Scheduling an Aircraft Repair Shop

Maliheh Aramon Bajestani and **J. Christopher Beck**
Department of Mechanical & Industrial Engineering
University of Toronto
Toronto, Ontario M5S 3G8, Canada
{maramon,jcb}@mie.utoronto.ca

Abstract

We address a scheduling problem in the context of military aircraft maintenance where the goal is to meet the aircraft requirements for a number of missions in the presence of breakdowns. The assignment of aircraft to a mission must consider the requirements for the mission, the probability of aircraft failure, and capacity of the repair shop that maintains the aircraft. Therefore, a solution both assigns aircraft to missions and schedules the repair shop to meet the assignments. We propose a dispatching heuristic algorithm; three complete approaches based on mixed integer programming, constraint programming, and logic-based Benders decomposition; and a hybrid heuristic-complete approach. Experiments demonstrate that the logic-based Benders variation combining mixed integer programming and constraint programming outperforms the other approaches, that the dispatching heuristic can feasibly schedule the repair shop in a very short time, and that using the dispatching solution as a bound marginally improves the complete approaches.

Introduction

For a many industries with expensive assets, repairing a failed asset is significantly more economical than replacing it. Furthermore, there is economic pressure to reduce the number of such expensive assets so as to not waste their capacity: companies often cannot afford an inventory of back-up parts. Under such conditions, high-quality schedules which integrate maintenance activities and "regular" activities can have a substantial effect on the performance of the overall system. For scheduling research, such problems raise the dual challenges of dealing with the uncertainty of breakdown and with simultaneously reasoning about maintenance and traditional scheduling metrics like makespan, through-put, and on-time performance.

Following the case study presented in Safaei et al. (Safaei, Banjevic, and Jardine 2010), we study the scheduling of a military aircraft maintenance facility. In the problem, there exists a flight program consisting of a number of missions called *waves*, each with a requirement for a specific number of aircraft of various types. At the beginning of the time horizon, an aircraft is either ready for a pre-flight check or is waiting to be repaired in the repair shop. The high-level goal is to assign aircraft to waves under the constraints that:

- Aircraft are subject to breakdowns, detected during systematic pre- and post-flight checks. We assume that once an aircraft fails, it leaves the system to be repaired during the next scheduling horizon.
- An aircraft can fly in multiple waves, provided the waves do not overlap and the aircraft does not breakdown.
- Aircraft in the repair shop must be repaired before being sent on a wave.
- The repair shop has limited capacity.

A solution determines an assignment of aircraft to waves and a schedule of repair jobs which maximizes the wave coverage, considering the known failure probabilities, the characteristics of the repair activities such as processing times and resource requirements, and the capacity of the repair shop. Wave coverage is the extent to which the aircraft requirements of the waves are met. We assume that the schedule is computed off-line at the beginning of the scheduling horizon. It is then executed as is, with no dynamic reaction to actual aircraft failures. We study dynamic re-scheduling in future work (Aramon Bajestani and Beck 2011). The flow of an aircraft during a single scheduling horizon is illustrated in Figure 1.

We explore several techniques to solve this problem: mixed integer programming (MIP); logic-based Benders decomposition (LBBD) using either MIP or constraint programming (CP); a dispatching heuristic motivated by the Apparent Tardiness Cost (ATC) dispatching rule; and two hybrid approaches which integrate the dispatching heuristic with complete approaches (MIP and LBBD).

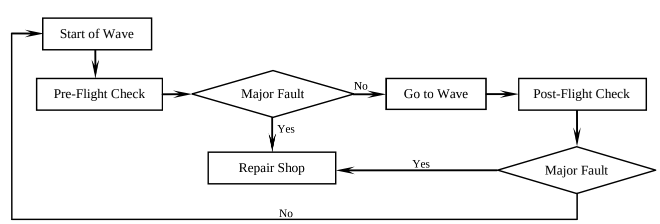

Figure 1: A Flow Chart Representing the Aircraft Flow among Waves, Checks, and the Repair Shop.

Our empirical studies indicate that the LBBD variation that combines MIP and CP outperforms the other complete approaches. The dispatching heuristic can provide the repair shop with a feasible schedule in a very short time, however the quality of the solution is likely to be insufficient. Finally, the integration of the dispatching heuristic and complete approaches results in a small reduction in the mean time to optimally schedule the repair shop.

The contributions of this paper are:

- Developing a logic-based Benders decomposition for scheduling an aircraft repair shop.
- Designing a dispatching heuristic algorithm guaranteeing a feasible schedule for the repair shop in a short time.
- Estimating aircraft availability and expected wave coverage by incorporating stochastic failure information.

The following section defines logic-based Benders decomposition and our problem. We then describe the proposed solution approaches and present our experiments and results. We end with related work on maintenance scheduling and conclude, discussing directions for the future work.

Background

In this section, logic-based Benders decomposition is introduced and the formal definition of the problem is given.

Logic-based Benders Decomposition

Logic-based Benders decomposition (LBBD) was introduced in the context of circuit verification and then generalized integer programming (Hooker and Yan 1995; Hooker and Ottosson 2003). The LBBD approach partitions the problem into a master problem (MP) and a set of subproblems (SPs). The former is a projection of the global model to a subset of decision variables, denoted y, and the constraints and objective function components that involve only y. The rest of the decision variables, x, define the subproblems. Solving a problem by Benders decomposition involves iteratively solving the MP to optimality and using the solution to fix the y variables, generating the sub-problems. The inference duals (Hooker 2005) of the SPs are solved to find the tightest bound on the global cost function that can be derived from the original constraints and the current MP solution. If this bound is greater than or equal to the current MP solution (assuming a maximization problem), the MP solution and the SP solutions constitute a globally optimal solution. Otherwise, a constraint, called a "Benders cut" is added to the MP to express the violated bound and another iteration is performed.

To our knowledge, this is the first work using logic-based Benders decomposition for maintenance scheduling.

Problem Definition

The problem is to assign aircraft to waves to maximize wave coverage while at the same time creating a feasible maintenance schedule. The problem definition is due to Safaei et al. (Safaei, Banjevic, and Jardine 2010).

Figure 2 is a snapshot of the problem at time 0, where circles represent aircraft. Three waves and their corresponding

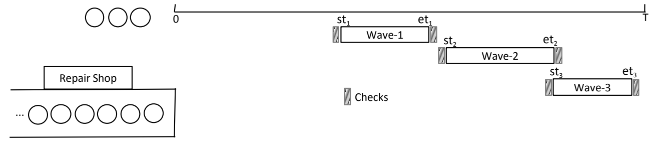

Figure 2: Snapshot of the Problem at Time 0.

pre- and post-flight checks are already scheduled. A number of aircraft (three in the diagram) are ready for the pre-flight check while others are currently in the repair shop awaiting maintenance before they can proceed to a pre-flight check. Failure is only detected during a check and we assume that a check will always correctly assess the status of an aircraft.

We adopt the following notation to represent the problem.

- N is the set of aircraft.
- K is the set of aircraft types. For each aircraft type $k \in K$, there are A_k aircraft ready at the beginning of the time horizon and λ_k is the mean failure rate.
- W is the set of waves and D is an ordered set of due dates consisting of the wave start-times plus a big value, B sorted in ascending order. Each wave, $w \in W$ has a start-time, $st_w \in D$, and an end-time et_w. Each wave requires a_{kw} aircraft of type k.
- R is the set of repair resources (called "trades"). The maximum capacity of trade $r \in R$ is C_r.
- J is the set of jobs (i.e., aircraft to be repaired). Each job is associated with a specific aircraft and I_k denotes the set of repair jobs for aircraft type k. M_r is the set of jobs requiring trade r. Each job requires one or more trades. The processing time and required capacity of job j on trade r is p_{jr} and c_{jr}, respectively.
- T is the scheduling horizon (i.e., the maximum time).

To find the probability of failure of an aircraft in pre- and post-flight checks for a specific wave, we need to track the complete history of the aircraft. For example, assume that a given aircraft is repaired and assigned to the first wave. There are three possibilities: the aircraft fails the pre-flight check; the aircraft passes the pre-flight check, flies the wave, and fails the post-flight check; or the aircraft passes the pre-flight check, flies the wave, and passes the post-flight check. Therefore, the availability of the aircraft for the second wave can be represented as a random variable whose expected value depends on the probability of these three different possibilities. Similarly the availability of the aircraft for subsequent waves depends on its entire path through the checks and waves. As the number of waves and aircraft increase, the size of the state space will become prohibitive. Furthermore, the scheduling decisions themselves impact the aircraft histories: the probability that an aircraft is available for the third wave is different depending on if it was repaired in time for the first wave or only for the second wave.

To break this loop and avoid the growth of the state space, we distinguish aircraft based on their type and use a recursive equation (Eqn (4)) to estimate aircraft availability and expected wave coverage. The failure rate, λ_k, is used to calculate the probability of failure associated with each aircraft type during pre- and post-flight checks, respectively: $\xi_k^{pre} = (1 - e^{-\alpha\lambda_k})$ and $\xi_k^{post} = (1 - e^{-\beta\lambda_k})$, where $\alpha < \beta$ to reflect deterioration of the aircraft through use. This formulation is different from that of Safaei et al. who used a constant probability of failure for all aircraft type both in pre- and post-flight checks.

Solution Approaches

In this section, the details of different solution techniques are provided. In our preliminary experimentation, a pure constraint programming model performed very poorly and so we did not pursue it.

Mixed Integer Programming

We propose a novel mixed integer programming model based on the common *time-indexed* formulation (Queyranne and Schulz 1994). This model is different from and significantly faster than that of Safaei et al. The variables are defined in Table 1 and the model is shown in Figure 3. We refer to this model below as *MIP*.

Var.	Definition
Z_{kw}	The number of aircraft of type k assigned to fly in wave w
x_{ij}	$x_{ij} = 1$ if the ith due date is assigned to job j
st_{jr}	The start-time of job j on trade r
U_{kw}	The number of aircraft of type k whose repair due date is st_w
E_{kw}	The expected number of available aircraft of type k for wave w
et_{jr}	The end-time of job j on trade r

Table 1: The decision variables (top) and inferred variables (bottom) for the MIP model.

$$\max. \sum_{w=1}^{W} \sum_{k=1}^{K} Z_{kw} \quad (1)$$

$$\text{s.t. } U_{kw} = \sum_{j \in I_k} x_{ij}, \text{ if } d_i = st_w \quad (2)$$

$$E_{k1} = (A_k + U_{k1})(1 - \xi_k^{pre}), \ \forall k \quad (3)$$

$$E_{kw} = (E_{k(w-1)} - Z_{k(w-1)} + U_{kw})(1 - \xi_k^{pre}) + \sum_{v=1}^{w-1} Z_{kv}(1 - \xi_k^{post})(1 - \xi_k^{pre}),$$
$$\text{if } st_{w-1} < et_v \le st_w, \forall w \ne 1, k \quad (4)$$

$$Z_{kw} \le \min(E_{kw}, a_{kw}), \ \forall k, w \quad (5)$$

$$\sum_{i=1}^{|D|} x_{ij} = 1, \ \forall j \quad (6)$$

$$st_{jr} + p_{jr} = et_{jr}, \forall j, r \quad (7)$$

$$et_{jr} \le \sum_{i=1}^{|D|} x_{ij} d_i, \ \forall j, r \quad (8)$$

$$\sum_{j \in M_r} c_{jr} \le C_r, \text{if } st_{jr} \le t < et_{jr}, \ \forall t, r \quad (9)$$

$$x_{ij} \in \{0, 1\}, \ \forall i, j \quad (10)$$

$$0 \le E_{kw} \le |N|, \ \forall k, w \quad (11)$$

$$st_{jr}, et_{jr} \in \mathbb{Z}^+ \cup \{0\}, \ \forall j, r \quad (12)$$

$$Z_{kw} \in \mathbb{Z}^+ \cup \{0\}, Z_{kw} \le |N|, \ \forall k, w \quad (13)$$

Figure 3: The Global MIP Model for the Aircraft Maintenance Scheduling Problem.

The objective function (1) maximizes the number of aircraft assigned to a wave subject to a bound on the number of aircraft required and the expected number available (Constraint (5)). Equation (2) calculates the number of aircraft of type k whose repair due date is st_w. Equation (3) calculates the expected number of available aircraft for the first wave. Equation (4) calculates the expectation for the other waves. The first term includes those aircraft available but not used for the previous wave and those newly arrived from the repair shop. The second term sums over all aircraft that become available because they have completed waves since the previous wave started. Constraint (6) ensures that exactly one due date is assigned to each job. As noted earlier, the ordered set, D, includes the wave start-times and a big value B. If a failed aircraft cannot be repaired in time for one of the waves, its due date is assigned to B so as to not to constrain the problem. Equation (7) calculates the end-time of the jobs. The end-time of each job is guaranteed to be less than or equal to the assigned due date by constraint (8). Constraint (9) enforces the capacity limit of each trade, where t denotes each time point during which job j executes.

Z_{kw}, the number of aircraft of type k that are assigned to fly in wave w, is a true decision variable: we can choose to send fewer aircraft on a wave than are currently (in expectation) available. In contrast, E_{kw} is the expected number of aircraft of type k available for wave w and is based on the probabilistic outcomes of previous waves and the number of newly repaired aircraft (U_{kw}).

The objective function (1) is not the expected wave coverage because each wave has specific plane requirements and the maximum wave coverage for each wave is 1. If the expected number of available aircraft E_{kw} is more than the requirement a_{kw} for a given wave, the extra aircraft do not fly the wave and so do not have any contribution to the coverage. By not flying "extra" planes, we increase the probability that they will be available for the next wave.

Logic-based Benders Decomposition

As the problem requires making two different decisions, assigning aircraft to the waves and scheduling repair jobs for failed aircraft, a decomposition approach may be well suited. A logic-based Benders decomposition (LBBD) method can be formulated where the master problem assigns aircraft to waves to maximize wave coverage and the subproblems create the maintenance schedules given the due dates derived from the master problem solution. We propose three variations: *Benders-MIP* and *Benders-MIP-T*, where the master problems are solved using MIP, the latter with a tighter sub-problem relaxation; and *Benders-CP*, with a constraint programming-based master problem. All models use CP for the scheduling sub-problems.

The Due-Date Assignment Master Problem (DAMP): MIP Model To formulate the master problem as a MIP model, we use a binary variable x_{ij} for each $j \in J$ and $i \in D$ with the same meaning as in the global MIP model. A MIP formulation of DAMP is as follows:

$$\begin{aligned} &\text{max. Objective (1)} \\ &\text{s.t. Constraints (2) to (6), (10), (11), (13)} \\ &\quad \sum_{j \in M_r} c_{jr} p_{jr} \leq C_r \max_{j \in M_r,\ i \in D} (x_{ij} d_i),\ \forall r && (14) \\ &\quad \text{MIP cuts} && (15) \end{aligned}$$

The master problem incorporates a number of the constraints in the global MIP model. It does not represent the start-times of jobs nor does it fully represent the capacity of the trades. As is common in Benders decomposition, the master problem includes a relaxation of the sub-problems (Constraints (14)) and Benders cuts (Constraints (15)).

The Sub-Problem Relaxation Constraint (14) is the relaxation of the capacity of a trade, expressing a limit on the area of jobs that can be executed. The limit is defined using the area bounded by the capacity of the trade and the time interval $[0, M]$ where M is the maximum due date assigned to the jobs on the trade. This relaxation is due to Sadykov & Wolsey (Sadykov and Wolsey 2006).

As there are a relatively small number of waves, we can tighten the relaxation by enforcing an analogous limit on multiple intervals: $[0, st_w]$ for each wave w, plus $[0, M]$. For each interval, the sum of the areas of the jobs whose assigned due date is less than or equal to the end-time of the interval must be less than the available area. Formally, the tighter relaxation replaces Constraint (14) with:

$$\sum_{j \in M_r,\ \sum_{i=1}^{|D|} x_{ij} d_i \leq st_w} c_{jr} p_{jr} \leq st_w C_r, \quad \forall r, w$$

The Benders Cuts Before defining the cut formally, we demonstrate the intuition with an example. Consider a due date set, $D = \{14, 17, 20, 100\}$, and, for a given trade with five jobs, the current master solution: $x_{21} = 1, x_{12} = 1, x_{43} = 1, x_{14} = 1,$ and $x_{15} = 1$. Job 1 is assigned to the second due date, 17, Job 2 has the first due date, 14, and so on. If the current solution is infeasible due to the resource capacity of the trade, then we know that at least one of the jobs must have a later due date. We can, therefore, constrain the sum of the consecutive x_{ij} up to and including the ones assigned to 1 to be one less than the number of jobs. In our example, the cut would be:

$$\begin{aligned} &(x_{11} + x_{21}) + (x_{12}) + \\ &(x_{13} + x_{23} + x_{33} + x_{43}) + (x_{14}) + (x_{15}) \leq 5 - 1 \end{aligned}$$

Formally, assume that in iteration h, the solution of the DAMP assigns a set, Q, of due dates to the jobs on trade r. Assume further that there is no feasible solution on trade r with the assignments in Q. The cut after iteration h is:

$$\sum_{j \in M_r} \sum_{i \in I^r_{jh}} x_{ij} \leq |M_r| - 1, \quad \forall r \tag{16}$$

where $I^r_{jh} = \{i' | i' \leq i, \text{ and } x^h_{ij} = 1\}$ is the set of due dates indices less than or equal to the due date index assigned to job j and $|M_r|$ is the number of jobs on trade r. The validity of this cut is proved in the following section.

The Due-Date Assignment Master Problem: CP Model We also formulate the MP using CP. Let d_j be the variable corresponding to the due date for job j.

$$\begin{aligned} &\text{max. Objective (1)} \\ &\text{s.t. Constraints (3) to (5), (11), (13)} \\ &\quad \text{GCC}([U_{kw}], [st_1, st_2, ..., st_W], [d_{j \in I_k}]),\ \forall k, w && (17) \\ &\quad \sum_{j \in M_r} c_{jr} p_{jr} \leq C_r \max_{j \in M_r} (d_j),\ \forall r && (18) \\ &\quad d_j \in \{st_1, st_2, ..., st_W, B\} && (19) \\ &\quad \text{CP cuts} \end{aligned}$$

Constraint (17) defines a global cardinality constraint for each aircraft type. The global cardinality constraint enforces that the cardinality variables (U_{kw}) count the number of times that each value in the due date set appears over the due date variables (d_j). Constraint (18) guarantees that the sum of processing areas for the set of jobs on the same trade does not exceed the maximum available area.

The CP cut is based on the same reasoning as with the DAMP MIP model. If J_r is the set of job indices on trade r and the assigned set of due dates is not a feasible solution for the SP, the cut will guarantee that in the next iteration at least one of the assigned due dates will have a greater value. Formally, the cut is:

$$\bigvee d_j > d^h_j, \quad \forall j \in J_r \tag{20}$$

where d^h_j is the due date assigned to job j in iteration h.

Job Scheduling Sub-Problem Given a set of due dates assigned to the jobs on a trade, the goal of the job scheduling sub-problem (JSSP) is to assign start-times to the jobs to satisfy the due dates and the trade capacity. The JSSP for each trade can be modeled using cumulative constraints (Hooker 2005). We use a CP formulation:

$$\begin{aligned} &\text{cumulative}([t_j | d^h_j], [p_{jr} | d^h_j], [c_{jr} | d^h_j], C_r),\ \forall r \\ &0 \leq t_j \leq d^h_j - p_{jr},\ \forall j, r && (21) \end{aligned}$$

where t is an array of variables such that t_j is the start-time of job j, d is an array of values such that d^h_j is the due date assigned to job j in master problem in iteration h. The variables p_{jr}, c_{jr}, C_r are as defined above. Constraint (21) enforces the time windows: the job cannot be started later than $d^h_j - p_{jr}$.

A Dispatching Heuristic

Since the problem modeled in this paper is NP-hard, solving it to optimality may be prohibitive. We therefore investigate a heuristic approach, inspired by the Apparent Tardiness Cost (ATC) heuristic, a composite dispatching rule that is typically applied to single machine scheduling with the sum of weighted tardiness objective (Pinedo 2005). The heuristic computes a ranking index for each job and sorts the jobs in ascending order of the index. The heuristic then iterates through the jobs, scheduling each job at its earliest available time. The ranking index we use is as follows:

$$I_j = ST(k_j)\exp(-\frac{FN_j}{FC_j}), \quad \forall j$$

If we let k_j denote the type of aircraft j, then $ST(k_j)$ is the start-time of the first wave that requires an aircraft of type k_j. FN_j is the fraction of the total number of aircraft of type k_j required by the first wave that requires k_j, and FC_j is the maximum proportion of the capacity needed by job j over all its required trades, as follows.

$$FC_j = \max_r(\frac{p_{jr}c_{jr}}{ST(k_j)C_r}), \quad \forall r$$

Intuitively, the earlier the start-time of the first relevant wave, the higher proportion of aircraft required by that wave, and the lower the proportion of capacity required before the wave, then the higher the job will be ranked. The exponential function is used to place more weight on the start-time.

Hybrid Heuristic-Complete Approaches

A hybrid heuristic-complete approach in which the heuristic provides an initial solution may improve the performance of the complete approaches. Therefore, a simple hybrid first runs the dispatching heuristic and then uses the objective value as a starting lower bound for the complete approaches. Assume that the heuristic finds a solution, S, with cost $f(S)$. Any of the complete approaches can now be modified by adding the following constraint:

$$\sum_{w=1}^{W}\sum_{k=1}^{K} Z_{kw} > f(S)$$

Theoretical Results

To guarantee the finite convergence of an LBBD model to a globally optimal solution, the Benders cuts must be valid and the master decision variables must have finite domains. A Benders cut is valid in a given iteration, h, if and only if (1) it excludes the current globally infeasible assignment in the master problem without (2) removing any globally optimal assignments (Chu and Xia 2004). The former guarantees the finite convergence and the later guarantees the optimality. As the decision variables in DAMP have a finite domain, it is sufficient to prove the satisfaction of the two conditions.

Theorem 1. *Cut* (16) *is valid.*

Proof. For condition (1), for the sub-problem in iteration h on trade r, by definition:

$$\sum_{j\in M_r}\sum_{i\in I^r_{jh}} x_{ij} = |M_r|$$

Consequently, cut (16) excludes the current assignment of master problem.

For condition (2), consider a global optimal solution S that does not satisfy cut (16) as generated in iteration h. As the cut states that at least one job must have a greater due date than it had in h, to violate the cut, all jobs in S must have equal or lesser due dates than they had in iteration h. However, because the sub-problem was infeasible in iteration h, any sub-problem with only equal or lesser due dates must also be infeasible as the available area on the trade is the same or less. Therefore, S must be infeasible and we contradict the assumption that S is globally optimal.

Therefore, the cut is valid. □

An analogous argument holds for cut (20).

Experimental Results

The next sub-section describes the problem instances and the experimental details. We then compare our solution approaches experimentally and present insights into each algorithm's performance through a deeper analysis of the results.

Experimental Setup

The problem instances have 10 to 30 aircraft (in steps of 1), 3 or 4 trades, and 3 or 4 waves. Five instances for each combination of parameters are generated, resulting in 420 instances (21 total aircraft counts by 2 trade counts by 2 wave counts by 5 instances).

Aircraft The number of aircraft types is equal to $\frac{|N|}{5}$, where $|N|$ is the number of aircraft. The aircraft are randomly assigned to different types with uniform probability. The number of aircraft of type k is n_k. The failure rate for each aircraft is randomly chosen from the uniform distribution $[0, 0.1]$. The failure rate for aircraft of type k, λ_k, is the mean failure rate over all aircraft of type k. The values of α and β are 1 and 3, respectively.

Waves The plane requirement for each wave is randomly generated from the integer uniform distribution $[1, n_k]$. The length of each wave was drawn with uniform probability from $[3, 5]$. To make an instance loose enough to permit feasible solutions yet tight enough to be challenging, a lower bound on the time horizon is needed. The sum of the processing areas of the jobs in each trade, r, divided by the trade capacity is denoted by S_r. $LB = \max_r(S_r)$ is a lower bound on the time required to schedule the jobs and we set the scheduling horizon to $T = 1.2 \times LB$. The end-time for each wave, et_w, is generated as $et_W = T - rand[0, 3]$ for the final wave, W, and $et_w = st_{w+1} - rand[0, 3]$ for $w < W$.

Trades The capacity limit for each trade is set at $C_r = 10$.

Repair Jobs Eighty percent of the aircraft are in the repair shop at the beginning of time horizon, resulting in $|J| = 0.8|N|$ repair jobs. The jobs are randomly assigned to the trades with replacement such that the number of jobs per trade is equal to $|J|/2$. Each job requires at least one

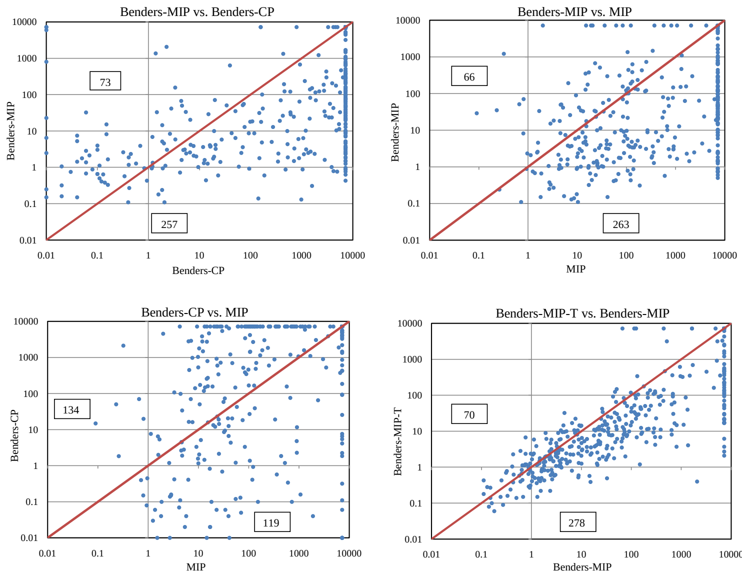

Figure 4: Run-times of the Four Complete Models

Method	Mean Time (s)	# Iter.	MP Time (s) per Iter.	SP Time (s) per Iter.	% Unsolved
Benders-MIP-T-Hybrid	995.25	14.15	71.24	0.13	11.90
Benders-MIP-T	1155.34	14.97	78.01	0.13	14.52
Benders-MIP	1842.61	16.75	111.81	0.08	23.10
MIP-Hybrid	3338.24	-	-	-	42.86
MIP	3566.19	-	-	-	46.90
Benders-CP	4181.72	34.71	123.04	0.02	52.86
Dispatching Rule	0.01	-	-	-	65.71

Table 2: The mean run-time, the mean number of master problem iterations, the mean run-time spent on the master problem and the sub-problems per iteration, and the percentage of unsolved problems for all approaches.

trade and some require more than one trade. The capacity of trade r used by job j, c_{jr}, is drawn from $[1, 10]$ while the processing time, p_{jr}, is drawn from $[r, 10r]$: jobs on trades with lower indices have shorter expected processing times than those on trades with higher indices.

All experiments were run with a 7200-second time limit. The MIP solver is IBM CPLEX 12.1 and the CP solver is IBM ILOG Solver/Scheduler 6.7.

Computational Results

Figures 4 show scatter-plots of run-times of the four complete approaches. Both axes are log-scale, and the points below the line $y = x$ indicate a lower run-time for the algorithm on the y-axis. The numbers in the boxes indicate the number of points below or above the line. Run-times are counted as equal if they differ less than 10%.

The graphs indicate a clear benefit for Benders-MIP over Benders-CP and MIP and for Benders-MIP-T over Benders-MIP while the performance comparison of Benders-CP and MIP is more even. Table 2 presents further data, sorted in ascending percentage of unsolved problems, for all algorithms.

Benders-MIP vs. MIP The Benders-MIP approach achieves a better run-time than MIP on 63% of the test problems and a worse run-time on 16%, achieving a lower mean run-time and solving a higher proportion of the problem instances. When the time horizon is short, the MIP approach is faster due to the tightness of the time-index model. With larger horizons and more jobs, the number of time-index variables grows, substantially reducing the performance.

Benders-CP vs. MIP The Benders-CP approach does better than MIP in terms of run-time on 28% of the instances while performing worse on 32%. When the CP master problem can be solved within the time limit, Benders-CP is superior to the MIP. However, Table 2 favors MIP in terms of the overall performance.

Benders-MIP vs. Benders-CP The Benders-MIP approach achieves a better run-time than Benders-CP on just over 61% of the test problems, performing worse on about 17%. The branching heuristics for Benders-CP (smallest to largest due date) often lead to an initial feasible master solutions with tighter due dates than the initial master solution in Benders-MIP. The tighter, globally infeasible initial solution means that the CP-based model requires more than twice as many iterations to find a globally feasible solution. However, when the initial master solutions are identical, Benders-CP often finds and proves optimality faster than Benders-MIP. We believe this is due to the different forms of the cut, but further research is necessary to fully understand this point. These observations suggest, however, that more intelligent branching heuristics may substantially improve Benders-CP, perhaps, to the point of being better than Benders-MIP.

Benders-MIP-T vs. Benders-MIP The tighter relaxation in Benders-MIP-T clearly speeds up LBBD: Benders-MIP-T has a better run-time than Benders-MIP on 66% of problems instances and worse on only 17%. The mean run-time, the percentage of unsolved problems, and the number of iterations decrease by 40%, 40%, and 11%, respectively. The solve time for the master problem decreases considerably compared to Benders-MIP, while the sub-problem run-time increases. This latter observation is because many sub-problems for Benders-MIP that can be quickly proved insoluble by the initial propagation of CP sub-problem model, violate the tighter relaxation in the Benders-MIP-T master problem. Therefore, in the tighter model, the sub-problem solver is not called on these "easy" sub-problems, increasing the mean run-time per sub-problem.

Incomplete and Hybrid Approaches The dispatching heuristic is fast, finding a feasible solution to all problems in an average of 0.096 seconds. However, it finds (but, of course, does not prove) an optimal solution in only 34% of the instances and Benders-MIP-T finds and proves optimality for these instances in 30.4 seconds on average. It seems that the heuristic can find the optimal solution only when the problem instance is relatively easy. The mean quality of the heuristic solution is 5% from optimal. In industries with expensive assets, such a reduction in solution quality can translate to costly under-use of a valuable resource (e.g., a fighter aircraft costs in the vicinity of 100 million dollars).

To evaluate the effect of combining the dispatching heuristic with the complete approaches, we examine using

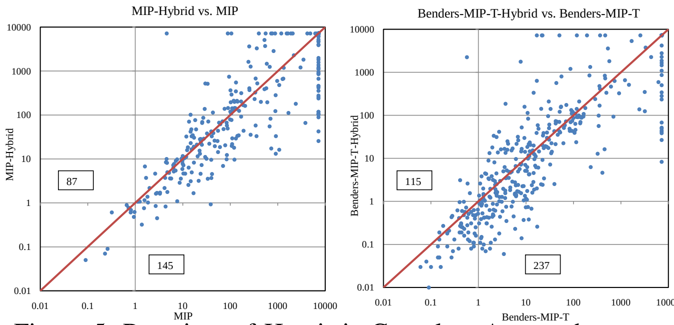

Figure 5: Run-time of Heuristic-Complete Approaches vs. Complete Approaches.

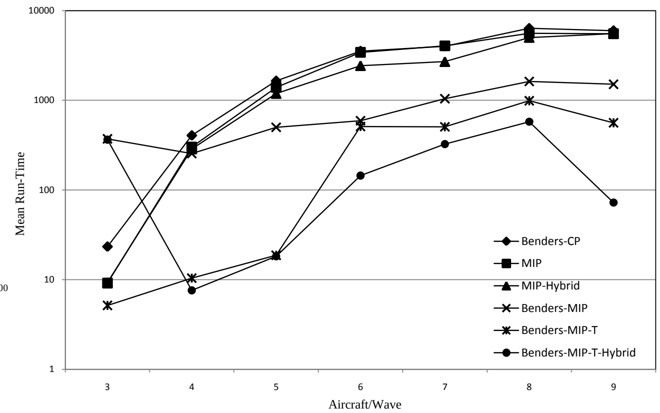

Figure 6: Mean Run-Time vs. Number of Aircraft per Wave ($|W| = 3$).

the hybrid heuristic-complete approach. A smaller feasible set is the direct consequence of defining a bound on the cost function. As the MIP model searches the feasible set, while LBBD methods explore the infeasible space, one intuition is that the MIP model should benefit more from using the heuristic solution. However, solving the master problem in LBBD requires searching in a relaxed feasible space and therefore the heuristic starting solution may also speed solving. Furthermore, as most of the LBBD run-time is taken up in solving the master problem, any such speed up is likely to impact the overall LBBD run-time.

Figures 5 and Table 2 show the effect of bounding the complete approaches with the dispatching heuristic solution. MIP-Hybrid achieves a better (worse) run-time than MIP on 35% (21%) of the problem instances. Benders-MIP-T-Hybrid has a better run-time than Benders-MIP-T on 60% of the instances and a worse run-time on 27%. The mean run-time decreases by 6% and 14% in MIP-Hybrid and Benders-MIP-T-Hybrid, respectively. As Table 2 indicates both hybrid methods are able to solve more problems to optimality. However, bootstrap paired-t tests (Cohen 1995) indicate that there is no significant difference in mean run-time at $p \leq 0.01$ for either hybrid.

Scalability Figure 6 shows our results as the number of aircraft per wave increases. We aggregate results by truncating $\frac{|N|}{|W|}$ and using the instances with three waves and both three and four trades. Note that each point represents 30 problem instances except $x = 3$ which has only 20 problems instances. We omitted $x = 10$ as we only have 10 problem instances for that point. The y-axis is log-scale.

The results show that the LBBD variations outperform the other techniques across all ratios. Interestingly, the LBBD hybrid approaches that include the dispatching heuristic show increasing performance with more aircraft per wave. This result suggests that the lack of statistical significance noted above, is mainly due to the results on the instances with a small number of aircraft per wave.

Summary The following observations are supported by the empirical results.

- The LBBD approach combining mixed integer programming and constraint programming outperforms the mixed integer programming model. Benders-MIP is over 300 times faster than MIP, on average, with a mean time ratio of 356.69. The time ratio for a given instance is calculated as the MIP run-time divided by Benders-MIP run-time.
- A tighter relaxation speeds up LBBD. Benders-MIP-T has a run-time over 50 times faster than Benders-MIP. (The mean time ratio is 51.02).
- A dispatching heuristic can provide the optimal solution for the easy problem instances. However, the mean percent relative error of heuristic is almost 5% overall, indicating that the dispatching rule by itself is not effective enough for industries with high equipment cost.
- Both MIP and LBBD benefit somewhat from a simple hybridization with the dispatching heuristic. While neither benefit is statistically significant overall, on problems with a larger number of aircraft per wave, the heuristic appears to have a strong positive impact on the LBBD approach.

Related Work

While there is a significant literature on maintenance scheduling for a deteriorating system, the majority of the articles assume an infinite horizon and provide tactical-level results, finding, for example, long-term maintenance frequencies or control-limits for maintenance depending on the state of components (Nicolai and Dekker 2008).

There are a number of pieces of work that examine operational level maintenance scheduling. Grigoriev et al. (Grigoriev, van de Klundert, and Spieksma 2006) took a deterministic approach to deal with the problem of cyclically scheduling maintenance activities assuming a fixed cycle length. Several formulations based on integer programming, flow formulation, and set-partitioning are proposed and their computational behaviour is investigated.

Budai et al. (Budai, Huisman, and Dekker 2006) considered the problem of scheduling preventive railway maintenance activities through a mathematical programming formulation. Maintenance jobs are assigned to different time periods to minimize the track down-time as well as the maintenance cost. The objective, unlike in many scheduling models, tends to force jobs close together.

Frost & Dechter (Frost and Dechter 1998) addressed the problem of optimally scheduling preventive maintenance of power generating units within a power plant. The goal was to find the duration and sequence of planned outages while minimizing operating and maintenance costs. The problem is cast as a continuous satisfaction problem and an "iterative learning" algorithm is used to solve the problem.

Haghani & Shafahi (Haghani and Shafahi 2002) studied the problem of scheduling bus maintenance activities. A mathematical programming approach for the problem is proposed to find the daily inspection and maintenance schedule in order to minimize the interruption in the daily bus operating schedule and to maximize the reliability of the system.

Safaei et al. (Safaei, Banjevic, and Jardine 2010) modeled the problem studied here as a MIP that includes an assignment problem and two network problems. The former assigned the aircraft to the waves, while the latter were used to calculate the expected number of available aircraft for the waves as well as the expected number of available workers for the repair jobs. The distinguishing feature of the problem is that the objective function (wave coverage) depends not only on the scheduling decisions but also on the outcomes of the pre- and post-flight checks. These two quite different components of the problem motivated the decomposition approach explored here.

Conclusion

In this paper, several algorithms are proposed to schedule an aircraft repair shop. The scheduling system is responsible for determining the start-time of maintenance activities to maximize the wave coverage considering the aircraft failure probabilities and maintenance capacity. Known statistics of probability of failure for each aircraft type are used to estimate the expected aircraft availability and therefore expected wave coverage.

The computational results showed that a logic-based Benders decomposition using a combination of mixed integer programming and constraint programming results in a three orders-of-magnitude speed up over a pure MIP model. A tighter relaxation in the master problem speeds up the logic-based Benders model another 50 times. A dispatching heuristic can provide a feasible schedule in a short time with five percent relative error on average and can also be used to speed-up the complete approaches, especially on harder problems instances with more aircraft per wave.

A focus of our future work is to consider the repair system in a dynamic environment. Above we assumed that the failed aircraft cannot be scheduled within the current time horizon. We are interested in relaxing this assumption such that new maintenance jobs can be scheduled as soon as failure is detected (Aramon Bajestani and Beck 2011).

Acknowledgments This research was supported in part by the Natural Sciences and Engineering Research Council of Canada, the Canadian Foundation for Innovation, the Ontario Research Fund, Microway, Inc., and IBM ILOG. Thanks to Nima Safaei for introducing the problem to us and numerous subsequent discussions.

References

Aramon Bajestani, M., and Beck, J. C. 2011. Scheduling a dynamic aircraft repair shop. In *Proceedings of the Scheduling and Planning Applications Workshop (SPARK)*. Submitted.

Budai, G.; Huisman, D.; and Dekker, R. 2006. Scheduling preventive railway maintenance activities. *Journal of the Operational Research Society* 57:1035–1044.

Chu, Y., and Xia, Q. 2004. Generating Benders cuts for a general class of integer programming problems. In *Proceedings of the First International Conference on the Integration of AI and OR Techniques in Constraint Programming (CPAIOR04)*, 127–136.

Cohen, P. R. 1995. *Empirical Methods for Artificial Intelligence*. The MIT Press, Cambridge, Mass.

Frost, D., and Dechter, R. 1998. Optimizing with constraints: a case study in scheduling maintenance of electric power units. *Lecture Notes in Computer Science* 1520:469–488.

Grigoriev, A.; van de Klundert, J.; and Spieksma, F. C. R. 2006. Modeling and solving the periodic maintenance problem. *European Journal of Operational Research* 172:783–797.

Haghani, A., and Shafahi, Y. 2002. Bus maintenance systems and maintenance scheduling: model formulations and solutions. *Transportation Research Part A* 36:453–482.

Hooker, J., and Ottosson, G. 2003. Logic-based Benders decomposition. *Mathematical Programming* 96:33–60.

Hooker, J., and Yan, H. 1995. Logic circuit verification by Benders decomposition. In Saraswat, V., and Van Hentenryck, P., eds., *Principles and Practice of Constraint Programming: The Newport Papers*. MIT Press. chapter 15, 267–288.

Hooker, J. 2005. A hybrid method for planning and scheduling. *Constraints* 10:385–401.

Nicolai, R. P., and Dekker, R. 2008. *Complex System Maintenance Handbook*. Springer Series in Reliability Engineering. chapter 11.

Pinedo, M. L. 2005. *Planning and Scheduling in Manufacturing and Services*. Springer Series in Operations Research. Springer.

Queyranne, M., and Schulz, A. 1994. Polyhedral approaches to machine scheduling problems. Technical Report 408/1994, Department of Mathematics, Technische Universitat Berlin, Germany. revised in 1996.

Sadykov, R., and Wolsey, L. A. 2006. Integer programming and constraint programming in solving a multimachine assignment scheduling problem with deadlines and release dates. *INFORMS Journal on Computings* 18:209–217.

Safaei, N.; Banjevic, D.; and Jardine, A. K. S. 2010. Workforce constrained maintenance scheduling for aircraft fleet: A case study. In *Proceedings of Sixteenth ISSAT International Conference on Reliability and Quality in Design*, 291–297.