and by congestion of both airports and airspace. If the reactionary delay is also considered then it can clearly be seen that a majority of the delays originate within the airlines.

Many of the recently founded low-cost airlines have delays in mind from the start and fly from less congested airports and have fleets consisting of only one type of aircraft to minimize maintenance issues and to simplify the scheduling task (Kohl, et al., 2004). For major carriers using only one aircraft type may not be an option, but there is still the matter of reactionary delay which is strongly dependent on the aircraft and crew schedules. Also these considerations are of great importance if one is to extend the current fleet or change destinations.

2.3 Airline scheduling

In order to understand how clever scheduling can prevent delays it is first important to understand how scheduling of aircraft and crew work. While the details of the scheduling may change from airline to airline they often solve the problem in similar ways. This process is described below.

Creating schedules for aircraft and crew in such a complex system as an airline is no easy task. The problem of scheduling is one that appears in very many situations and already the seemingly simple task of job scheduling is known to be NP-complete for as little as two resources. The problem that airlines are faced with is more complex and constrained and consists of designing a schedule that is legal, fair and cheap. As can be expected there are tradeoffs between these three criteria, where legality can be seen as a hard constraint but where fairness must be weighed against cost. These problems apply not only to airlines and related research can be found, for example, on what is known as the Nurse Scheduling Problem which aims to create schedules for nurses that satisfy these constraints (Goodman, 2009).

The problem for airlines is further complicated by the fact that many activities, most commonly flights, start and end at different locations. It is important to stress the fact that these schedules also contain activities such as maintenance or training that start and end in the same location. Such activities may only be carried out at certain locations. For instance, aircraft maintenance is most commonly performed at only a few of the airports that an airline services. Apart from these additional constraints there are the various rules and regulations, which may vary between airlines, which affect both crew and aircraft and thus the legality of the schedule. One such rule that is often in place dictates that certain crew members may only fly certain types of aircraft.(Andersson, et al., 1998)

Because of the presence of such rules, but also to limit the complexity of the problem, the scheduling is commonly divided into several sub-problems. These problems are then solved individually to create the final solution (Sohoni, 2006).

The first step, referred to as *pairing*, is concerned with combining resources, such as crew and aircraft types, with activities such as flights. For crew a set of consecutive flights is known as a *duty*. A pairing for crew is a set of duties that not violate any rules. It is important to understand that in this step only *types* of aircraft and crew are considered, not the actual crew members or aircraft themselves. A pairing can therefore be seen as an anonymous schedule.

Solving the pairing problem is usually done by representing it as Set Cover problem which can then be solved using several methods, such as column generation (Barnhart, 1994). Also algorithms of a more random nature such as genetic algorithms have been experimented with for solving this problem (Kornilakis & Stamatopoulos, 2002) (Zeren & Özkol, 2012). However, the random algorithms are still in a very experimental stage and need plenty of improvement before they can be used in reality.

It is also worth noting that it is usually not possible to solve the pairing problem for crew until after it has been solved for aircraft. This is because it is often only after this step that it is known what type of aircraft that will be used for a specific flight. Since, as mentioned before, certain crew members may only fly specific types of aircraft this must be known before crew pairing can be made.

In the second step, known as *Tail Assignment* for aircraft and *Rostering* for crew, specific resources are assigned the pairings that was created in the previous step (Kohl, et al., 2004). This produces final schedules for each individual resource.

It may seem like a waste of resources to split pairing and tail assignment or rostering, but this is usually not the case. With the pairing solved airlines can still make strategic decisions regarding the schedule without having to worry about the effects on individual resources and how pairing fits individual activities such as maintenance or vacations. Finally, rostering is not something that is done in the same way for all airlines. Some airlines feature complex models where crew can bid for certain trips and where senior crew members are given priority over less senior coworkers. By creating this split the pairing step can still be performed without taking these further complications into account.

When creating crew and aircraft schedules it is usually assumed that the schedule will be executed with few or no disruptions at all (Schaefter, et al., 2001). Several schedules are generated, and usually the one with the lowest cost for execution is chosen. In reality, schedules are almost never executed without disruptions and airlines must adjust them as they are carried out. Such adjustments have the potential to be very expensive and these costs are not considered when choosing schedules. This means that schedules with some inherited robustness such as extra slack between flights may have been the better choice, but they are discarded due to their higher initial execution cost.

Some airlines today use what is known as *degradable scheduling* to prepare for major disruptions. This technique is also sometimes known as creating *snow plans* or simply as *schedule thinning*. These kinds of schedules are designed in tiers where low-priority tiers can be discarded in case of major disruptions. This gives airlines clear pre-planned strategies for recovery and allows for creation of lower bounds for revenue. Furthermore, when discarding a tier, resources become available to aid in the execution of the remaining tiers. This is a major improvement over not having these strategies, but it provides no information of what can be improved or how likely a certain degradation is. Degradable scheduling is covered further in 2.5, Robust scheduling.

2.4 Airline recovery

During airline operations delays and other events may cause disruptions to the schedule and must be handled. This phase is known as *Disruption Management* and can be simplified as the process of fulfilling the following three conditions (Kohl, et al., 2004)

1. Get passengers and their luggage to the destination on time.
2. Minimize costs for crew, compensation, accommodation etc,
3. Get back to the original plan.

There is an obvious contradiction between points one and two and this also illustrates why the recovery problem is fundamentally different from the scheduling problem. In the scheduling phase a "good" schedule can rather easily be defined as one with a low cost associated with it. During recovery the definition of "good" is not as simple as softer factors such as customer satisfaction must be carefully regarded.

Finally, the third point is somewhat debatable. Returning to the original schedule is something that is best classified as conventional wisdom and might not be needed. There may be benefits associated with returning to schedule, for instance if there are union rules that prevent many changes of crew schedules. Another reason suggested by Kohl et.al. is simply to limit the complexity of the problem. Returning to schedule limits what decisions that can be taken, and the end result is known to be good. Not returning to schedule gives a vast array of choices to make and their outcomes may not always be good.

In order to return the disruptions that occur airlines have several choices. The most common ones are listed below.

- Delay flight departure until aircraft/crew is ready. This is the most simple of all approaches and flights are simply delayed until they are ready to take-off. Scenario A in Figure 5 is an example of this.
- Aircraft /crew swap. If there is another aircraft/crew at the base that can fly the disrupted fight then the two can be swapped in order to prevent propagated delay. This process can solve the problem at hand, but may introduce new problems later on. How swaps can reduce delay can be seen in Figure 5 where the situation from Figure 4 is resolved using a swap.
- Calling in standby/reserve crew. Most schedules include standby crew that are on duty, but not assigned to a particular flight. These crews can be used to relieve stressed crew that cannot make their connection. Such crew is usually only available

at the major hubs of an airline and is preferably used in such a way that both they and the crew they replace return to base at the end of the duty, otherwise additional costs for hotel nights etc are introduced.

- Deadheading. This is the term for crew members that are onboard an aircraft but not working. This crew can for example be reserve crew that deadhead to a certain location in order to fly a plane back or vice versa.

- Ferry flights. Flights from one location to another without any passengers onboard are known as ferry flights. These are very costly but may be necessary in order to make an aircraft schedule work.

- Cancelling flights. While not an attractive option it is sometimes best to cancel flights. The effects of cancelling one flight varies greatly with the schedule but often leads to even more cancellations since both aircraft and crew can no longer make their next connection because they are now at the wrong airport.

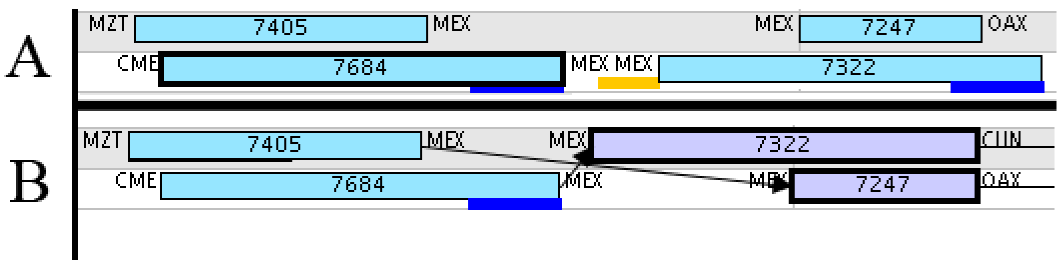

Figure 5. A. The propagated delay from Figure 4, part C. B. A crew and flight swap is made for flights 7322 and 7247 (arrows indicate *old* schedule) which prevents the delay from propagating.

The process of finding the right recovery choices is often a complex task since resolving one disruption might have effects that cause infeasibilities elsewhere in the schedule. Also disruptions caused by regional events such as bad weather affect several flights at once which may force the planners to solve several disruptions at once, further complicating the problem. For this thesis a recovery approach developed by The Technical University of Denmark and Jeppesen Systems AB in collaboration with the European Union and British Airways, which is part of the DESCARTES project, is used, for more details on this solution, see (Løve, et al., 2001) and (Tiourine, et al., 2010) .

Sometimes solving disruptions here and now may cause disruptions later. For instance, if the crew for flight 7247 in Figure 5 were supposed to end their duty in OAX they will now find themselves in CUN instead. If we assume that the next flight in their schedule returns them to OAX, they can end their duty unless there are rules that prevent them from flying that flight. If this is the case then this flight is suddenly without crew and might need to be cancelled. This shows that one needs to take the whole schedule into account when resolving disruptions since their effects can spread far beyond the disrupted flights. This also hints as to why airlines are sometimes reluctant to giving precise reasons for delay and cancellations to passengers. Few passengers would understand why their flight is suddenly without crew as they expect that both crew and aircraft must exist in order to sell tickets to a flight. Explaining that they are without crew because of a delay to a different flight several hours earlier will probably not improve the situation.

Finally, the problem of having many aircraft types appears once again. Since crew is usually only licensed to fly one particular aircraft and because aircraft vary in capacity, the numbers of valid swaps decrease as the numbers of aircraft types increase. Also, reserve crews become less useful as an airline might have to have one reserve crew for each aircraft type.

In reality, a solution to a disruption must not only take aircraft and crew into account, but also passengers and luggage, further complicating the problem. To an outsider the concept might even seem to happen backwards. It would seem that airlines should first accommodate their passengers, and then solve their own problems. However this is not financially possible or even feasible as aircraft and crew issues need to be resolved before it can be determined how passengers can be accommodated. However, the cost of passenger recovery should be considered when recovering crew and aircraft as an expensive crew recovery might be balanced by getting all passengers to their destination rather than choosing a cheap crew recovery but having to pay for hotels and other forms of compensation for the passengers.

To further complicate the problem, some decisions are taken by the pilots, rather than the planners. An example of such a decision is diversion of a flight due to poor weather. This is a safety issue that the pilot must assess in the air and cannot be determined from the ground. If the pilot chooses to divert the flight, this creates a disruption of the schedule that the airline must resolve.

The process of recovery is also very different from scheduling in the fact that it needs solving in real-time in order to prevent further disruptions from occurring (Schaefter, et al., 2001). This is an area where software for decision support is believed to become increasingly useful to airlines in the future. The area is the focus of significant research but airlines are generally slow to catch on and start using such software.(Lan, 2003).

There has, however, been an increased interest in these problems, as the costs of poor disruption management have risen greatly with increased congestion and a few eventful years for delays. Such software has also proven to be extremely useful in getting back to plan after major disruptions. Examples include the usage of software known simply as CrewSolver at Continental Airlines, which was used for a large snowstorm on December 29th 2001. By using the CrewSolver Continental Airlines could return to normal operations several days before that of competing airlines. Over the following year it is estimated that the same solver saved Continental $40 million (Yu, 2003). While much progress has been made in the field of automation since then, and even though the benefits are clear, the recovery software is by some estimated to be where the planning software was some 15 years ago (Belobaba, et al., 2009).

Even with sophisticated recovery tools the airlines must decide on a strategy for recovery. Examples of variables that make up a strategy include whether long delays are preferred compared to cancellations. Furthermore rules for crew such as a limit on the number of schedule changes allowed per month can lead to strategies where few crew changes are made unless faced with serious disruptions, to make sure that such disruptions can be dealt with later on. This may however lead to sub-optimal recovery during normal operations since the recovery tools might not always make use of the available staff.

2.5 Robust scheduling

It is not hard to see that airlines would like to minimize the number of times they need to perform disruption management, since it is both costly and hard. In order to achieve this, airlines can perform pre-emptive actions to limit or mitigate the effects of disruptions. These actions are additional constraints that may be added during the scheduling process in order to obtain schedules more resistant to disruptions or that reduce the ripple effects of delays and cancellations. A few of these are described below:

- **Round trips**. An airline may schedule flights so that they as often as possible fly back and forth to a destination. If this is done then if major disruptions occur then the airline can cancel both flights and be sure that they can execute the schedule as before after that.

- **Crew follows aircraft**. While it might be more cost efficient to change what crew that fly a specific aircraft it is also something that propagates disruptions. If a flight is disrupted then it may affect several other flights if, for instance, cockpit and cabin crew both fly different aircraft afterwards. By having crew work the same aircraft for the entire duty the delay propagation is at least constrained to those flights.

- **Allow for slack in the schedule**. While clearly not cost efficient, some extra slack between activities can absorb smaller delays and prevent delay propagation. If no disruption occurs then such slack leads to sub-optimal utilization of resources. It is therefore important to figure out where to place slack for it to be used optimally. Returning to Figure 5 we can see that adding slack to flight 7322 which suffers from propagated delay is not very useful since we can use the slack already present between flights 7405 and 7247 by swapping flights. This means that smart allocation of slack between two flights may be used by other flights and resources can be better utilized.

- **Degradable Scheduling**. By creating schedules that are built with a pre-decided degradation plan, airlines can plan ahead and have fallback-schedules and rules for prioritization in place before disruptions occur. When disruptions do happen the airlines can fall back to another schedule where certain flights are cancelled. This frees some resources that can be used to solve the disrupted flights that remain in the schedule. By having the degradable schedule in place from the start, many of the problems associated with cancellations and delays disappear. The downside is that one might degrade too much and therefore excessively cancel or delay flights.

Kang even goes so far as to suggest that airlines could also sell their tickets with different reliability levels and, by doing so, mitigate the customer unpleasantness that occurs from having your itinerary disrupted (Kang, 2004).

- **Planning for cancellations and alternate schedules**. Airlines can keep alternate schedules of "what-if"-scenarios for the flights most likely to be disrupted or for major events that may occur. Examples of such events include imminent snow storms, and other weather related events but also events such as union actions. These problems can then be solved without the added time-pressure of disruption management. This is similar to the concept of degradable scheduling and can be seen an operational version of the same idea. The major difference lie in that planning for "what-if" scenarios focuses more on when additional resources will be needed and works with shorter time spans as well as more real-time oriented data.

By using these measures airlines hope to minimize the number of disruptions that occur. All of these techniques come at a cost however, and there is always a tradeoff between schedule robustness and schedule execution cost.

2.6 Simulation

One method for deriving possible outcomes of an executed schedule, and the costs associated with them, is to use computer simulations. This is a commonly used tool when attempting to estimate or analyze the performance of complicated systems. When airports are constructed or modified simulation tools are often used to simulate the effects of passenger flows (Gatersleben & van der Weij, 1999) or ground operations, such as de-icing of aircraft etc (Norin, 2008). Simulation is also a commonly used tool to keep airline crew training up to date.

Much work has also been done in the field of airspace simulation. Several simulation tools exist that can realistically capture the behavior of over-crowded airports and the massive delays, both primary and propagated, that these incur. They are further discussed in 2.7, Related work and 3.5, Airport and airspace simulation. Together these simulation tools cover many of the aspects of an airline, but one area that remains largely uncovered is simulation of crew and their effects on airline schedules. There are several reasons for this:

- Crew is very strongly dependant on aircraft. Due to this dependence a crew simulator would not be very beneficial without also having aircraft simulation.

- While aircraft behave in similar ways for all airlines this is not the case for crew. The rules that govern crew and their actions are often quite complex, and more importantly, differ greatly from one airline to another. This makes the task of simulating crew even more difficult since it is a lot harder to make a tool that is of use to many airlines.
- Due to the complex nature of the crew rules, crew simulation also requires a complex model to determine when and where an *event* such as arriving late for work, or disease, causes an actual *disruption* of the airline schedule, such as a delayed flight.
- Crew events are not considered to be a major source of disruptions.

When faced with the above facts it is perhaps hard to see why crew simulation needs to be performed. The problem is complex and the impact is expected to be small, and this is usually the reason why many simulation tools lack a crew model.

The importance of crew appears when attempting to simulate recovery. This can be done by adding recovery to existing aircraft simulations. This may seem like a logical and straightforward solution, but by doing so one major issue is overlooked. If one considers the complex rules that govern crew, it is very likely that crew will greatly limit the choices available during recovery. This, in turn, affects the actions that airlines can take when faced with *any kind* of disruption. Therefore, if no crew model is present, or if this model is greatly simplified, then the recovery performed by such a simulation might not be feasible in the real world. This is a good reason not to dismiss crew simulation on the grounds that the direct effects are small.

This highlights a major hurdle in creating a realistic airline simulation: To involve simulation of the recovery performed by the airline. Recovery is, even with the assistance of sophisticated software, a process which requires making decisions that seems to require human intelligence. This is especially evident in situations where no solution to the problem exists. In this case the solution that is “least bad” must be chosen. While it is possible to model such a choice it is not usually something that is done by recovery software today, instead such choices are left to the planners. This means that a simulation tool must attempt to model such choices in order to realistically reflect the real world.

2.7 Related work

In the field of airline scheduling and robustness much interesting work has been done, apart from the ones already mentioned, including several models to evaluate robustness and to estimate propagated delay. Notable examples include (Schaefter, et al., 2001) who introduces an element of uncertainty to the scheduling process, (Lan, 2003) who suggests new planning techniques for minimizing delays and (Lapp, et al., 2008) that uses a recursive algorithm to create estimates of delay propagation of an existing schedule. Others such as (Smith, 2004) propose that limitations to the scheduling problem can actually increase the overall robustness and cost efficiency of airline schedule. Smith demonstrates this concept by limiting the number of aircraft types that may serve a particular airport, a technique known as “station purity”. Many others such as (Michelle, et al., 2012) improve further on these techniques using more integrated models for aircraft routing and crew pairing. A common factor for much of the related work is that the presented methods do not take into account the process of recovery. Recovery can potentially have a large effect on the schedule and therefore change the outcome drastically.

The two major platforms for aircraft simulation are MEANS, the MIT Extensible Air Network Simulation, developed at Massachusetts Institute of Technology (MIT) (Melconian, 2001), and SimAir, developed at Georgia Institute of Technology. MEANS is a tool for simulation of aircraft and airport operations and can be used to simulate effects of factors such as poor weather and airport congestion (Clarke, et al., 2007). MEANS has since its creation been extended in several ways including simulation of the United States National Airspace System to allow for better simulation of airspace congestion within the United States (Whittaker, 2006).

SimAir is another simulation tool available that can realistically simulate airline operations, including unscheduled maintenance and recovery, but unfortunately also SimAir lacks a complete crew model (Rosenberger, et al., 2000). SimAir has since its creation been used to validate new techniques in airline operation (Schaefer, 2006).

Attempts to create simulations that simulate crew and recovery have been made in the past. (Rabbani, 2004) attempted to extend MEANS to include schedule generation, aircraft simulation as well as airline recovery. A rough crew model and a model for recovery was created for this simulation. The rules that govern crew are based on the rules set by the Federal Aviation Administration, a part of the U.S. Department of Transportation. While this was a major step in simulation techniques this simulation did not reflect the true complexity of crew rules and recovery, and as a result was therefore less useful to airlines.

3. Framework Design

The simulation problems described in 2.6 and the lack of crew simulation discussed in 2.7 have been central in the design of the simulation framework presented in this thesis. The aim is to create a framework that realistically captures the behavior of an airline including advanced recovery. This provides a much needed additional dimension to the simulations present today, as schedules are not expected to be executed without undergoing several changes.

Such a framework calls for sophisticated models of both aircraft and crew, but also for powerful recovery tools. Because of this the framework has been constructed by adding simulation capabilities to recovery software rather than the other way around, as is usually the case in related work. This allows the framework to make use of already existing industry-standard models which feature complex rule sets, which can be used to perform advanced recovery in a very realistic manner.

This chapter provides an overview of the simulation framework and its internal components. The various components are described and their role in the framework is discussed. For details on the implementation and modification of the included components see Chapter 4, Framework Implementation.

3.1 Simulation model

Before any simulation can be made the environment which needs to be simulated must be considered. The real world environment must be studied and a model created that captures the features that need to be simulated. The available data and tools must also be studied in order to ensure good results (Law, 1990) (Robinson, 2004). In this particular case it is easy to determine that the simulation should depict the day to day operations of an airline. However, this leaves a number of questions open for debate. Firstly a level of realism must be decided upon. For the purpose of this thesis the framework is limited to capturing only aspects that are deemed to have a significant impact on the final aircraft schedule are to be simulated. This includes unexpected events such as airport closure but also regular everyday events such as crew checking in for work, which may also be sources of delay.

While it is likely possible to simulate operations on a smaller scale, including more exact models of how crew move in airports, individual taxi times for aircraft depending on what runway they land on etc, these factors are deemed small and hard to model in a realistic fashion. Large amounts of data not readily available to the public would need to be collected and analyzed before such a model could be created.

Finally it needs to be highlighted that the simulation framework presented in this thesis will not feature a passenger model. This limits the predictions of revenue and disrupted flights somewhat but still allows for understanding of delays and other important factors. However, disrupted itineraries and other passenger-related values can still be estimated using this framework as explained in chapter 6, Discussion.

The high-level model can be seen in Figure 6. It features a controller, two modules for simulation and one module for disruption management. The controller requests simulation and disruption management from these modules and by doing so moves the simulation clock forward.

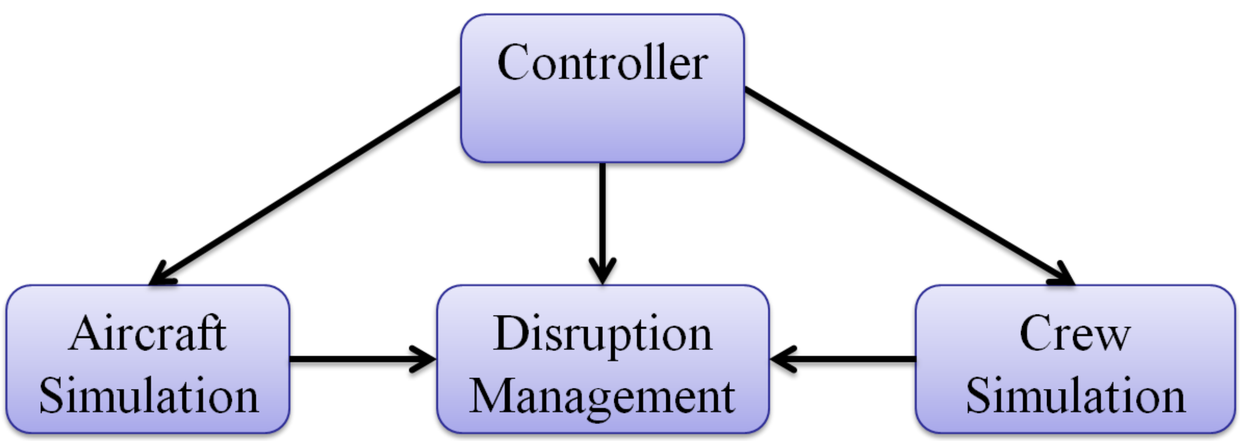

Figure 6. High-level simulation model

3.2 Data set

Important aspects such as the aircraft fleet, the structure of the flight network are all factors that dictate what happens when disruptions occur and therefore have a large impact on the output of the simulation and must therefore also be decided upon before a model can be created.

In order to simulate an airline the data that is used as input to the simulation, such as flight schedules, must be deemed trustworthy to ensure realistic results of the simulation. Also data of actual airline performance under similar conditions must exist, so that the simulation results can be validated. It is only after this has been achieved that the simulation can be deemed realistic enough to simulate the effects of other situations where estimates are desired.

This gives rise to yet another problem: The data described above is usually business sensitive information. Very few, if any, airlines are comfortable with sharing their data in such an open manner. While flight times can be gathered from public sources, such as EUROCONTROL or the American Bureau of Transportation Statistics (BTS), exact schedules are not known, especially not for crew. Also, when collecting data from public sources, only the flights actually flown are presented. This gives little or no information about recovery actions performed by the airlines. This conflict between usefulness and classification usually forces researchers in this area to make a choice of either using realistic, but not real, data generated with regards to some airline, or using real data from an airline, or parts of a real airline, but without publically presenting sensitive results.

For the purposes of this thesis, a third option has presented itself. In August 2010 the Mexican airline Mexicana de Aviación, commonly and henceforth only referenced as Mexicana, ceased operations (Bloomberg, 2010). As a result of this both historical data as well as actual schedules which are no longer considered business sensitive have been made available to the author. For the purposes of this thesis this data will be used as input.

It should however be pointed out that publically available data is very useful when determining the simulation environment. For instance, information about congestion at various airports can be determined from such data, and this is a very important factor for a successful simulation.

IATA Airport Code	Ratio
MEX	67%
CUN	17%
GDL	17%
LAX	14%
MTY	13%

Table 1. IATA Airport codes and the percentage of all flights departing or arriving at this airport.

Mexicana was in many cases a typical hub-and-spoke airline. Such an airline is defined as an airline that has flights that take off or depart from the airline's major hubs, rather than attempting to connect all of the serviced cities to each other. This can be troublesome for passengers, as they always need to travel through one of the hubs, but usually provides more smooth operations for the airline.

The three hubs along with a few of the more frequent destinations can be seen in Table 1. The percentage indicates percentage of flights that *arrived at or departed from* that airport. For a more holistic view, the most operated airport pairs of Mexicana can be seen in Figure 7. When comparing these sources of data it becomes clear that Mexicana serviced several airports through its three major hubs.

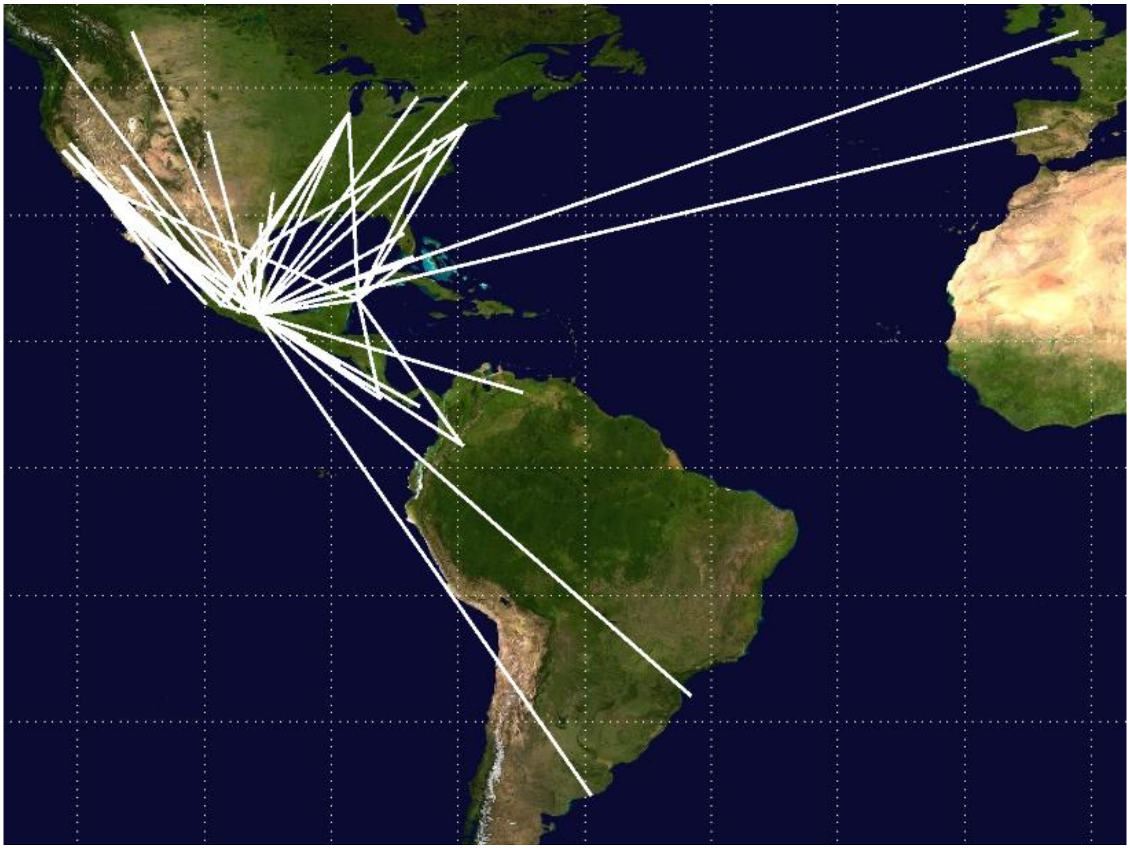

Figure 7. The most frequently operated airport pairs by Mexicana.

3.3 Simulation controller

The *Integrated Simulation Controller* (ISC) is the heart of the simulation and its role is to drive the other components forward. It is also up to ISC to monitor the simulation and to act on various problems that are encountered during simulation. How exactly this is done is configurable by the user and may change from one simulation to another, or even during the course of one simulation.

Users configure ISC by creating an XML-file that describes the different steps taken by the simulation. This format allows for as simple programming-like environment while at the same time maintaining human-readability. This allows users to dictate the actions taken by ISC without knowing how the components work internally or having much programming experience.

ISC advances the simulation in small time-steps and in each of these steps the different parts of the simulation that are specified in the current configuration are run until they match the current simulation time and then halted. Each part of the simulation is responsible for submitting updates caused by simulated events to any other parts of the simulation framework that may require them. This keeps the simulation synchronized over time. Once such a component has finished running and has updated all relevant components they signal ISC that they are done.

Any recovery-components that exist in the simulation framework are called just like any other part of the simulation, and just like the other components they may create new disruptions. While this might seem strange at first it can be justified by the fact that if no solution exists then a recovery tool might still update the schedule to limit the effects of a previous disruption. This will fix problems associated with the original disruptions but new problems may also arise.

As a final precaution ISC may also directly step in and modify the schedule. This can be utilized in a situation where there are infeasibilities in the immediate future of the simulation. In such a situation it would not make sense to advance the time further and ISC can therefore be configured to step in and perform cancellations of flights, or other actions, in such an event. Since this can technically be considered recovery it is not a recommended action, but might unfortunately be necessary in order to simulate the decisions taken by planners where recovery tools can´t solve the problem at hand.

Before proceeding to the next step of the simulation ISC also performs a final check that the simulation is feasible within the immediate future. If this is not the case then ISC halts and the simulation is aborted. It is also worth noting that ISC does not halt if infeasibilities are found further ahead in time. This is due to the fact that it is not uncommon for flight schedules to have errors that are solved as the day progresses. For instance, a cancelled flight might actually free up the resources needed to fix another disruption.

3.4 Disruption management

The central data store for the simulation framework created by this thesis is the data model used by Jeppesen Disruption Management (JDM) (Jeppesen Systems Ab, 2010). This data model models not only flights, crew and aircraft but also the rules that are in effect for this airline. By using this as the central data store the recovery tools in the JDM suite can perform recovery without any modification, and thus allowing for efficient and well tested recovery.

The framework proposed in this thesis is built for use with the JDM suite to allow for recovery similar to that used in the industry. This imposes a restriction on availability for this tool but this restriction is considered to be outweighed by the associated benefits.

The benefits gained from JDM are many. JDM features, as mentioned before, a model for representation of both aircraft and crew along with a complex model to describe the rules that govern these areas. The rules are formulated using a programming-like language known as RAVE and can be updated or changed at any time. Jeppesen also provides rule sets that contain rules similar to those used by real airlines thus creating a solid ground for realistic crew simulation. By using these resources the simulation framework can support the introduction of new rules at any time as well as allowing for simulations with less constrained rules.

Furthermore JDM contains industry-leading optimization software for recovery support. This software makes use of sophisticated algorithms for solving recovery problems. This is used by the simulation framework to achieve automated recovery of both aircraft and crew and makes for very good simulation of airline recovery.

3.5 Airport and airspace simulation

Airport and airspace congestion are two topics that have been frequently studied over the past two decades. Such congestion leads not only to passenger delay and loss of revenue for airlines but also to increased pollution due to excess fuel burn for aircraft (Simaiakis, 2009). Along with suggested improvements to systems for air traffic control and new ideas for airline schedules, these studies have also included sophisticated tools for simulation of traffic of entire air systems, including behavior of individual aircraft during all stages of a flight.

Several of these tools have been used to successfully simulate airport operations under various circumstances. For the purposes of this thesis the MIT Extensive Air Network Simulation (MEANS) will be utilized for aircraft and airport simulation. MEANS was developed as a part of a study of nonstop routings (Melconian, 2001) and has been extended since. MEANS was chosen for this thesis due to its extension-friendly program structure and because of the many different sources of disruptions, such as weather, congestion, air traffic control, etc. that MEANS is capable of simulating.

MEANS is a discrete-event simulation tool that models an entire airspace. Aircraft are modeled during all major stages of a flight. These stages are taxi in and out, take off, en route, landing, as well as at gate. Delays can be introduced at any of these stages and affects the simulation in different ways. A summary of the types of the delays introduced at each stage are summarized below

- Gate – Delay due to turnaround times
- Taxi – Delay due to congestion of runways or taxi-lanes
- Take-off and Landing – Delay due to air traffic control or congestion
- En route – Delay due to airspace congestion or weather (strong winds etc)

Apart from this MEANS can also initiate so called *Ground Delay Programs* or *Slot Restrictions* that delay flights from leaving the gate if they are scheduled to arrive at a congested airport. Such flights must wait to be allocated a landing slot at the affected airport before giving permission to depart. The same applies for aircraft leaving a congested airport.

How these delays occur vary greatly with the parameters. MEANS allows for customization of airport capacities, flight times, taxi times, weather data as well as additional air traffic, known as *padding flights.* Padding flights are flights that are not part of the simulated airline but may create congestion at certain airports and therefore affect the outcome of the simulation.

The weather data along with a tower module within MEANS, which acts as airport control, decide actual the capacities of the airports at any given time. These capacities determine how many aircraft may land or take off. Further description of MEANS, its internal structure and its inputs can be found in “MEANS - MIT Extensible Air Network Simulation” (Clarke, et al., 2007).

3.6 Crew event generation

Just as MEANS acts as an event generator for aircraft in the simulation framework, a similar structure is needed for crew. While the reasons for the events that affect crew are too many to list, the events themselves can be categorized rather easily; crews are either delayed, or they do not show up for the scheduled activity.

The effects of crew events are determined by the rules present in disruption management and the disruptions that they may cause must therefore be generated by this part of the framework. It is important to highlight the difference between an *event* and a *disruption*. An event is characterized by something happening such as a crew connection between flights taking longer than usual. A disruption only occurs if this actually has an effect on the actual schedule, that is to say, the next flight is delayed due to the late arrival of crew.

This separation allows a Crew Event Generator (CEG) to consider only the sources of events and the events themselves. While the sources of delay can be almost anything it is important that the event generation performed by CEG can be easily adjusted and tuned, to be fitted to real scenarios.

3.7 Output

The number of parameters that can be utilized to analyze a schedule are many, and more can be expected in the future. There are many different ways of determining whether the execution of a schedule can be determined to be good or not and not all of these can be implemented. To cope with this the simulation framework stores activities performed during each step in a way that allows them to be played back afterwards.

Also, the database that contains all simulation information is stored at the end of the simulation. Information stored here includes scheduled and actual start and end times for all activities, all assignments of crew and aircraft etc. For flights additional data such as when the aircraft leaves the gate, takes off etc. are stored. This allows for data mining of important parameters to be added as a post-processing tool. As a result simulations can be analyzed not only as they progress but also afterwards. In addition to this the simulation framework also generates a summary of selected key values at the end of a simulation that can be easily accessed though a web browser.

4. Framework Implementation

The components mentioned in Chapter 3, with the exception of the crew event generator, may seem like they can be used without much modification, but this was not always the case. During the creation of this framework several adjustments have been made. Many of the components need to be provided with very specific inputs in order to properly simulate the desired environment, and also some adjustments needed to make data sharing between the different components possible.

This chapter describes the modifications made to the different parts, and how the relevant input data was obtained. It also describes, in detail, how the Crew Event Generator was created and how it works.

4.1 Data set

The data used for the purposes of this thesis is, as previously stated, real data used by Mexicana. This offers great benefits compared to generating data since no assumptions need to be made, neither does the realism of the data be validated. It does, however not come without some problems.

During airline operations it is not uncommon that the current schedule contains disruptions over the next couple of days. This is natural since, for instance, a flight swap might solve problems today but create problems next week. Also, planners normally operate on specific time-windows and/or specific fleet types, where one group of planners is responsible for each such area. These different parts are synchronized at given times, but in between these times there may be inconsistencies in the data set.

The data used for this thesis is a live snapshot from a day of operations. Because of the reasons described above some work was devoted into cleaning the data set so that simulations could be performed without having to handle disruptions caused by previous actions. While this caused changes to the schedule all departures and their times are intact. This means that all of the flights flown during simulations are the same as the ones that would normally have been flown. Normal day-to-day disruptions were not altered as they would be expected in an ordinary schedule. Some of these are severe and not always possible to solve without cancelling flights.

It is not uncommon for airlines to utilize different systems for various parts of operations. In the case of Mexicana, this applied to the crew check-in process which was largely handled manually. Crew check-in is performed by crew upon arriving to work, and consists of swiping a card in a terminal to signal to the system that they are ready for duty. Late check-ins or crew members not checking in at all are causes of disruptions and since this needs to be simulated the Mexicana data had to be updated with rules and historical check-ins for crew.

Finally the two subsidiaries of Mexicana, Mexicana Click and Mexicana Link, were ignored as no crew data was available for these companies.

4.2 ISC

ISC, the *Integrated Simulation Controller*, is a project developed by Jeppesen Systems AB, but where development had been halted. This thesis is built upon that same controller, but has made several adjustments and improvements to it to allow for more realistic and stable simulations.

ISC is a tool written in C++ that uses XML files as simple recipes for simulation. ISC is as previously described often not directly involved in updating the current state of the simulation, but acts more as a communications hub that notifies the other components of when and how they should run.

When constructing the framework the functionality of ISC has remained similar. The most obvious exceptions include the addition of a stage where ISC can delay flights along with a more clever way of cancelling infeasible flights. This allows ISC to hold flights that are waiting for crew that are late for work or have tight connections and if ISC is forced to cancel flights it now considers more factors such as round trips, to avoid propagating the consequences of the cancellation to other parts of the schedule.

While the functionality of ISC has remained the same, much work has been done to improve the stability of the simulation process. As it is often needed to simulate the same conditions over and over to guarantee a certain level of statistical significance then it is vital that the simulation can run unsupervised for a long time. There were a few situations where this could not be guaranteed and much work has been made to ensure that it now is.

4.3 MEANS

MEANS and its connection to ISC were already in place at Jeppesen Systems AB before work began on this thesis, and this connection was used without modification. Using this connection ISC signals to MEANS that it should refresh its data from the data model in JDM, as it may have changed since MEANS was last run, and simulate for a given amount of time. After this time has passed MEANS sends the results of the simulation back to JDM.

In order for MEANS to simulate it needs data regarding weather, taxi times as well as flight times. In the absence of such data MEANS will attempt to estimate flight times, taxi times and simulate fair weather. While this estimation is better than what one might expect it still includes some assumptions on flight times etc.

In order to reduce errors from such assumptions values for both flight times and taxi times were extracted from previous Mexicana flights, available in the data. The data was gathered from all locations where Mexicana had flown at least 10 flights from one airport to another and was gathered independently for each departure and destination airport. The results produced from this are therefore believed to be free from significant statistical errors and take into account that flights times may differ with direction. Jet streams are constant flows of air that airlines may make use of while flying from one destination to another, but not when flying back, since these winds never change direction. This can cause significantly different flying times for flights between the same airports depending on direction.

Since weather data is strongly dependant on the situation that needs to be simulated this data was not generated but is rather something that can be customized by the user.

4.4 Crew event generator

The crew event generator (CEG) is the one single component that has taken up the most of the time taken to develop this framework. Simulation of crew is an area that has remained largely uncovered and there are limited previous studies that can be built upon. Therefore simulation techniques needed to be carefully studied as well as deciding what events to actually simulate. Furthermore, the exact causes for delays related to crew are not existent in publicly available data, nor are they present in the data from Mexicana. It is often only known *that* a flight is delayed rather than *why*. While IATA has defined a set of delay codes for various events they are often grouped presented in groups, such as “Flight Operations and Crewing”.

When constructing the crew event generator, two events were focused upon. Firstly a crew member could be delayed for an activity, second they could be absent. These two very simple events manage to capture most of the delays that originate from crew. However, it is not the events themselves, but their causes that are of interest. These causes may vary from situation to situation and must therefore be easily modeled by the user of the simulation framework. For delays the crew event generator should notify other parts of the system when the delays occur, but for absence there must also be the option of notifying the system in advance of the activity. This can for instance be used to model events such as calling in sick in the morning.

As for the actual simulation technique it must support a constant change of data as well as parameters from external sources. This is because the flight data is known to change over time as delays and recovery is performed, and the parameters that govern the simulation may also change over time, for instance to simulate what happens if an airport terminal is being rebuilt, or if influenza is spreading.

The simulation also obviously needs to support different parameters for different locations in the world. It is not reasonable to assume that the conditions for, say business of airport, are the same at all airports. Even so, the daily activities of a crew member must be able to be modeled in a way that does not require a specific model for each location.

At first, a model that could capture all these features was developed, that would also allow users of the framework to describe how crew should be simulated on a general level. After deciding upon such a model the transitions within the model were determined. The workday of a crew member can be broken down into three major parts:

1. Get to work and check in
2. Perform activities such as flights or ground duties
3. Get home, or to a hotel, and rest.

While this model is very rough it captures many of the aspects already. The three points define states which the crew member must reach within a given time frame, and if this fails then delays are generated. Furthermore, crew may fail to reach a state entirely and thus not showing up for work or having to go home during a workday due to sickness or other reasons.

This simple model is far from usable, but it highlights a few important features. Activities of crew can be modeled as simple states, where transitions between states describe actions, such as getting to work. The time required to perform such a transition may be different for different locations, but the overall model remains unchanged.

4.4.1 Crew simulation model

In order to successfully simulate a system as the one described in 4.4 two simulation techniques were focused upon. The first one is a discrete event simulation where the simulation is defined as a set of states with connections in between them. Crew members move between the different states according to given rules and generate events as they do so. The simulation is known as a discrete event simulation because once an event is handled the simulation's internal clock jumps forward to that of the next event.

The other model that can be utilized is a queue simulation. This model is similar to the discrete event simulation but states are defined as queues with capacities and crew members move between the states by adding themselves to the end of the queue that takes them to their next state. The simulation then removes crew members at the head of the different queues according to the capacities and this is the way that crew members move in the model.

When comparing these two there are a few differences that can be seen. The first thing that can be noted is that the queue simulation is by design built to simulate congestion as many crew members in one queue will create a bottleneck. This could be useful when simulating things such as check in or passport control where queues naturally occur. On the other hand tuning the performance of such queues can be very hard. While the distribution of times in order to pass such a queue can rather easily be measured it will take some work to convert that into a flow and a capacity. Furthermore, not only people present in the current simulation will affect queues and this requires either adding dummy members to the model, a process known as *padding*, or adjusting the capacities.

A second point worth noting is that the queue simulation needs to update the state of all the queues in every simulation step. The number of queues could potentially be large as it grows not only with different airports but also with different connections between flights etc. The discrete simulation on the other hand needs only generate when each crew member arrives to its next state, and then that crew member may be safely ignored until that time.

For these reasons a discrete event simulation was chosen as base for this simulation. It may very well be that a well tuned queue simulation would give a more exact simulation, especially for situations where much congestion is expected in common areas such as customs or passport control.

4.4.2 Customization

As the crew event generator may be used to simulate several different situations for several different airlines it is of great importance that it must be customizable. A user must be able to completely redefine the parameters for the simulation without having to change and re-compile the code. The simulation has been built with this purpose in mind and consists of several parts that the user is in full control of. The parts of the model are as follows:

- **States** define positions in the model in which a crew member has completed some task and is ready to perform the next. Examples include "Checked-In" which indicates that a crew member has arrived to work and "Flight" which marks a crew member as being ready to depart on a flight.
- **Connections** define how crew members may move from one state to another. Movement between states usually takes time and exactly how much can be sampled from various random distributions. What distribution that is sampled is also customizable and may vary with airport as well as crew type. For instance the distribution that controls how long it takes to move from one aircraft to another may vary greatly from large airports to smaller airports. Another example may be that pilots changing aircraft sample from another set of distributions with a higher mean values, to simulate the fact that they must also spend a few minutes filling out paperwork between flights.
- **Failure connections** define connections to secondary states that the crew member may find themselves in if the original connection fails. For instance, a crew member that is scheduled to start their day, modeled as a connection between states "Off-Duty" and "Day Start" may not do so if they have become ill. In this case they will have moved to a new state, "Sick", instead of the desired "Day Start".
- **Distributions** are common random distributions that the user may use to specify the times taken to move from one state to another.
- **Tests** are small Boolean functions that the user can create that describe when a connection fails to a failure connection. These tests may be of a random nature to indicate, for instance, the probability of getting sick.
- **Events** are the heart of the simulation and are generated in states. Not all states need to generate events, however. Events come in three flavors:
    - **Instant events** are events that occur when the crew member arrives at a state, such as "checked in".
    - **Delay events** occur when a crew member is late for something, such as a connection flight. The event starts when the crew member should have arrived at the new state and ends when they actually arrived
    - **Full day events** are events that take up the remainder of the day, such as sickness.

The parts described above can be used by an end user by creating an XML-file which defines the model to be used in the simulation. This file is parsed by CEG and then used as the simulation model. The model may also be changed during simulation to accommodate for changes in the world. This can be used to model temporary things such as influenza, which would affect a sickness test, or construction work at the airport, which would perhaps have an impact on some of the distributions.

When defining the model the user can also bind certain states to activities in the flight schedule. This indicates that all crew members scheduled for such an activity must have reached the given state before it can start. If some crew member would not make this connection the crew model must be designed to generate delay events for that state, otherwise the schedule will be executed as normal.

4.4.3 Operation

While ISC runs CEG in small time steps CEG operates slightly differently internally. Here CEG generates events for all crew members up to their next scheduled activity. Events include both switching from one state to another as well as the events described above.

The events that happen within the given timeframe are published by CEG. This means that the events are handed over to a part of CEG known as Message Dispatchers that create messages to be sent to other parts of the system to implement the events. The dispatchers are separated from the generation for two reasons. Firstly it makes it easier to adapt the CEG to work with new systems, and secondly because most systems have different handling of events that regard crew, such as check-in, and aircraft, such as delaying a flight.

The pre-generation of events allows for very fast run times, as many of the events are pre-generated. However, since several other parts of the simulation update the schedule it is not uncommon that CEG needs to re-generate the events for a crew member. It can happen that a crew member has already begun moving through the model when his/her events are re-generated. This may lead to crew checking-in early if their flight was delayed etc. This is the desired behavior as it is what happens in the real world.

The process can be visualized in the following steps:

1. Refresh Data. CEG checks the next upcoming action in the database and compares this to the same action last time CEG was run. If these differ, either because the schedule is changed or because the action is done, CEG removes all future events for this crewmember.

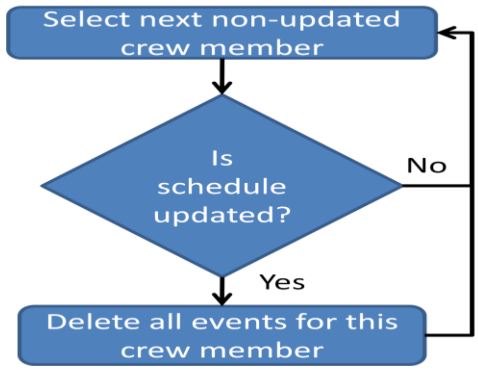

Figure 8. Step 1 of CEG: Refresh data

2. Generate Events: For all crewmembers that have no events, either because they were deleted in Step 1 or because they have not previously had any scheduled actions, CEG generates events up to the next scheduled action.

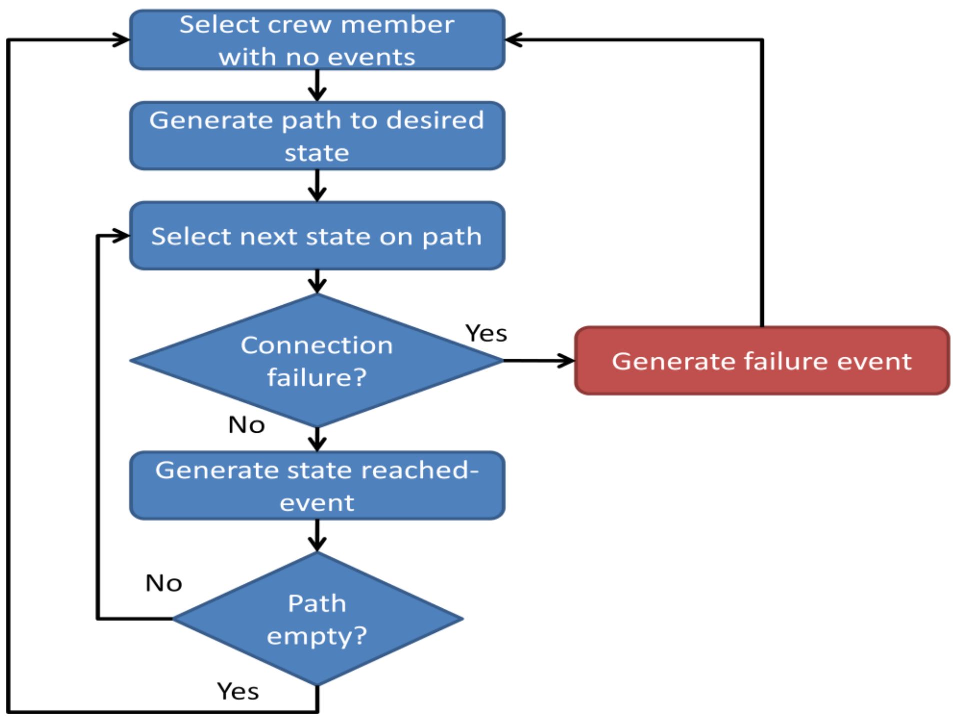

Figure 9. Step 2 of CEG: Generate Events

3. Update and Dispatch: In the third and final step CEG inspects the events for each crew member, sorted in a chronological order with the earliest event first. If the event occurs within the given time frame, the local model is updated and the event is sent to the dispatchers. This mechanism is what makes crew members arrive early or late to work if their flight is changed close to departure. Events generated for the first departure time would then already have affected the model when the new time is published.

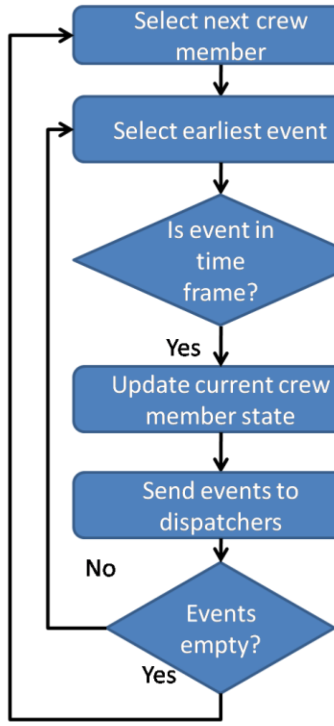

Figure 10. Step 3 of CEG: Update and Dispatch

Once all events are generated and the appropriate events have been dispatched, CEG reports back to ISC and waits for the next invocation.

4.4.4 Mexicana event model

For the purposes of this thesis a model for the crew event generator was built to simulate crew for Mexicana. The created model is a simplification of a crew member's workday, and contains only major states, such as Flight and Ground Duty that can be used to represent activities for all types of crew. The model models crew arriving to work and their transition between different states during their workday. The model also includes failure to show up to work, both for common diseases as well as for more wide-spread seasonal disease such as influenza. Influenza serves as an example of how this model can be used to generate special events and was included to demonstrate such behavior.

A sketch of the model can be seen in Figure 11. The state "Flight" covers all flights that are flown by Mexicana, and crew members will find themselves in this state both when working a flight and when deadheading. The state "OAG" on the other hand represents all flights flown by other airlines, which may be utilized for deadheading. These are fundamentally different from own flights since they cannot be controlled by the own airline and that crew have to check in earlier for these flight. The name OAG is an abbreviation for "Official Airline Guide", a source of timetables and other data for airlines and has become a common word to describe flights by other airlines(OAG Navigation, 2012).

Delay events are created whenever a crew member is late for Flight or Ground Duty. Check-In events are generated whenever a crew member checks in, regardless of whether this is late or not. The states "Leave Home" and "Traffic" define how much time a crew member has between leaving their home and their first activity (Flight, Ground Duty or OAG) and thus affects how early or late crew members check in for work. Finally the state "Well" exists as an opposite to calling in sick. This state makes sure that the test for sickness only is tested once per duty period.

Values for how long it takes to perform a state change and the ratios of failure due to sickness where obtained from the available data from Mexicana. This data was gathered for two distinct crew types: Cockpit (Pilots and First Officers) and Cabin (Flight attendants). The values for these were expected to differ as cockpit-crew was assumed to be less prone to calling in sick. The reasons for this are believed to be many but could include the fact that cockpit crew are better treated by the airline since their job is more mission-critical compared to that of cabin crew. The assumption that these times differ was confirmed by the available data, however the reasons for the difference remains a speculation.

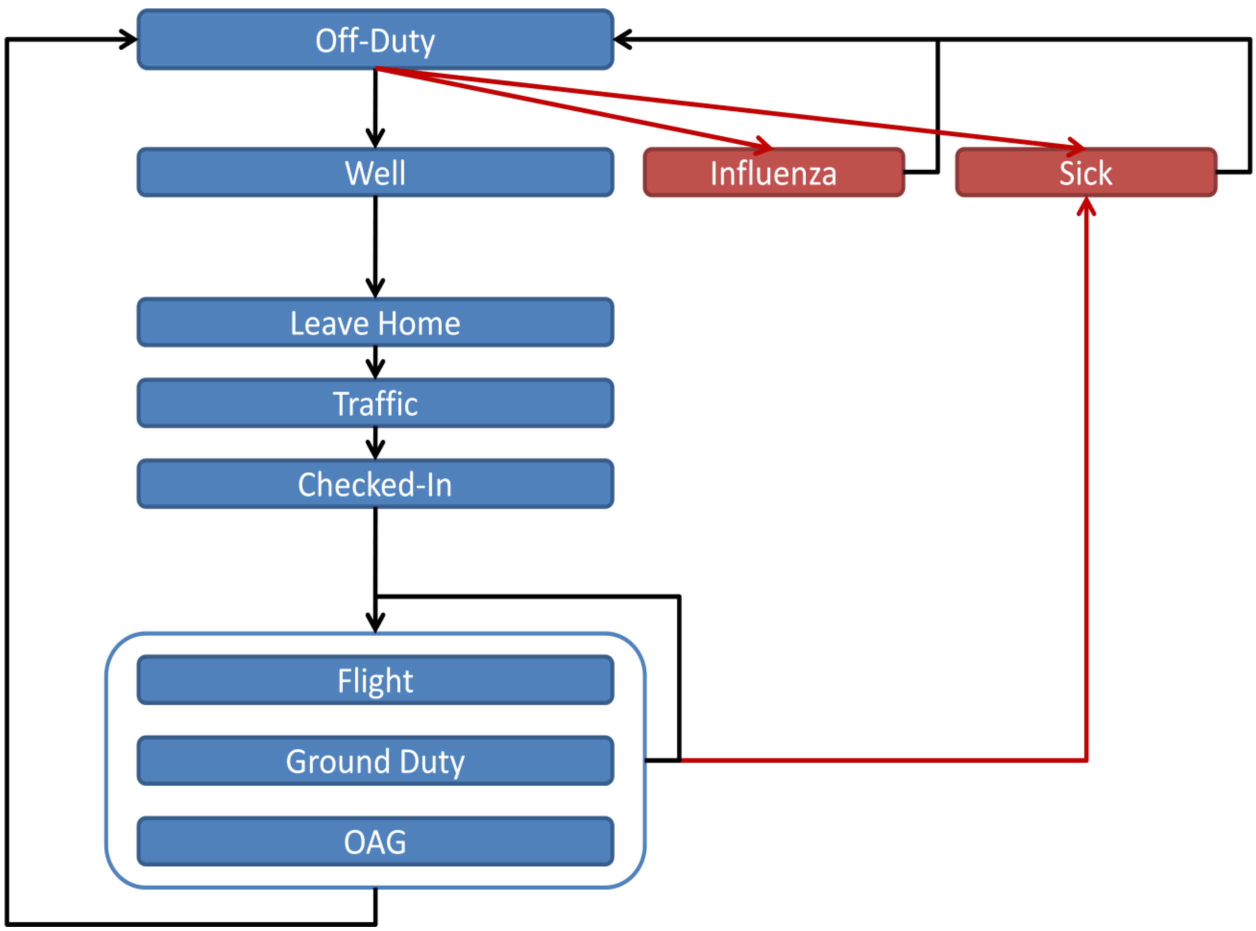

Figure 11. The crew model used for simulation. States Flight, Ground Duty and OAG represent the actual workday of a crewmember and several of these may be passed in one day. States in red indicate failure-states

5. Results

This thesis shows how a framework for simulation of the operations of an airline can be created by combining existing tools for simulation and recovery with a simplified model for a crew workday. The result is a framework that can be used to simulate the effects of various events that may cause the airline to operate under both regular and irregular conditions. By including a recovery tool used by airlines the effects of solving disruptions can be studied and key figures related to costs and performance can be obtained.

This chapter shows comparisons between simulated and real data to demonstrate how well the framework can represent the real world. A theoretical example of poor weather conditions is also evaluated. Finally, non-functional requirements, such as the performance of the simulation, are studied.

5.1 Outcome

The simulation framework created by this thesis if fundamentally different from much of the related simulation work as it treats disruption management and airline schedule recovery as something that is naturally occurring rather than something that only happens in extreme cases. Using recovery like this several robustness factors such as propagated delay can be estimated in a realistic environment where recovery is performed for small as well as large disruptions. This compares favorably to the operations of airlines where recovery is carried out on a day to day basis.

The framework also allows for studies of variables that are related to both cost efficiency and robustness that depend on recovery. The percentage of stand-by crew used is such a variable. If standby-usage is high then the schedule is perhaps somewhat over-tuned and may be operating near its limits.

This framework also allows for repeated simulations over the same data that can be used to identify the likelihood of a schedule performing poorly, as well as identifying the events that lead up to the poor results. All steps of the simulation are stored as the simulation progresses so that they may be studied afterward. Repeated simulations using different strategies for recovery may also show how these strategies differ given different circumstances.

5.2 Comparison between simulated and real data

In order to validate the framework the results of a simulation were compared to actual data from Mexicana. Some of the general attributes of the simulation can be found in Table 2.

The data presented below is gathered from two runs, each of three days of simulated time, and compared to two sets of three days of real data. The simulation framework was run with only minor events and the real data used for comparison is taken from periods with the lowest average delay, as this is assumed to have been executed during similar conditions.

Parameter	Value
Aircraft	60
Flights/day	~200
Total aircraft ground activities/day	~10
Total number of crew in simulation	~2200
Total crew ground activities/day	Average: 90 (varies from 40-200)

Table 2. Parameters of simulation

The simulation was performed using time-steps of 15 minutes. This provides excellent granularity and is a rough estimation of how long it takes actual planners to react and deliver a solution to most problems.

5.2.1 Flight simulation

As a first verification of the data gate-to-gate times for all flown legs were gathered. This time includes taxi times, both in and out at respective airports, as well as the total flight time. One leg is identified by its unique flight number and may be flown several times during the selected period. It is important to stress that legs are always one-way and that the returning flight has another unique identifier. A comparison between the total gate-to-gate time for simulated and actual data can be found in Figure 12. Each dot on the line represents one unique flight number.

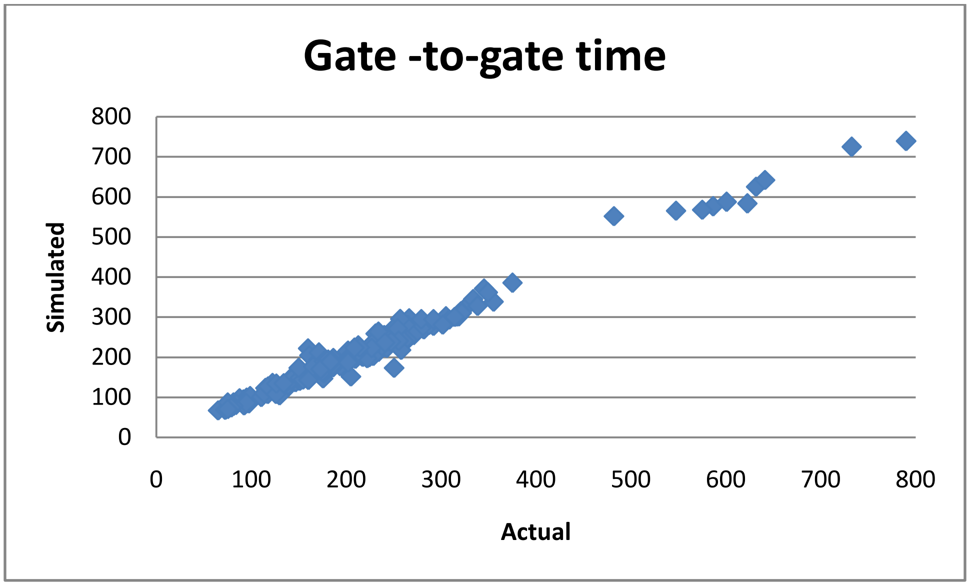

Figure 12. Actual gate-to-gate times (X-axis) compared to simulated gate-to-gate times (Y-axis)

The graph shows a strong correlation between simulated and actual gate-to-gate times with only a few minor variations. On closer inspection the simulated times actually fall a few minutes short of the actual flight times, something which, for instance, can be seen near the 300 minute mark in actual flight times, where most simulated flights are completed in around 290 minutes. This is a consequence of the fact that the simulation was performed under near-ideal circumstances with good weather and low traffic around airports. As a result of this the time for taxi-in and taxi-out are somewhat shorter in the simulated time compared to that of the actual recorded flight times. These times can be seen in Figure 13.

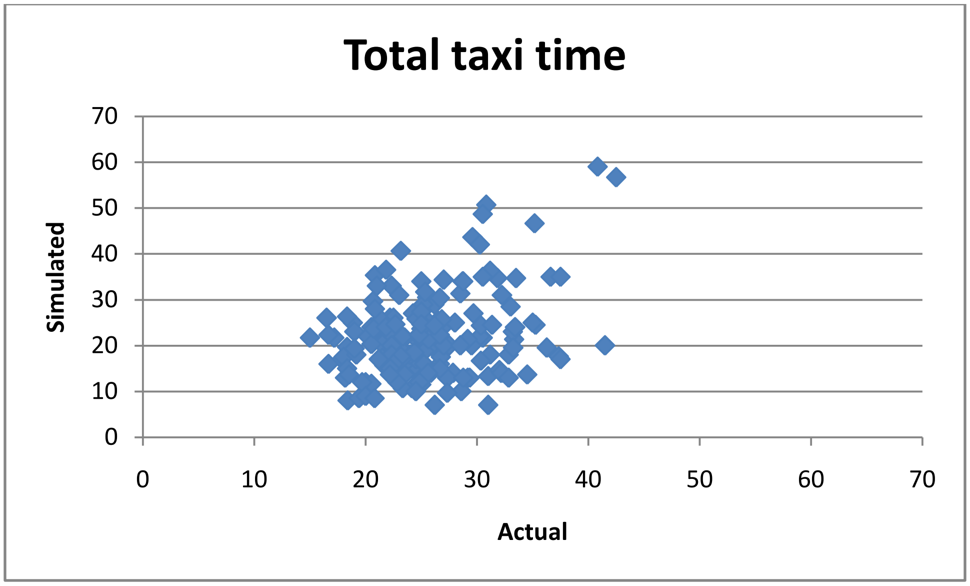

Figure 13. Actual taxi times (X-axis) compared to simulated taxi times (Y-axis)

Here it is evident that simulating during ideal circumstances in many cases underestimates the time required to perform taxi. Actual taxi times mostly fall within the 20-30 minute span while their simulated counterparts populate the 10-30 minute span. In this figure the randomness noise of both simulation and operations also becomes more visible. The simulated data is not always lower than the actual data, two points notably stick out and show taxi times as high as 60 minutes. However, the fact that the real times of these values are amongst the highest still makes these values feasible.

In conclusion it is obvious that even though MEANS has commonly been used to simulate behavior of entire airspaces or airports it is also a very good tool for simulating the aircraft of individual airlines.

5.2.2 Simulated events

Several events occur in the simulation that does not affect flight times. A comparison of such events can be found in Table 3.

Value	Actual occurrence	Simulated occurrence
Cockpit crew sickness	0,29%	0,22%
Cabin crew sickness	0,66%	0,64%
Cancelled flights	1,16%	1,08%

Table 3. Comparison of occurrence of events in actual data and simulated data

It may seem strange that both the actual data and the simulated data contain cancelled flights. As mentioned before the schedule was not perfect and this is what gives rise to the cancelled flights in the simulated data as some disruptions cannot be solved without cancelling flights. For the real data cancellations for similar reasons along with cancellations due to low booking together make up the slightly higher percentage.

5.2.3 On- time performance

As was shown in 5.2.1, Flight simulation, the simulated and actual flight times show a strong correlation and that the taxi times of the simulation are commonly slightly lower than those that are found in real data. In order not to underestimate the total gate-to-gate time, a second set of taxi times was generated where taxi times at all stations were increased by 10 minutes, both for taxi in and taxi out. Both are used below when comparing to the available historical data.

It is easy to believe that, due to the strong correlation in flight times, the delays of the simulation match those of the actual data, but this is not the case. In order to demonstrate this Figure 14 shows average departure delay per hour of day in local time for the airline. The data presented here is, as previously stated, average values for *all three days* of simulated/actual data.

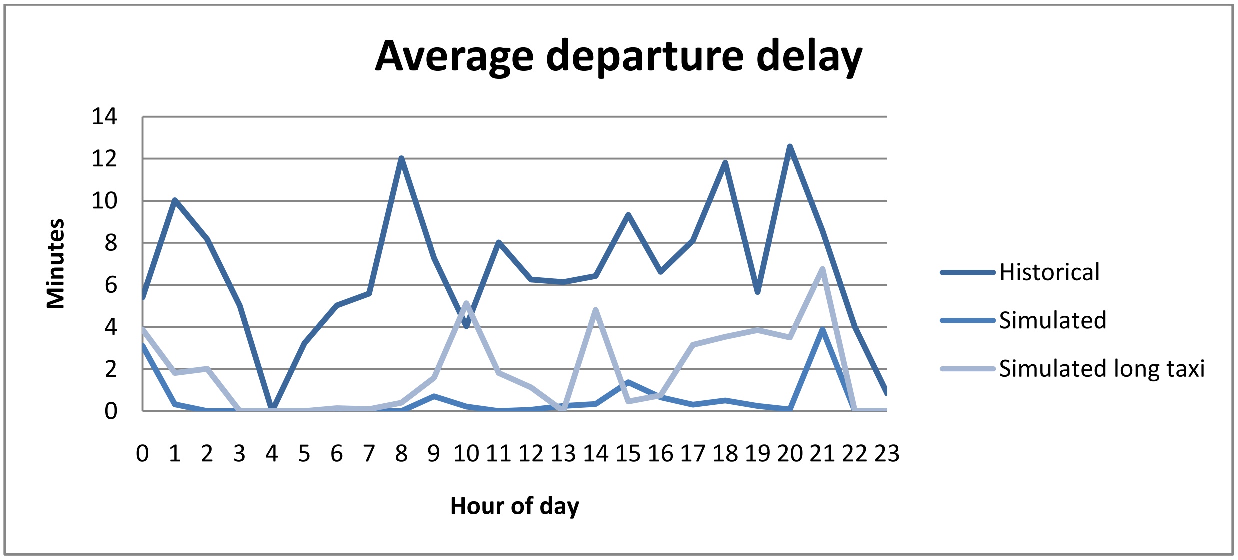

Figure 14. Average departure delay for all three data sets

Looking to the historical values the hypothesis of delays propagating though the day appears to be correct, and we can see times increasing as the day progresses. However, the simulated values do not follow the same pattern and most flights depart on time.

It also becomes evident that the simulations suffer less from departure delay compared to that of the historical data, even when the total flight time is greater and regardless of taxi time distribution. We can also see several spikes in the delay for all three data sets that occur in similar places. These are indications of periods of time where little can be done to mitigate delays. This may indicate times in the schedule where there are few possibilities for recovery, such as possible flight swaps etc. To better understand the departure delay, the average arrival delay for the same data sets is also presented in Figure 15.

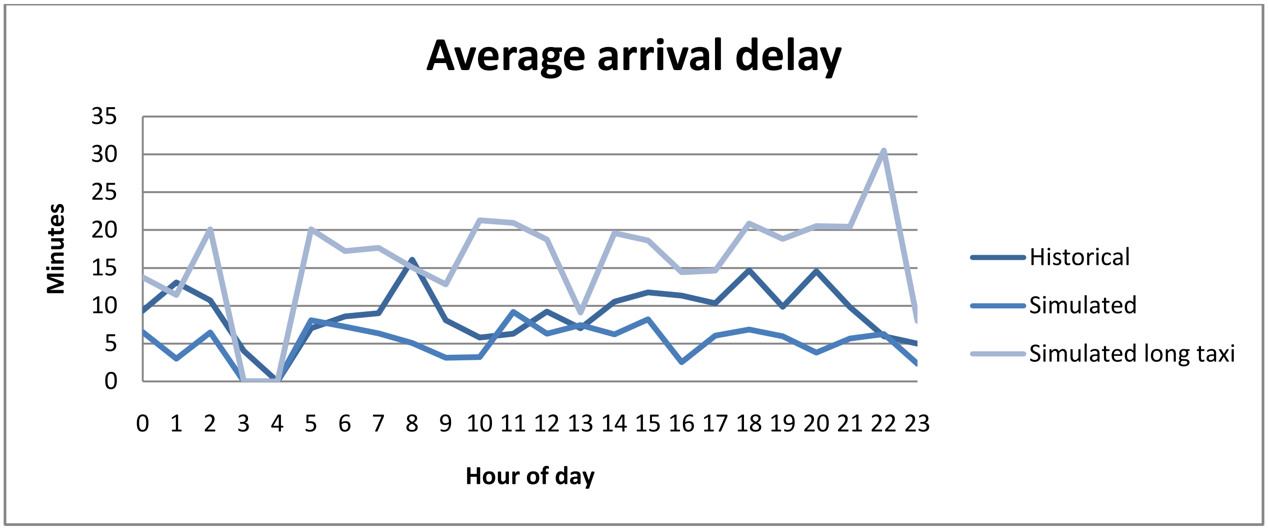

Figure 15. Average arrival delay for the three data sets

Here the effects of the two distributions for taxi times become obvious, with one simulation underestimating the total flight time and the other overestimating it. When combining the data in the figures 10 and 11 a trend starts to appear. Looking only to the historical data we can see shapes that appear in both graphs, for instance between 05:00 and 08:00. These shapes show that in periods of high arrival delay there is also a large departure delay, a sign that delay is being propagated.

For the simulated values these shapes do not reappear as frequently. One can however spot similar conditions in the simulation using longer taxi times around 02:00, 10:00 and 22:00 which are amongst the higher values for both departure and arrival delay. It is also worth noting that while 22:00 has the highest arrival delay this delay is not propagated, most likely because very few flights take off this late. It could however meant that earlier delay propagated to the current flight which then also suffered from additional delay, causing it to be delayed further on departure.

The images presented above represent the average delay, in minutes, of all flights. They do however not indicate how many of the flights that were delayed. Figure 16 shows the distribution of delayed flights, also presented per hour of day.

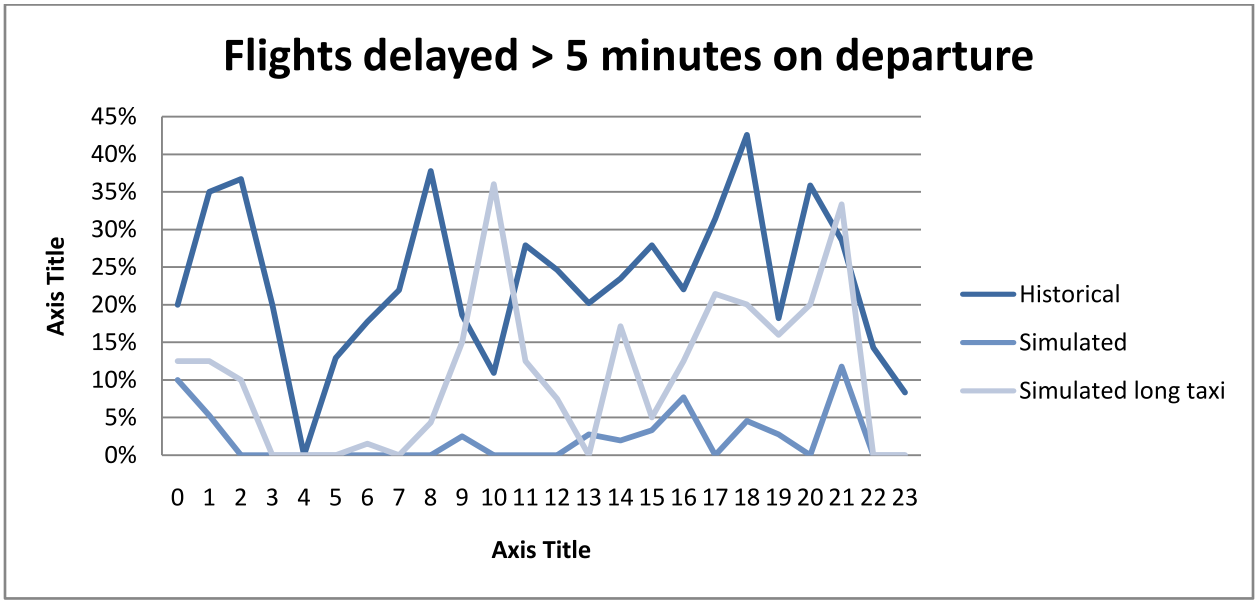

Figure 16. Percentage of flights delayed more than 5 minutes on departure

Here we see once again how delay is being prevented from propagating as several spikes are present in the simulated data but are seldom propagated to the following hour. We can also see around 10:00 how recovery can also create delay at points were delay was not previously present. In total, the numbers of flights delayed are fewer than in the historical data, but the overall shapes of the delays stay the same.

The reason for the overall decrease in propagated delay lies in the recovery performed during simulation. In every step of the simulation recovery is performed to prevent delay from propagating. During the course of the simulation this means that hundreds of aircraft swaps were done, sometimes to save just a few minutes. In reality this would probably be considered somewhat over-enthusiastic and many of these actions may not have been feasible in real time. During simulation all decisions are performed instantly whilst considering the entire schedule. In reality these types of choices are made by humans that cannot foresee all effects of their decisions, neither can they make an aircraft swap instantly.

Also one could imagine that there are several small events that are not simulated which could have delayed flights a few minutes on departure. Examples include luggage handling, passengers arriving late for their flights, slow boarding, and other airport related factors. Since such events are not simulated the simulated departure delay stays very close to zero when a flight is not affected by propagated delay. In reality this value is believed to be somewhat higher.

These reasons can aid in understanding why the delays in the simulated values are lower than the ones in the actual operational data. However, the data presented here also shows the positive impacts of good recovery even during regular operations, but it is important to keep in mind that the delay in all data sets is considered small. It is therefore not reasonable to conclude that propagated delay can be eliminated by recovery tools, as the costliest and longest sources of delay occur during major disruptions, something that is not captured during the course of these simulations.

What can be concluded is that the simulated values resemble those of the real data, and that the effects of recovery are clearly visible. This indicates that the framework performs as intended and simulates similar results given similar circumstances but with the additional dimension of recovery.

5.3 Simulation of poor conditions

Part 5.2 proves that the simulation framework performs well under good circumstances but it is also vital that the simulation can perform well under poor conditions. The purpose here is not to correlate simulated values to an actual scenario, but rather to see that the framework can operate under excessive stress. For this, a theoretical scenario was created that involved a tropical storm sweeping in over Mexico from the Pacific Ocean. The storm affects the capacity of several airports, but has the largest effects on MEX and CUN which are hit by the storm with several hours of strong winds and low visibility as a result. The simulated storm diminishes as it enters the Caribbean. This happens during the course of one day with the storm appearing early in the day, and disappearing at night, local time, and after this the simulation progresses under fair conditions.

The results of one such simulation is presented below to show how the simulation framework can be used to estimate key values for such a scenario. It is however recommended that several simulations are done, and compared to avoid any statistical errors, before the output can be relied on. The results presented here are therefore meant to be interpreted as a proof of concept, rather than a prediction of what would actually happen during such a scenario.

The reasons for the scenario are many. Firstly it is a scenario which will certainly force the airline to take hard decisions in order to be able to continue operations. Two of three hubs are directly hit which will thoroughly test the capacities of the recovery tools. This period is followed by a period of relative calm to ensure that after the external events have ceased the airline should return to normal operations.

Second, this is a situation where different recovery strategies can perform in very different ways. This thesis will not make any comparison between different such strategies but this simulation illustrates an example of a situation where the framework can be used to evaluate the actual optimization performed. The strategy used for this simulation focuses greatly on not propagating the delays to the following day, a quite common choice in practice. This means that flights are rather cancelled than delayed to the following day.

Finally, it is not an unlikely scenario for a Mexican airline to be in. Tropical storms appear along both coasts of Mexico and have large effects on the airline operations. Testing of performance during such extremes would be useful to many airlines in order to evaluate their current state of operations.

5.3.1 Simulation results

As can be expected, such a scenario cannot be carried out without cancelling and/or delaying flights. This scenario is an example of very stressful times for an airline, but the stress is mainly not due to resource shortage, but because of decreased operational capacities. This is seen almost instantly in the simulation where 15 of the 197 flights are cancelled within the first 24 hours. These hours are also subject to severe delays. The graph below shows the departure delay for the first 24 hours.

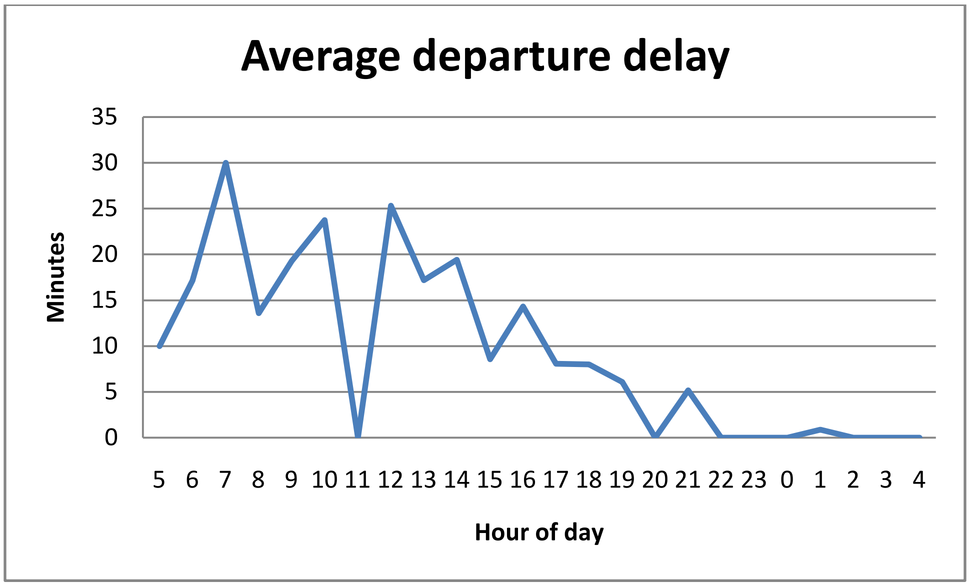

Figure 17. Average departure delay for the first 24 hours. Note! The simulation starts at 05:00!

Here we can see that the weather affects the airline almost instantly, even before the storm has first hit, which happens at 07:00. This is because the weather is worsening but also because airport control at both MEX and CUN estimates that the future demand for landing slots will be larger than their respective predicted capacities and issue slot restrictions in advance. This means that several flights are held at the ground before they are allocated a slot for landing or take-off at one of the affected airports. It's also worth noting that this is the average delay for all flights, even for flights not affected by the storm. Flights delayed at MEX and CUN were often delayed for one or two hours, before being assigned a slot.

One hour, 11:00, needs some additional explanation since it shows zero delay. Only five flights were scheduled to depart during this hour: One was cancelled, one was not flying between the affected airports, and the three remaining flights were all assigned slots which matched their departure. Therefore, rather amazingly, all flights during this hour did indeed depart on time, in spite of the bad conditions. Had there been more scheduled flights at this time, however, zero delay would not have been likely. This is a good example of why one should run several simulations, since even with the current setting it is unlikely that all flights depart on time.

The effects of poor weather on the simulation are easily seen with average delay times of up to 30 minutes at 07:00, which is a moment in time when CUN is enforcing slot restrictions and MEX have announced that they be doing the same at 09:00, keeping many planes on the ground. By 14:00 both MEX and CUN have sufficient landing slots but the effects of them still linger. In, fact this creates a larger flow to the third base, GDL, which also suffers slightly from poor weather, causing it to enforce a slot restriction between 14:00 and 19:00. The consequences are however small as the capacity of GDL is only slightly below the demand, and planes are therefore not kept waiting for as long. Finally, the next day, business continues as usual and the following two days are executed with an average departure delay of less than five minutes for all hours of the day.

The fact that this simulation did not create a lack of resources, robustness related values such as standby-usage do not indicate much stress. During this three-day simulation only around 8% of standby-times for captains were used, simply because there were already an excess of resources that had been made free due to the cancellation of previous flights. The standby crew that was used usually covered for sick coworkers, not due to the weather.

Another important factor in airline scheduling, and one that can rise greatly during irregular operations is the cost of layovers. In short, this means that the airline must pay for hotels, meals and transportation for their crew if they have to stay the night in away from base. Table 4 shows the total number of crew members that spend their night away from any of the three bases: MEX, CUN or GDL. Here the importance of good recovery is shown once again. Unlike what may be expected, the first night actually has less layovers than in the original schedule. While many flights were cancelled due to the poor weather, they were cancelled in a manner that kept the schedule consistent and crew at base. If one analyzes the cancellations it becomes obvious that many of the cancellations are round trips which produces this effect. The decreased number of layovers during day 1 and day 3 has their roots in the cancellations made, on the first day due to poor weather and also during the third day due to the problems in the schedule. The increased layovers on day 2 are a result of tuning to the schedule and a few early deadheads to cope with the cancellations on day 3.

Day	Original schedule	Simulated schedule
1	520	510
2	503	506
3	523	512

Table 4. Nights spent away from base

5.4 Simulation stability

During simulation there is always the chance of the simulation creating an infeasible state. This may be due to poor modeling, errors within the simulation or poor recovery techniques. Infeasible states are states that contain real errors, such as flights with no crew, or states which are not deemed to be realistic, such as schedules with a majority of all flights cancelled. Since the simulation framework is meant to be able to be run in an unattended way, it is important that such states are avoided as this would cause a need for supervision. In some cases, such as severe weather, there may be situations where there are no feasible solutions and mass-cancellation is the only way. However, during normal operations where conditions are fair it is assumed that such states should never be reached.

ISC is configured in such a way that it stops if direct errors in the simulation are discovered. Therefore, the simulation stability can be tested by allowing the simulation to run without supervision for a long time. The simulation framework is then expected to run to completion and all states along the way should be deemed feasible.

As work has progressed during this project several simulations of longer runs have been executed without problems, and simulations have been running for days of real time without encountering problems or infeasible states. This indicates that the framework is indeed very stable and can function unsupervised.

One might argue that this test is more a test of the performance of the recovery algorithms rather than the simulation. While this is to some extent true, it also shows that the simulation does not generate severe disruptions when they are not intended to occur.

5.5 Simulation performance

In order for a simulation to be useful it is imperative that a simulation can be performed in reasonable time. Exactly what is meant by this is strongly dependant on the situation at hand. During strategic planning, that happens several months before a schedule is executed, the run-time of a simulation is of less relevance compared to during the recovery phase. For this reason it is also hard to determine what is classified as “good”.

Furthermore, the time taken to simulate depends very much on the simulated environment. The time taken to simulate obviously varies with the number of flights and activities per day as well as the number of crew members involved. Also, the time varies with the amount of recovery that needs to be performed as a feasible solution might take some time to produce. The parameters for the simulation can be seen in Table 2. Parameters of simulation.

Using these settings described in Table 2 when simulating for fair conditions simulation of one “time step” takes an average of two minutes. During these two minutes more than 90% of the time is spent performing recovery and the actual generation of events and simulation is finished in a matter of seconds.

As previously stated, these simulations used a time step of 15 minutes and shorter steps are usually not necessary. Simulations using time steps up to 30 minutes also provide adequate results but show a tradeoff between realism and simulation speed. Using steps larger than this is not recommended as this sometimes causes recovery to be performed too late resulting in less optimized recovery or sometimes even errors in the simulation. Also it is not reasonable to believe that human planners would react that slowly.

Using 15 minute time steps simulation is roughly seven times faster than real time and one day of operations can be simulated in less than four hours. This time window is large enough to allow for recovery personnel to simulate the possible effects of bad weather, airport closure and other events that may occur in the reasonably foreseeable future. If a larger time step is used then the framework can also be used to simulate the effects of various recovery options for a couple of hours into the future. For instance a ten minute simulation could provide information on the upcoming 2.5 hours if using 30 minute time steps. While this may not provide the planner with a direct solution it will aid them in making a more informed decision.

For actual operations it is recommended that several simulations using the same parameters but with varying random seeds are performed in parallel across several machines in order to obtain a statistical level of security. Certain events, such as disease may have severe effects if certain crew members, usually captains, are affected but can be quite manageable otherwise. By running several simulations the user will also get an understanding for the estimated probabilities of the various outcomes.

6. Discussion

This thesis has presented a simulation framework which allows airlines to estimate several parameter related to schedule robustness and cost. These values provide indications of how the schedule is likely to perform during the given circumstances. Previously airlines could only calculate the cost of all assignments in the schedules and other factors such as robustness had to be derived from previous experience.

By looking to this the estimates produced by this framework, airlines can better select what schedule to use to avoid major costs caused by irregular operations. This can lead to more smooth operations along as the need for recovery diminishes with increase robustness. Smooth operations and increased robustness can also lead to less propagated delay, which can increase on-time performance and customer satisfaction.

6.1 Results

In chapter 5 results of the simulation framework were presented to showcase examples of variables that can be obtained from simulation. Using the simulation framework the inherent robustness of the schedule was tested and parameters such as gate-to-gate times observed. These parameters aligned with that of historically available data, which demonstrates that the framework is indeed able to capture the operations of an airline in a good way. The delay times were lower than that of the historical data, a result of the simulation being able to take instant decisions while considering the entire schedule, something that is not possible for human planners.

Returning to the figure 3 from chapter, sources of primary delay, it becomes evident that the simulation framework covers most of the areas that cause delays, with the exception of Security, Other Airport and Miscellaneous. However, large sources of delay, such as the ones originating within the airlines themselves, would still benefit from further investigation. For instance, aircraft events or gate-related events could be included to better model Airline and Other Airport delays.

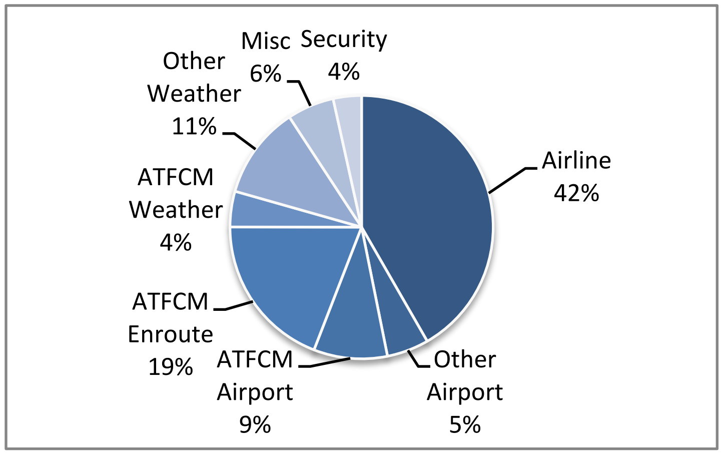

Figure 18. Sources of primary delay in Europe 2010 (EUROCONTROL, 2010)

A second scenario was also presented where the effects of severe disruptions showed much higher rates of cancellation, as well as much propagated delay. This shows that the framework is indeed able to capture the effects of severe disruptions with much delay propagation. It also shows that even with the aid of recovery tools propagated delay cannot be eliminated during certain times.

The framework provides access to variables such as estimates of delays, delay propagation and standby-usage which allows for better understanding of the strain placed on the schedule and where extra resources may be needed in order to further improve it. In summary, the simulation framework allows for more informed decisions when choosing which schedule to implement. It also offers access to better metrics for comparing two different schedules to each other before execution.

The variables presented in this thesis, for evaluation of performance and robustness, are believed to be only a few of many interesting values that can be observed to determine schedule performance. For instance, the value of standby-usage only serves as an estimate of the average strain on the schedule. It is possible that analyzing how this parameter changes over time may lead to the discovery of both situations where standby-crew are not needed, as well as times where there is a greater need. For this reason the framework supplies data of all stages of simulation so that new variables may be gathered from the data as a post processing feature. This is very useful as simulations that performed very differently given seemingly similar circumstances can be stored and analyzed as new metrics for comparison are studied.

Finally, the impact of changed circumstances can be determined before they happen, in order to make way for smooth transitions. For instance, consider an airline which is currently negotiating a new lower limit for the number of changes to the schedule allowed to be made for pilots. This may not seem like a major change of circumstances, but it does have an impact on how recovery can be performed. By simulating this while negotiations are still in progress the current schedule can be tested under the new circumstances. If changes are needed these can be found before the rule is enforced, and additional recovery situations can be avoided.

Another important feature is also that the airlines can now seek to minimize the estimated cost function. Previously a lower cost could mean lower cost after execution but could also be the absolute opposite as schedule robustness might be severely degraded if the original schedule is overly-optimized. Using the simulation framework the total expected cost of executing the schedule can be calculated and this can potentially be of great use as optimization of the schedule can, to some extent, be automated and poor schedules be disregarded at an early stages.

The fact that the framework can also help discard poor plans early on may be one of the most value adding features of the framework. This leaves planners with more time to choose between the good plans and actions, rather than finding out which the good plans are for themselves. This allows them to focus more on value adding tasks and hopefully implement better solutions.

6.2 Shortcomings

The framework proposes a step forward for airline simulation, but there are still aspects of airlines that are not covered by the framework. These would contribute to the end result of the framework and give more insightful results. A few of them are presented below:

6.2.1 Passenger model

The most notable shortcoming is perhaps the lack of a passenger model and passenger recovery. It is believed that the best solution for crew and aircraft is close to the solution that is best also for passengers since both models still strive to minimize delays and cancellations, although being able to consider passenger itineraries and compensation would undoubtedly change the results of this framework.

Such values can to a certain extent be estimated form the values of the simulation framework by comparing previous itineraries and percentage of tickets sold per flight to the results of the framework. However, this can at best be seen as a rough estimate and should not be relied on to a great extent.

Furthermore, and perhaps more importantly, just as a crew model was justified by the constraints it enforced on recovery the same can to some extent be translated to a passenger model. Passengers are not bound by rules in the same way as crew, unless airline policies state such “rules”, but they do affect the seats available for deadheading and the costs for cancellation. It is therefore possible that the simulation framework sometimes over-estimates variables such as the number of seats available for deadhead use and that those factors may have lead to a different result.

6.2.2 Airline operations

The simulation framework simulates advanced recovery as it is likely to be performed by the airline, but there are still many factors of airline operations and recovery which are not simulated. A notable example of this is that airlines can increase the cruise speeds of aircraft that have been delayed on departure in order to minimize the delay on arrival. This will cause the fuel burn during the flight to increase, as flights are normally flown at the most fuel efficient speed, but might be more cost efficient on the whole since costs due to delays are removed.

Such actions are not simulated by the simulation framework and apart from this they also have a slight effect on the observed flight times in the historical data. These types of decisions are taken by planners and could be added to ISC, but require a much more sophisticated decision-model than what is currently present. Furthermore, this would require an update of MEANS to allow dynamic switching to a data set containing faster flight times for these cases.

Apart from the above, airlines often perform recovery in a slightly different manner than what is done within this framework. Due to the complexity of the problem and because several planners work with the same schedule, recovery is often only performed on parts of a schedule. Examples of such parts can be individual fleets rather than for all aircraft at once, or day-by-day rather than over the foreseeable future. It may seem more clever to regard all aircraft as a bigger picture is considered but, recalling Smith´s (Smith, 2004) work using station purity, this may not be the case. This is therefore an area that could benefit greatly from further studies as new findings would have the potential to improve both simulation and airline operations.

These are but a few examples of situations where the decisions made by planners have been simplified. The topic of cancellations has also been mentioned previously and is still something that ISC performs rather naïvely. A decision model where flights can be prioritized based on revenue or passenger itineraries could, for instance, improve these decisions even further.

6.3 Performance and usage

The performance of the simulation framework has proven good enough for the types of simulations presented in this thesis but it is possible that higher speeds are required in the future. It is worth pointing out that it is not believed that entire months of running will need to be simulated. This is because airline schedules are often cyclical and because effects of poor scheduling or recovery become visible in a much shorter time span. Therefore, a simulation of one or two weeks of operations with good results can provide information on a much larger time span than just the simulated time.

As previously stated the simulation framework spends most of its time performing recovery and when this is performed is decided by the input to the framework. A better scheduling of recovery which could potentially improve the speed greatly rather than just performing recovery every time step, which is the current setting. An even better solution would include some modification of ISC to trigger recovery only when certain disruptions occur.

Finally, for disruption management ISC can be configured to simply propagate delay flights for all types of disruption instead of using recovery tools. In this setting simulation is *very* fast and shows immediate results of a recovery within minutes. It would not make sense to use this for any predictions longer than, say, an hour but may prove useful in stressful times.

6.4 Further applications

The focus during this thesis has been to create a framework that can be used when creating airline schedules and when performing recovery. However, during the development of this framework several other areas have seen benefits from such a framework. Because they are not the primary goal of this thesis they have been left out up until now, but deserves some mentioning as it shows more areas where a framework such as the one presented in this thesis can be of use.

6.4.1 Planner training

Using the Jeppesen Disruption Management as a data center was a choice that was made to get easy access to rule sets and optimization features. It does however come with another bonus feature, namely that the simulation framework interacts with the system as external systems and planners normally would. This means that simulation progress can be observed and visualized though the tools that planners use in their everyday work.

By removing the automated recovery step and instead halting the simulation framework after each time step a human planner is allowed to step in and perform recovery instead. This setup can be used to train future planners and let them get familiar with both the tools as well as the knock-on effects of choices that they make without actually jeopardizing airline operations. It also allows training of planners for severe situations so that they are better prepared when such events occur.

Today planners are trained using theoretical exercises and by watching experienced planners perform recovery. Using the simulation framework for this would add a level of hands-on experience with the tools and might even shorten the training process as training for severe events can be performed on-demand rather than when they occur in the real world.

6.4.2 Strategic decisions

The simulation framework is not only a tool to study schedule performance, but it can also be used to evaluate recovery procedures. Recovery, even with the aid of powerful tools, requires careful consideration of many factors such as whether delays should be preferred to cancellations and when these rules should followed. Such strategies are likely to behave very differently considering the current weather, or airspace situation. Using the simulation framework airlines can test several different strategies and compare them to each other without jeopardizing actual operations.

6.4.3 Optimization

As discussed previously in 5.4, Simulation stability, a stable simulation is not only an indicator that the simulation tools is performing well but also that the recovery tools are doing the same. This has also made the simulation tool useful when testing new versions of the optimizer or new rule sets for crew. It could also potentially be used to perform more extensive tests of rule sets before they are brought into use, making the transition from one system to another smoother.

6.5 Summary

In short, the simulation framework provides a good base for simulations where several different situations can be tested. The framework can be used to evaluate different techniques and strategies for scheduling and recovery under various conditions as well as evaluation of the current performance. The results match those of historically available data and severe disruptions can be evaluated without the introduction of unexpected side effects. The framework also provides both summaries of important parameters as well as data on all actions performed in the thesis. This allows for monitoring of key values between simulations as well as in-depth studies when these are needed. Finally, the framework can be of use in other areas, not within the focus of this thesis such as planner training.

7. Future work

The airline industry is a very interesting area with very many topics left to explore. Several of these areas have direct relation to this thesis and could make use of the work from this thesis. The ones most appealing to the author have been presented below, although it is easy to imagine several other areas that are worth investigating.

7.1 Passenger model

One of the most striking features of this thesis, and much the related work is, as previously mentioned, that passengers are left out. The reasons for this are, as previously explained, that aircraft and crew recovery is usually more costly and that it is hard to perform passenger recovery unless aircraft and crew recovery is performed first: only then will the airline know what flights will be cancelled. However it is not hard to imagine that the passenger costs, especially during more severe disruptions, may contribute greatly to the cost function. This means that a slightly different approach that is more passenger-friendly might be the better choice, but would be ignored by this framework.

Even if the inclusion of passengers in the model is not believed to cause any major changes to the schedule chosen the cost of passenger recovery is something that airlines will want to know. Also, the sooner these effects are known, the sooner the airline can start to make plans for possible hotel bookings etc. By making such decisions early on airlines could possibly save on the recovery costs as well.

7.2 Airline modeling

As previously mentioned the decisions taken by airlines have been simplified for the purposes of this thesis. Such decisions can however make a significant difference to the end result. By including better decision support the simulation framework could perform more clever cancellations and perhaps also save on delays when choosing to fly certain flights above normal cruise speed.

Also the simulation framework could be modified to better reflect how planners act in the real world, as described in 6.2.2. This would undoubtedly increase the realism of the framework and also enable the framework to be used to evaluate how planners working routines could be improved.

Another interesting aspect is studying the consequences of the current state of airline operations. As previously seen the simulation framework suffered less from propagated delay compared to historical data. This indicates that there may be room for improvement in airline operations. By creating several simulations using slightly different decision models then new strategies and approaches can be evaluated.

7.3 Parameters for robust scheduling

As previously mentioned the framework presented in this thesis can evaluate schedule robustness by, for instance, estimating the overall cost of executing the schedule. While this is indeed very useful and can be used to separate robust schedules from less robust schedules there is still a great opportunity for finding the key parameters that create the robustness. An in-depth study of why certain schedules fail when similar schedules do not may provide very interesting answers as to exactly which techniques contribute the most towards robust scheduling. Obtaining such parameters would provide some much needed scientific grounds for major decisions that are today made only using past experience.

7.4 Usability

While there is a lot of input that needs to be fed to the simulation framework in order for it to function properly, much of this input, such as flight and taxi times, is expected to remain static between simulations. Variables such as weather conditions and simulation parameters are on the other hand something that needs to be configured from one run to another. This thesis has not focused on making such changes easy to configure for a new user. Instead, these settings are defined in files that are parsed by the simulation framework. The same files are also monitored for changes during the simulation to allow for changes to conditions during simulation.

Allowing for access to these settings in an easier and more intuitive fashion than opening them and editing them manually would allow for less experienced users to also make use of the simulation framework. Allowing for planners to access settings and run simulations from the same interface that they use for planning would also make simulation an integrated part of their work. Such work unfortunately was not prioritized during this thesis but would surely contribute a lot to the usefulness of the framework.

8. Works Cited

Air Traffic Association of America Inc., 2010. *Economic Report,* Washington DC: Air Traffic Association of America Inc..

Andersson, E., Housos, E., Niklas, K. & Wedelin, D., 1998. Crew Pairing Optimization. In: G. Yu, ed. *Operations Research in the Airline Industry.* Boston: Kluwer Academic Publishers, pp. 228-256.

Barnhart, C., 1994. *Branch-and-Price: Column Generation for Solving Huge Integer Problems,* Atlanta: Georgia Institute of Technology.

Belobaba, P., Odoni, A. & Barnhart, C., 2009. *The global airline industry.* West Sussex: John Wiley & Sons.

Bloomberg, 2010. *Mexicana de Aviacion Files for Bankruptcy in Mexico.* [Online]
Available at: http://www.bloomberg.com/news/2010-08-03/mexicana-files-for-bankruptcy-protection-says-flights-won-t-be-affected.html
[Accessed 02 05 2012].

Clarke, J.-P., Melconian, T., Bly, E. & Rabbani, F., 2007. MEANS MIT Extensible Air Network Simulation. *Simulation,* 83(5), pp. 385-400.

Cook, A., Tanner, G. & Anderson, S., 2004. *Evaluating the true cost to airlines of one minute of airborne delay,* Brussels: EUROCONTROL.

EUROCONTROL, 2010. *Coda Digest, Delays to Air Transport in Europe,* Brussels: EUROCONTROL.

Gatersleben, M. R. & van der Weij, S. W., 1999. *Analysis and Simulation of Passenger Flows in an Airport Terminal.* Schipol, Winter Simulations Conference.

Goodman, M. D., 2009. *A Grasp-Knapsack Hybrid for a Nurse-Scheduling Problem.,* Cardiff: Cardiff University.

International Air Transport Association (IATA), 2012. *FACT SHEET: Industry statistics.* [Online]
Available at: http://www.iata.org/pressroom/facts_figures/fact_sheets/Documents/industry-stats-mar2012.pdf
[Accessed 29 03 2012].

Jeppesen Systems Ab, 2010. *Jeppesen Disruption Management.* [Online]
Available at:
http://ww1.jeppesen.com/documents/aviation/commercial/JeppesenDisruptionManagement2010

11.pdf
[Accessed 22 05 2012].

Kang, L. S., 2004. *Degradable Airline Scheduling: an Approach to Improve Operational Robustness and Differentiate Service Quality,* Cambridge: Massachusetts Institute of Technology.

Kohl, N. et al., 2004. *Airline Disruption Management - Perspectives, Experiences and Outlook,* Copenhagen: Techical University of Denmark.

Kornilakis, H. & Stamatopoulos, P., 2002. Crew pairing optimization with genetic algorithms. *I.P. Vlahavas and C.D. Spyropoulos (Eds.),* pp. 109-120.

Lan, S., 2003. *Planning for Robust Ariline Operations: Optimizing Aircraft Routings and Flight Departure Times to Achieve Minimum Passenger Disruptions,* Cambridge: Massacgusetts Institute of Technology.

Lapp, M., AhmadBeygi, S., Cohn, A. & Tsimhoni, O., 2008. *A Recursion-based Approach to Simulating Airline Schedule Robustness.* Ann Arbor, Proceedings of the 2008 Winter Simulation Conference.

Law, A. M., 1990. *Secrets of Successful Simulation Studies.* Atlanta, Institute of Industrial Engineers.

Melconian, T., 2001. *Effects of increased nonstop routings on Airline Cost and Profit,* Cambridge: Massachusetts Institute of Technology.

Michelle, D., Gary, F. & Cheng-Lung, W., 2012. Robust Airline Schedule Planning: Minimizing Propagated Delay in an Integrated Routing and Crewing Framework. *Transportation Science,* 1 May, pp. 204-216.

Norin, A., 2008. *Airport Logistics - Modeling and Optimizing the Turn-Around Proces.* Norrköping, LiU-Tryck.

OAG Navigation, 2012. *OAG Dynamic.* [Online]
Available at: http://www.oagaviation.com/Solutions/Aviation-Data/OAG-Schedules-Data/OAG-Dynamic
[Accessed 22 05 2012].

Rabbani, F. F. R., 2004. *Implementation of an Airline Recovery Model in an Event-Based Simulation,* Cambridge: Massachusetts Institute of Technology.

Robinson, S., 2004. *Simulation: The Practice of Model Development and Use.* 1st ed. West Sussex: John Wiley and Sons, Ltd.

Rosenberger, J. M. et al., 2000. *SIMAIR: A stochastic model of airline operations.* Atlanta, Georgia Institute of Technology.

Schaefer, A. J., 2006. *Improving airline operational performance through schedule perturbation,* Berlin: Springer Science + Business Media.

Schaefter, A. J., Johnson, E. L., Kleywegt, A. J. & Nemhauser, G. L., 2001. *Airline Crew Scheduling under Uncertainty,* Pittsburgh: University of Pittsburgh.

Schumer, C. E. & Maloney, C. B., 2008. *FLIGHT DELAYS COST PASSENGERS, AIRLINES AND THE U.S ECONOMY BILLIIONS,* Washington DC: Joint Economic Committee.

Simaiakis, I., 2009. *Analysis and Control of Airport Departure Processes to Mitigate Congestion Impacts,* Cambridge: Massachusetts Institute of Technology.

Smith, B. C., 2004. *Robust Airline Fleet Assignment,* Georgia: Georgia Institute of Technology.

Sohoni, M. G., 2006. *Operational airline reserve crew planning,* Berlin: Springer Science.

Stojković, M., Soumis, F. & Desrosiers, J., 1998. The Operational Airline Crew Scheduling Problem. *Transportation Science,* 1 August, pp. 232-245.

Whittaker, C. J., 2006. *Design and Implementation of a Sector-Based Airspace Model for the MIT Extensible Air Network Simulation,* Cambridge: Massachusetts Institute of Technology.

Yu, G., 2003. A New Era for Crew Recovery at Continental Airlines. *Interfaces*, 1 January, pp. 5-22.

Zeren, B. & Özkol, İ., 2012. An Improved Genetic Algorithm for Crew Pairing Optimization. *Journal of Intelligent Learning Systems and Applications,* 4(1), pp. 70-80.