FACULTEIT ECONOMIE
EN BEDRIJFSKUNDE

TWEEKERKENSTRAAT 2
B-9000 GENT
Tel. : 32 - (0)9 – 264.34.61
Fax. : 32 - (0)9 – 264.35.92

WORKING PAPER

A Hybrid Scatter Search Heuristic for Personalized Crew Rostering in the Airline Industry

Broos Maenhout [1]
Mario Vanhoucke [2]

March 2007

2007/454

[1] *Faculty of Economics and Business Administration, Ghent University, Gent, Belgium*
[2] *Faculty of Economics and Business Administration, Ghent University, Gent, Belgium*
Operations & Technology Management Centre, Vlerick Leuven Gent Management School, Gent, Belgium
broos.maenhout@ugent.be • mario.vanhoucke@ugent.be

D/2007/7012/25

ABSTRACT

The crew scheduling problem in the airline industry is extensively investigated in the operations research literature since efficient crew employment can drastically reduce operational costs of airline companies. Given the flight schedule of an airline company, crew scheduling is the process of assigning all necessary crew members in such a way that the airline is able to operate all its flights and constructing a roster line for each employee minimizing the corresponding overall cost for personnel. In this paper, we present a scatter search algorithm for the airline crew rostering problem. The objective is to assign a personalized roster to each crew member minimizing the overall operational costs while ensuring the social quality of the schedule. We combine different complementary meta-heuristic crew scheduling combination and improvement principles. Detailed computational experiments of all characteristics of the procedure are presented. Moreover, we compare the proposed scatter search algorithm with an exact branch-and-price procedure and a steepest descent variable neighborhood search.

Keywords: *airline crew rostering; meta-heuristics; scatter search*

1 Introduction

The crew scheduling problem in the airline industry is extensively investigated in the operations research literature since scheduling crew efficiently can drastically reduce operational costs of airline companies (Kohl and Karisch, 2004). The airline crew scheduling problem typically consists of assigning duties to crew members securing the safety of all flights and constructing a roster line for each employee minimizing the corresponding overall cost for personnel. In Europe, crew costs typically constitute the second largest expense after fuel costs, typically amounting 15-20% of total airlines operations costs (El Moudani et al., 2001). However, in the modern work environment, apart from minimizing the operational personnel costs, roster solutions additionally need to incorporate individual preferences since these have an important impact on the service quality (Ernst et al., 2004). Next to these often conflicting objectives, the assignment has to consider a multitude of restrictions forced by governmental regulations, union agreements, and company-specific rules restricting the construction of the flight duties as well as the crew members' lines-of-work (Cappanera and Gallo, 2004; Caprara et al., 1998a; El Moudani et al., 2001; Ernst et al., 2004; Kohl and Karisch, 2004; Thiel, 2004). Due to its complexity, the airline crew scheduling problem is usually decomposed into the airline crew pairing problem and the airline crew rostering problem which are likely to be solved sequentially. Both problems are known to be NP-hard (Yu, 1998).

The *crew pairing problem* aims to find a set of (anonymous) crew itineraries or pairings that covers the crew needs on each flight of the planning period minimizing the associated crew costs. A pairing is basically a sequence of flight duties, starting and ending at the same crew base and has to be assigned to one or more crew members working in one or more crew positions (ranks). The construction of pairings is constrained by many feasibility rules, i.e., company, union and regulatory rules (e.g., the maximum elapsed time, the minimum connection time between flights, the maximum number of working days, and the maximum number of flight legs) and is performed anonymously without consideration of a crew member's individual needs or desires (Hoffman and Padberg, 1993; Kohl and Karisch, 2004; Makri and Klabjan, 2004; Medard and Sawhney, 2004; Thiel, 2004; Vance et al., 1997).
In the *crew rostering problem*, the constructed set of pairings is assigned to individual crew members and sequenced to individual crew rosters considering all governmental rules, union- and company agreements as well as other activities such as pre-assigned activities (e.g., vacation, requested and granted off-periods, office, simulator/training, medical examination), ground duties, reserve duties (during which the employee must be available to replace another crew member who cannot work his/her assigned pairing), and off-duty blocks. In this step, all the pairings need to be assigned to as many crews as required by the flights in the pairing and each crew member receives a roster. The creation of rosters for individual crew members usually takes place 2 to 6 weeks before the flights are operated (Kohl and Karisch, 2004; Thiel, 2004).

Even though crew pairing problems are different from airline to airline with respect to rules and costs, the main characteristics remain the same. The rule structures are comparable and objective functions are based on pay and reflect mainly real costs. On the other hand, the crew rostering process generally aims not only at minimizing the operational cost for the airline company but also at maximizing the social quality as perceived by the crew members. In contrast to pairing, crew rostering can be done in various ways following different approaches. The so-called 'bidlines approach', especially used by American airline companies, first generates anonymous lines-of-work or bidlines which are assigned to individuals after an elaborated bidding process. By expressing a bid, i.e., a global preference, for the roster lines, a crew member knows exactly what he/she will get if the bid is granted. A drawback is the rather rigid structure of this approach since problems arise when some bidlines cannot be assigned entirely to crew members due to conflicts with pre-assignments and vacation days. The 'personalized rostering approach', however, directly constructs individual rosters for each crew member. This approach, particularly exercised by European airline companies, is based on a fair-and-equal share principle satisfying certain quality preferences involving workload (e.g., free days, night flying duties). In this approach the crew members express their preference for certain attributes of their rosters without knowing exactly how their rosters will look like (Kohl and Karisch, 2004; Medard and Sawhney, 2004; Thiel, 2004).

Most of the published work involving the airline crew scheduling problem relates to crew pairing (e.g., Anbil et al., 1991; Anbil et al., 1992; Andersson et al., 1998; Barnhart and Shenoir, 1996; Barnhart et al., 1994; Barnhart et al., 1995; Barutt and Hull, 1990; Desaulniers et al., 1997; Gershkoff, 1989; Graves et al., 1993; Hjorring and Hansen, 1999; Hoffman and Padberg, 1993; Housos and Elmroth, 1997; Lavoie et al., 1998; Levine, 1996; Vance et al., 1997). Compared to the crew pairing problem, crew rostering has received much less attention in the academic literature (Kohl and Karisch, 2004). The reason is that most of the cost benefits can be achieved by having productive pairings that minimize costs. Several approaches proposed in the literature make use of mathematical programming and column generation techniques where the subproblems, i.e., generating rosters for individual crew members, are solved as constrained shortest path problems (e.g., Cappanera and Gallo, 2001; Fahle et al., 2002; Freling et al., 2001; Gamache and Soumis, 1998; Gamache et al., 1998; Gamache et al., 1999; Jarrah and Diamond, 1997; Kohl and Karisch, 2000; Lasry et al., 2000; Sellmann et al., 2002). However, the crew rostering problem is known to be complex and difficult and that is why most of these approaches solve the problem heuristically. Others suggest meta-heuristic approaches to solve the airline crew rostering problem (e.g., Campbell et al., 1997; Christou et al., 1999; El Moudani et al., 2001; Gamache et al., 2007; Kerati et al., 2002).

In this paper, a scatter search procedure is presented to solve the airline crew rostering problem. In section 2, we describe the specific problem characteristics and formulate the problem as a generalized set partitioning type problem. In section 3, we explain the algorithmic details of our procedure. In section 4, we present computational results of all facets of the proposed meta-heuristic procedure and compare our procedure with the exact branch-and-price procedure of Gamache et al. (1999) and a variable neighbourhood search based on the principles of Hansen and Mladenović (2001). In section 5, conclusions are drawn and directions for future research are given.

2 Problem description and formulation

2.1 Problem description

In this paper we investigate a procedure to solve the airline crew rostering problem using the personalized rostering approach. The input for the crew rostering problem consists in general of crew information, activities to be rostered, rules and regulations, and objectives for the creation of a crew schedule.

When producing personalized rosters, each crew member's personal records (e.g., hours flown, trainings), qualifications (e.g., list of destinations, language proficiency), pre-assigned activities (e.g., office duties, training, medical checks), and vacation days are given.

The set of activities which are to be assigned consists of pairings, reserve duties, and pre-assigned activities. This set of activities covering the service requirements is assumed to be given, i.e., the crew pairing problem has been solved. Hence, each duty has a specific start time and a specific duration, i.e., service time.

The rules and regulations monitoring the quality of a schedule can basically be divided in horizontal and vertical rules (Kohl and Karisch, 2004).
Horizontal rules only imply the regulations governing the quality of a single crew member's roster line and/or restricting the sequencing of activities. The horizontal rules are concerned with the attributes of the crew member for which the roster is generated and with the properties of the assigned activities. In the following some important types of rules are given.

- Several provisions of the collective agreements affect which pairings and rest periods (breaks) can be assigned to each employee. Crew members can only be assigned to pairings and reserve duties compatible with the crew's qualifications and pre-assigned activities. For example, a crew member is qualified if he or she possesses a safety and rescue certificate for the type of aircraft to be used for the pairings or all crew members must possess the necessary visas. Hence, only activities consistent with the qualifications of a crew member and which do not overlap with pre-assigned activities can be assigned to the employee (Gamache et al., 1999; Kohl and Karisch, 2004; Medard and Sawhney, 2004).
- There are many rules governing the sequencing of activities and rest periods, e.g., 18 (22) hours of minimum rest after a day (night) activity, the number of consecutive working days or the number of consecutive (over)night duties (Gamache et al., 1999).
- All airlines have regulations restricting the amount of activities and rest days to be assigned to a single crew member. Often, weekly and/or monthly constraints are imposed concerning e.g., the number of overnight duties, the number of complete days as a rest period, the average weekly rest time, the number of total service time, the number of flight time.

Vertical rules involve the regulations concerning several roster lines and crew members. These rules mainly govern the composition of the crew assigned to a pairing. Each pairing must be staffed correctly by the correct number of cabin and service personnel. The composition of the crew assigned to a pairing must comply with certain requirements concerning the qualifications of individual members and the number of crew members performing each function on board of the plane, e.g., the number of experienced crew members, language qualification, and crew members who must/cannot fly together (Gamache et al., 1999).

These constraints can typically be treated as hard constraints whenever their violation may impair the security of the flight (e.g., crew qualifications, national legislation concerning duration of work and rest times) or as soft constraints preserving the social quality of the schedule (e.g., internal company rules, declared assignments preferences by the crew staff). In most airline companies the qualitative indicators have been integrated into the collective agreement and as such, they have been considered as (soft) constraints in the model definition (Gamache et al., 1999; El Moudani et al., 2001; Medard and Sawhney, 2004).

Airline companies usually set different kind of objectives, e.g., objectives related to real costs, objectives in order to preserve the quality of the schedule, and objectives related to individual preferences. In our procedure, the objective function consists of three major components. As a *first objective* we aim at the minimization of open time, i.e., unassigned activities. Moreover the objective is to maximize the total duration of pairings that is covered by the regular crew members during the planning period. Pairings not covered are then assigned to crew members working reserve blocks or doing over-time. When not all pairings can be staffed properly by the regular crew members, the understaffed pairings are assigned to supplementary crew which need to be hired at a fixed crew cost or to so-called freelancers. These employees do not receive a fixed salary but get paid for how much they fly per hour. As a *second objective*, the solution needs to ensure impartiality and fairness to all regular crew members. To implement the equal assignment criteria, the deviations from the average or ‘standard’ values of the involved resource consumption constraints are penalized. As a *third objective*, the crew members can express their preferences for certain roster attributes. The crew members can not only express their preferences for specific pairings and reserve duties but also for more general scheduling preferences as the moment (day/time) they want to be scheduled.

2.2 Problem formulation

The formulation of the airline crew rostering problem is based on the Dantzig-Wolfe decomposition breaking up the original problem formulation into a master rostering problem and a subproblem obtaining a feasible crew roster respecting all horizontal rules (Ryan, 1992). The master rostering problem is modelled as a generalized set partitioning type problem. Solving the problem generates exactly one roster for each regular crew member and at most one roster for freelance and extra personnel such that the demands of the activities are met, the solution satisfies constraints between several crew members and the objective is optimized (Gamache et al., 1999; Kohl and Karisch, 2004). The master airline crew rostering problem can be formulated as follows,

Notation

Crew members

$C^r / C^e / C^f$	set of regular/extra/freelance crew members to be scheduled (index *i*, $i \in C^r / i \in C^e / i \in C^f$)
S	set of skill categories (index *k*, $k \in S$)
$b^r_{ik} / b^e_{ik} / b^f_{ik}$	1 if regular/extra/freelance crew member *i* has skill category *k*, 0 otherwise

Individual crew rosters

$F^r_i / F^e_i / F^f_i$	set of feasible roster lines for regular/extra/freelance crew member *i* with respect to all hard horizontal rules (index l, $l \in F^r_i / l \in F^e_i / l_h \in F^f_i$)

Activities

A	set of activities to be scheduled (index *j*, $j \in A$ with $A = P \cup R \cup Q$)
P	set of pairings to be scheduled (index *m*, $m \in P$)
R	set of reserve duties to be scheduled (index *n*, $n \in R$)
Q	set of pre-assigned activities to be scheduled (index *o*, $o \in Q$)
R_{jk}	required number of crew members of skill category *k* for activity *j*
t_j	service time of activity *j*

Other parameters

$a^r_{ijl} / a^e_{ijl} / a^f_{ijl}$	1 if roster l for regular/extra/freelance crew member *i* covers activity *j*, 0 otherwise
c^r_{il}	penalty cost of roster l for regular crew member *i* for the complete period

c_k^e	penalty cost of hiring extra personnel of skill category k for the complete period
c_k^f	penalty cost of hiring freelance personnel of skill category k per hour
c_{jk}^u	penalty cost of understaffing activity j in skill category k
c_{jk}^o	penalty cost of overstaffing activity j in skill category k

<u>Decision variables</u>

$y_{il}^r / y_{il}^e / y_{il}^f$	1 if regular/extra/freelance crew member i is assigned to roster l, 0 otherwise

Model formulation

$$\text{Min} \quad \sum_{i\in C^r}\sum_{l\in F_i^r} c_{il}^r y_{il}^r + \sum_{i\in C^e}\sum_{l\in F_i^e}\sum_{k\in S} c_k^e b_{ik}^e y_{il}^e + \sum_{i\in C^f}\sum_{l\in F_i^f}\sum_{j\in A}\sum_{k\in S} c_k^f b_{ik}^f t_j a_{ijl}^f y_{il}^f + \sum_{j\in A}\sum_{k\in S} c_{jk}^u t_j v_{jk} + \sum_{j\in A}\sum_{k\in S} c_{jk}^o w_{jk} \tag{1}$$

$$\text{s.t.} \quad \sum_{i\in C^r}\sum_{l\in F_i^r} b_{ik}^r a_{ijl}^r y_{il}^r + \sum_{i\in C^e}\sum_{l\in F_i^e} b_{ik}^e a_{ijl}^e y_{il}^e + \sum_{i\in C^f}\sum_{l\in F_i^f} b_{ik}^f a_{ijl}^f y_{il}^f + v_{jk} - w_{jk} = R_{jk} \qquad \forall j\in A; \forall k\in S \tag{2}$$

$$\sum_{i\in F_i^r} y_{il}^r = 1 \qquad \forall i\in C^r \tag{3}$$

$$\sum_{i\in F_i^e} y_{il}^e \le 1 \qquad \forall i\in C^e \tag{4}$$

$$\sum_{i\in F_i^f} y_{il}^f \le 1 \qquad \forall i\in C^f \tag{5}$$

$$y_{il}^r \in \{0,1\} \qquad \forall i\in C^r; \forall l\in F_i^r$$

$$y_{il}^e \in \{0,1\} \qquad \forall i\in C^e; \forall l\in F_i^e$$

$$y_{il}^f \in \{0,1\} \qquad \forall i\in C^f; \forall l\in F_i^f \tag{6}$$

The objective (1) minimizes the total penalty cost of assigning a feasible roster to regular, extra, and freelance personnel. The penalty cost c_{il}^r of a roster line assigned to regular crew members is the sum of the crew preferences for roster attributes, i.e., preferences for assignable activities and for the preferred rest and working periods, and the accounted penalty costs when deviating from the average or 'standard' constraint values in order to incorporate fairness between the crew members. The penalty costs c_k^e of a roster assigned to extra crew members is a fixed lump sum representing the wage cost when hiring these people for the complete scheduling period. When assigning activities to freelance personnel, the penalty cost is equal to an hourly wage cost c_k^f times the number of working hours scheduled for the respective freelance crew member. Furthermore, the objective minimizes the duration of understaffed activities and the number of times more personnel than needed is scheduled for an activity. Constraint (2) requires that each activity is adequately staffed for each skill category. To ensure mathematical feasibility, two slack variables v_{jk} and w_{jk} are associated with constraint (2) penalizing under- and overstaffing respectively. Constraint (3) requires that exactly one roster is assigned to each regular crew member. Constraint (4) and (5) indicate that no more than one roster can be assigned to extra or freelance crew members respectively. Constraint (6) is the binary constraint assuring the integrality of the variables.

The subproblem is to obtain a feasible roster line for a crew member and can be modelled as a resource constrained shortest path problem (Gamache and Soumis, 1999). This is a minimum cost flow problem for which a separate network needs to be constructed for each crew member. All crew members' networks are acyclic having a source and a sink node representing the beginning and end of the month. The nodes in the network, except the source or the sink node, represent pairings, reserve blocks, rest periods and/or pre-assigned activities. Any activity compatible with the crew member's qualifications is represented in the network. Each arc of the network represents a link between two consecutive activities and/or rest periods. The network structure typically incorporates the succession and qualification rules for a single crew member. In figure 1 the pairings, reserve duties, and pre-assigned activities for an example problem are indicated. This problem is transformed to a network for a single crew member as displayed in figure 2.

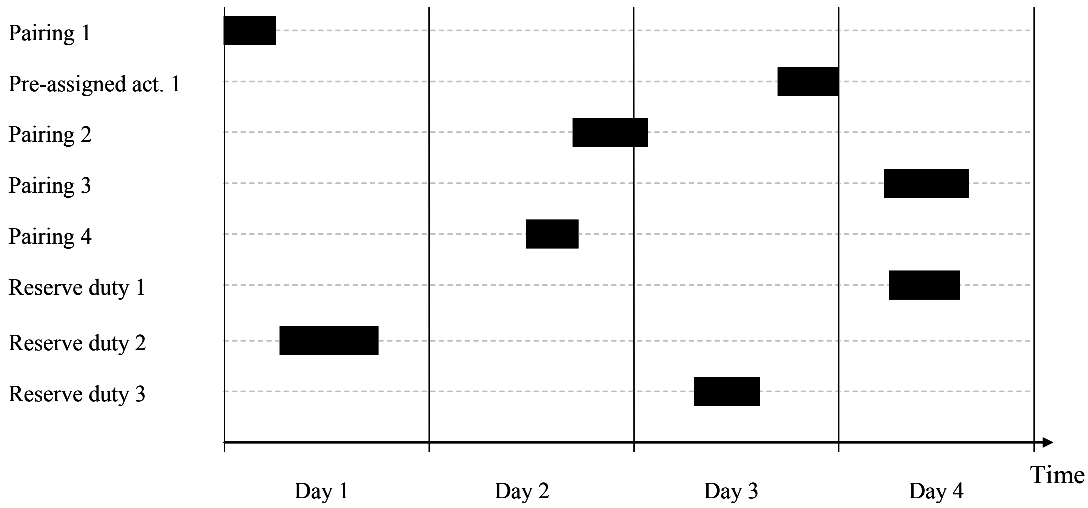

Figure 1. An example problem instance

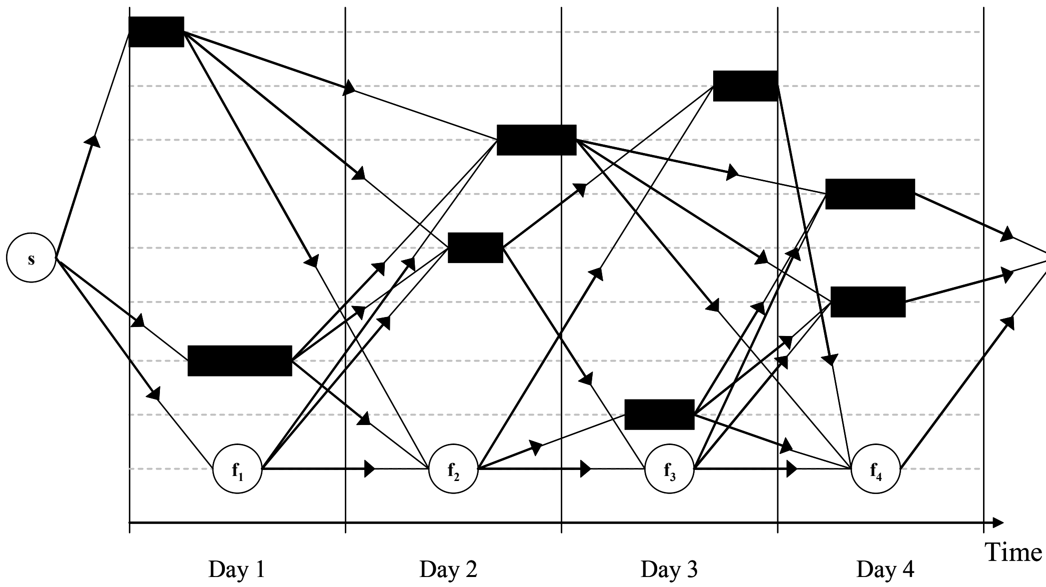

Figure 2. Example network structure for a crew member
(s – source node; t – sink node; f_d – free day or rest period)

However, the structure of the network is not sufficient to guarantee that all horizontal constraints on roster construction are satisfied by paths in the network. Hence, some constraints need to be taken into account during the calculation of the shortest path. A number of restrictions cannot be modelled directly in the network structure because activities are accompanied by multidimensional resource vectors which consumption is accumulated along paths and constrained at intermediate nodes (e.g., service time, flight time). For each of these resources, fairness values are postulated in order to obtain an equal distribution of the workload among the regular crew members. These fairness measures are calculated as the minimum of two values, i.e., the total resource consumption of all eligible activities for the specific resource averaged over the regular crew members and the maximum allowable resource consumption for a single crew member described by governmental, union-, or company rules.

An overview of the different pseudo-polynomial methodologies developed for solving resource constrained shortest paths can be found in Irnich and Desaulniers (2004). We implemented a dynamic programming approach for solving the resource constrained shortest path problem at hand (Desrosiers et al., 1995, Gamache et al., 1999). This algorithm generates feasible roster lines incorporating each employee's peculiarities and incorporates additional penalty costs to relax some horizontal constraints in order to ensure fairness and/or to violate soft regulations (Desrochers and Soumis, 1989; Ioachim et al., 1998).

3 Solution procedure

Evolutionary algorithms are adaptive heuristic search procedures imitating the ideas of natural selection and the survival of the fittest. As such, they represent an intelligent exploitation of a random search within a defined search space to solve the problem under study. Scatter Search (Glover, 1998) is a population-based meta-heuristic in which solutions are intelligently combined to yield better solutions. This meta-heuristic template appeals to strategic designs where other approaches resort to randomization. The scatter search methodology is very flexible and has successful applications in several application areas, since each of its elements can be implemented in a variety of ways and degrees of sophistication. For an overview of the basic and advanced features of the scatter search meta-heuristic, we refer to Glover and Laguna (2000) and Marti et al. (2006). The generic framework of the scatter search methodology is as follows, i.e.,

```
Algorithm Scatter Search
   Diversification Generation Method
   While Stop Criterion not met
      Subset Generation Method
      Solution Combination Method
      Improvement Method
      Reference Set Update Method
   Endwhile
```

In the following, we describe the implemented meta-heuristic principles for solving the airline crew rostering problem. The dedicated search heuristic is performed until a maximum number of schedules is evaluated.

3.1 Data representation and fitness function

In evolutionary heuristics, the population elements are encoded in a problem-specific data structure. The employed data representation in this paper directly reflects the activities the crew member is assigned to over the planning period. This binary encoding indicates a 1 if the corresponding crew member is scheduled to operate the corresponding activity and 0 otherwise. Other meta-heuristic approaches utilize a non-binary representation where for each activity the operating crew member is indicated (Christou et al., 1999; El Moudani et al., 2001; Kerati et al., 2002). An example of the employed data structure is provided in figure 3. The activities (the pairings, the reserve duties, and the pre-assigned activities) and the crew members (the regular crew members, the extra personnel, and the freelance personnel) are displayed respectively on the horizontal and the vertical axis in the matrix.

	Activity / Crew member	Pairings: 1	…	$\lvert P \rvert$	Reserve duties: 1	…	$\lvert R \rvert$	Pre-assigned activities: 1	…	$\lvert Q \rvert$
Regular crew members	1	1	…	0	0	…	0	0	…	1
	…	…	…	…	…	…	…	…	…	…
	$\lvert C^r \rvert$	0	…	0	0	…	0	1	…	0
Extra personnel	1	0	…	0	1	…	0	0	…	0
	…	…	…	…	…	…	…	…	…	…
	$\lvert C^e \rvert$	1	…	0	0	…	1	0	…	0
Freelance personnel	1	0	…	0	1	…	0	0	…	0
	…	…	…	…	…	…	…	…	…	…
	$\lvert C^f \rvert$	0	…	1	0	…	0	0	…	0

Figure 3. The data representation

These chromosomes serve as input for a fitness function which evaluates the quality of the solution encoded in each chromosome. The solution quality is calculated as a weighted average of the operational costs and the social quality of the population element as described in section 2.1.

3.2 The Diversification Generation Method

In this step, a large set of solution vectors is initialized. Since useful information about the structure of optimal solutions is typically contained in a suitably diverse collection of elite solutions, a diverse set of initial solutions is generated (Glover and Laguna, 2000). Kapsalis et al. (1993) and Reeves (1995) pointed out that introducing high-quality solutions, obtained from a heuristic technique, can help a meta-heuristic to find good solutions more quickly. Hence, we utilize a constructive heuristic which schedules the crew members in a random sequence by solving the identified subproblem taking appropriate preferences and penalty costs into account (Moore et al., 1978; Byrne, 1988). We implemented three different methodologies initializing the population elements, i.e., the first generates all solutions randomly, the second generates all solutions heuristically, and the third strategy generates a subset of the solutions with the constructive heuristic and a subset randomly. A subset of these initialized population elements are designated to be reference solutions. $Refset_1$ contains the best b_1 population elements in terms of solution quality. $Refset_2$ contains the b_2 most diverse solutions with respect to the solutions incorporated in $Refset_1$. The distance between two solutions is a measure for diversity and is calculated as the number of different assignments between the two solutions.

3.3 The Subset Generation Method

After the initialization phase, scatter search operates on this reference set by combining pairs of reference solutions in a controlled, structured way. In the subset generation method two elements of the reference set are chosen in a systematic way to produce points both inside and outside the convex regions spanned by the reference solutions. Glover and Laguna (2000) suggest to create new solutions out of all two-element subsets. Choosing the two reference solutions out of the same cluster (i.e. $Refset_1 \times Refset_1$ or $Refset_2 \times Refset_2$) stimulates intensification, while choosing them from different clusters ($Refset_1 \times Refset_2$) stimulates diversification. The way each pair of solutions is selected from these three combinations of reference sets has been varied along three different subset generation strategies. The *completely random selection* selects each pair of solutions at random from one of the reference sets $Refset_1$ or $Refset_2$. The *controlled random selection* selects solution pairs from both reference sets in a controlled way. More precisely, this mechanism is determined by pre-defined probabilities that a pair of solution elements comes from $Refset_1 \times Refset_1$, $Refset_2 \times Refset_2$ or from $Refset_1 \times Refset_2$. The solutions are randomly selected out of the determined reference sets. The *controlled selection based on fitness/diversity value* employs the same mechanism to determine a pair comes from $Refset_1 \times Refset_1$, $Refset_2 \times Refset_2$ or from $Refset_1 \times Refset_2$. However, this selection strategy selects the combination elements based on their fitness value when element of $Refset_1$ and/or the distance when element of $Refset_2$. Fitter population elements of $Refset_1$ and more diverse population elements of $Refset_2$ with respect to $Refset_1$ have a higher probability of being selected.

3.4 The Solution Combination Method

At each evolution cycle, the two selected population elements exchange information in such a way a new individual is created with attributes of both the parent solutions. In this paper, solution points are combined in different ways as described below based on crew-based combination methods or path relinking. Analogous to natural selection, all these heuristic principles thrive on the idea that the parent solutions will pass their good characteristics on to the newly created solution points. Hence, the algorithm preserves or even improves the good characteristics of the parent solutions as the population evolves.

3.4.1 Crew-based crossover operators

The *one-point crossover operator* randomly selects a crossover point between 1 and the number of regular crew members, such that the individual crew rosters before the crossover point are copied from the one parent and the individual crew rosters after the crossover point are copied from the other parent (Beasley and Chu, 1996; Aickelin, 2000).

The *crossover operator with best tournament selection* creates a child schedule that combines the best individual crew rosters from both parents. We have implemented this crossover operator in two different ways into our algorithm discerning between two notions of 'best' individual crew rosters, i.e., best in terms of social quality and in terms of overall solution quality (operational costs and social quality) (Burke et al., 2001).

The *crossover operator with random tournament selection* selects randomly a crew roster out of the two parent solutions for each crew member (Aickelin, 2000; Burke et al., 2001; Dias et al., 2003).

3.4.2 Path relinking approach

The path relinking approach (Glover, 1998) combines solutions using convex linear or non-linear combinations. The process of generating linear combinations may be characterized as generating paths between solutions in Euclidian space (Glover and Laguna, 2000; Glover, 1998). A path between solutions will generally yield new solutions that share a significant subset of attributes contained in both parent solutions, which can differ according to the path selected. The moves introduce attributes contributed by a guiding solution and/or reduce the distance between the initiating and the guiding solution. The goal is to capture the assignments that frequently or significantly occur in high quality solutions, and then to introduce some of these compatible assignments into other solutions that are generated by a heuristic combination mechanism.

In combining the parent solutions, a feasible roster line will be generated for each crew member by solving the identified subproblem with appropriate costs. The accounted costs are the result of a weighted average based on a (negative) cost given to the assigned activities present in both parent solutions maintaining the good characteristics of the initiating solution and introducing attributes of the guiding solution, the social quality, and the operational costs. In this paper, we have implemented this combination mechanism in two different ways, i.e., a *restricted path relinking* approach where only the activities present in the guiding and initiating solution are considered and a *total path relinking* approach where all activities compatible with the crew member's qualifications are taken into consideration.

3.5 The Improvement Method

The improvement method applies heuristic processes to improve the total operational costs, fairness among the crew members and the crew members' preferences of the newly generated solution points. To that purpose, we implemented different complementary local search algorithms based on the findings of Ahuja et al. (2002), each focusing on a different part of the scheduling matrix, i.e., the single crew roster-based local search (focusing on the whole roster line of a single crew member), the period-based local search (focusing on a single period for all crew members), the activity-based local search (focusing on a single activity for all crew members), and the schedule-based local search (focusing on the whole schedule for all crew members). The single crew roster-based local search was already implemented in a heuristic procedure for the crew scheduling problem by Moore et al. (1978) and Byrne (1988) making a single pass over an employee list. The drawback of this method is the poor roster quality for the employees scheduled last. The period-by-period construction of a crew schedule was already investigated in a single pass heuristic by Nicoletti (1975), Buhr (1978), Tingley (1979), and Sarra (1988) but was unable to cope with the potential difficulties on succeeding days. Giaferri et al. (1982) combined the single crew roster-based local search and the period-based local search in order to overcome problems of both single pass heuristics. Maenhout and Vanhoucke (2005) exploited these improvement methods in a meta-heuristic procedure for the nurse scheduling problem.

In the following, we describe these different improvement methods and illustrate the respective appropriate network structures on an example problem instance modeling the problem as a multicommodity flow problem (Cappanera and Gallo, 2004). Figure 4 displays the neighbourhoods of the different local search heuristics exercised on an initial crew schedule. The faded nodes and arcs are not taken into account when solving the various minimum cost flow problems. The dark arcs and nodes designate the respective paths and activities crew members can possibly be assigned to. In the example four crew members are to be scheduled over a period of 4 days. All crew members have a source (s_i) and a sink (t_i) node. The activities which need to be assigned to the crew members are displayed in figure 1. The pre-assigned activity needs to be covered by crew member 3. All activities require 1 crew member except from pairing 1 which needs 2 crew members. Crew members have a rest day (displayed as a circle node) when they are not assigned to an activity that day. The horizontal regulations are limited to the sequencing rules modelled in the network structure of figure 2. Crew member 1, for example, works pairing 1 and 2 and has a day off on day 4.

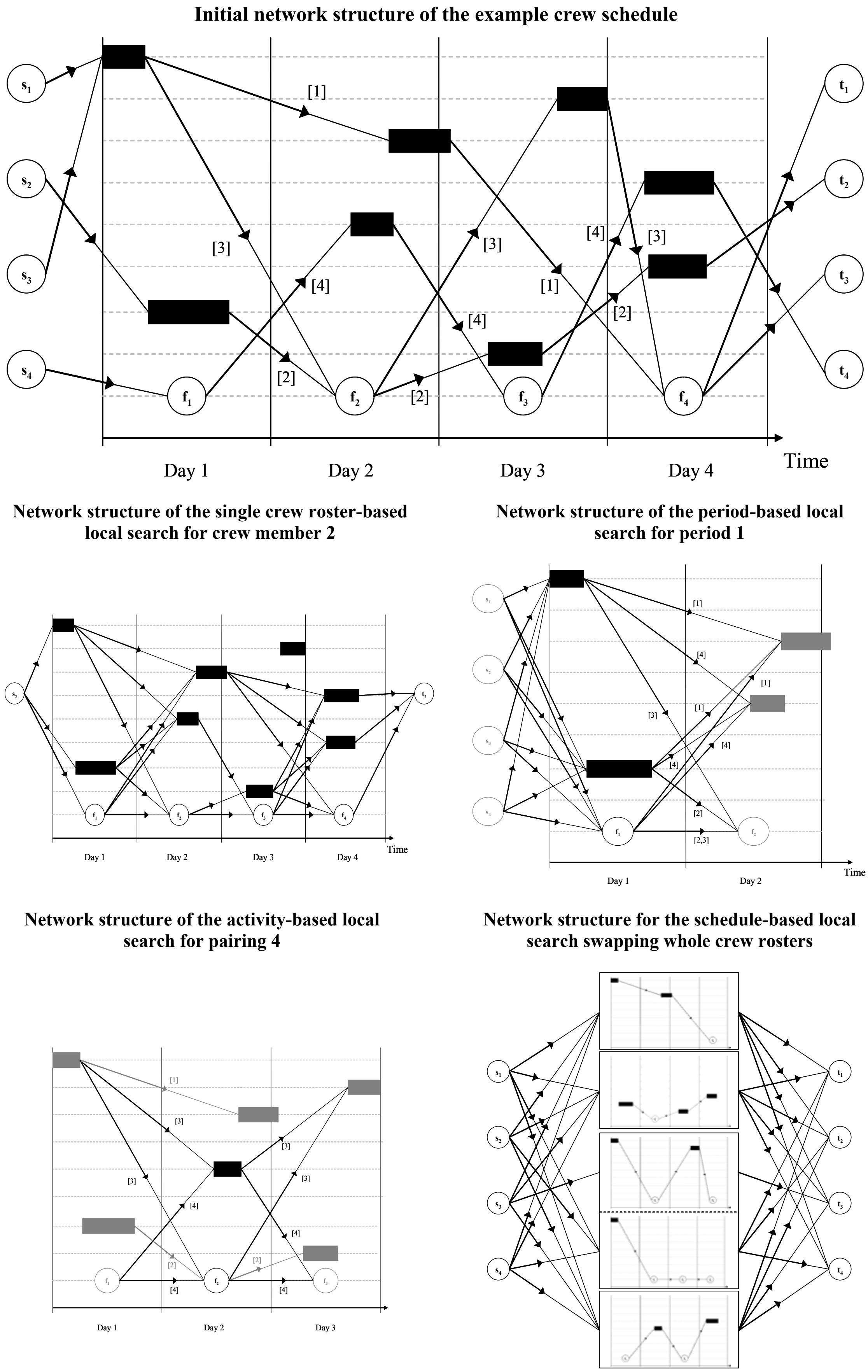

Figure 4. Illustration of the different improvement methods on an example crew schedule

The *single crew roster-based local search* optimizes the line-of-work of a single crew member, given the roster lines of the other crew members. This local search schedules the crew members in a random sequence one by one by solving the identified subproblem with appropriate penalty costs for under- and overstaffing, fairness penalty costs, and crew preferences. In figure 4 we have illustrated the appropriate network structure when optimizing the crew roster of crew member 2.

The *period-based local search* optimizes the schedule for one period given the assignments of the crew members in all other periods. The length of this period during which the assignments are optimized is determined such that at most one activity can be assigned to a crew member within the period (e.g., one day). In this way, activities can linearly be assigned to the individual crew members based on crew preferences, fairness, and appropriate penalty costs for under- and overstaffing. To that purpose, the roster schedule over the period is converted to a linear assignment problem matrix. The cost matrix consists of the penalty cost and crew satisfaction of the whole roster line if the activity under consideration would have been incorporated. In constructing the matrix, each activity has a number of columns that is equal to its staffing requirements. Moreover, dummy activities and dummy crew members are added to allow the under- and overstaffing of activities. The cost of assigning dummy crew members to activities is equal to the penalty cost of understaffing the activity. When non-dummy crew members are assigned to dummy activities, these crew members do not contribute to the actual required staffing of the activities considered in the respective period. These crew members will overstaff activities or will be assigned a rest period. Since the problem is a minimum cost flow problem, the dummy activities represent the best assignment for a particular crew member. The cost of this assignment is equal to the minimum assignment cost of the (feasible) activities or the rest period the crew member can be assigned to plus the penalty cost for overstaffing the particular activity. Of course, certain activity assignments are excluded from the matrix in order to cope with all the constraints, taking into account the fixed assignments in the other periods of the current solution. For each period of the planning horizon, a linear assignment problem is solved by means of the Hungarian method (Kuhn, 1955). In figure 4 we have displayed the appropriate network structure when optimizing period 1 of the crew schedule. In our example a period corresponds to 24 hours, i.e., one day. We have to optimally re-assign pairing 1 and reserve duty 2 given the crew members' assignments on the other days.

The *activity-based local search* optimizes the assignment of a single activity given all the other assignments of the crew members. This local search is implemented similarly to the period-based local search but focuses on only one activity in contrast to the period-based local search which focuses on all activities within the period. This local search defines a linear assignment problem for each activity. In figure 4 we have displayed the appropriate network structure when re-assigning pairing 4 over the crew members.

The *schedule-based local search* aims to improve the crew members' satisfaction by swapping (parts of) lines-of-work between crew members. This problem is solved by defining a linear assignment problem that optimally re-distributes the lines-of-work of the current schedule among the crew members based on Kuhn's method (1955). This re-distribution mechanism has only effect on the social quality of the schedule. The algorithm tries first to swap complete lines-of-work and tries then to swap parts of the roster lines (i.e., parts of two and three periods, whether or not consecutive) between the crew members taking only the pairings and reserve duties into consideration. Each constructed LAP matrix consists of all possible swapping alternatives between crew members taking into account the crew member's qualifications and possibly fixed assignments. The assignment cost consists of the cost of the complete roster line after the swap would have been executed. In figure 4 we have displayed the appropriate network structure when re-assigning the complete crew rosters over the crew members. Since only pairings and reserve duties are swapped between crew members the crew roster of crew member 3 is adapted when assigning the roster line to one of the other crew members leaving out the pre-assigned activity for crew member 3.

3.6 The Reference Set Update Method

The population evolves over time with the entrance of new solution points and the drop-out of old solutions, searching to improve the quality of the best known solution. A new solution may become a member of the reference set either if the solution point has a better objective function value than the solution point with the worst objective function value in $Refset_1$ or if the solution point is more diverse with respect to $Refset_1$ than the least diverse solution point in $Refset_2$. The reference set undergoes a static or a dynamic update. In contrast to a static update where the reference set is updated only after a number of solution points is generated, a dynamic update evaluates each possible reference set entrance instantly. In both cases "better" solutions replace the worst solution point in the concerned reference set and the ranking is updated. During the search, diversity in the reference set is maintained through the use of the artificial tiers in the reference set and by preventing the duplication of solution points and/or the entrance of highly resembling solutions.

4 Computational Results

In this section we present computational results for the meta-heuristic procedure under study, carried out on a Dell computer with a Dual Core processor 2.8 Ghz and 2 Gb RAM. In section 4.1, we describe the specific settings and characteristics of the problem instances and the parameter settings of the described meta-heuristic procedure. In section 4.2, we present detailed computational results showing the performance of the implemented meta-heuristic principles. In section 4.3, we compare our scatter procedure with a variable neighbourhood search based on the principles of Hansen and Mladenović (2001) and the exact branch-and-price procedure of Gamache et al. (1999).

4.1 Test design

4.1.1 Problem instance characteristics

To construct a robust procedure we generated several artificial test files based on real-world problem characteristics from literature (Caprara et al., 1998; Gamache et al., 1999; Cappanera and Gallo, 2004; Kohl and Karisch, 2004). We have generated 200 problem instances of 10 regular crew members, 10 extra crew members (on the average), and 10 freelance crew members (on the average), and 90 activities (40 pairings, 40 reserve duties and 10 pre-assigned activities). The set of extra and freelance crew members, i.e., C^e and C^f, are determined based on lower bounds described in Caprara et al. (2003) which calculate the minimum number of needed personnel for each skill category. These lower bounds indicate that the number of employees must be at least equal to the maximum daily request or to the ratio between the overall number of activities in the planning horizon and the maximum number of activities which can be assigned to a crew member. The wage costs for employing extra and freelance crew members amount respectively to 5,000 and 100.

The problem instances typify the characteristics and attributes of the activities and crew members. The crew members have an explicit skill competency and operate as cabin personnel (e.g., captain, first officer) or service personnel (e.g., steward, hostess). Moreover, for each crew member multiple qualifications are described, i.e., his/her seniority, the types of aircraft he/she is able to operate on, the type of airport he/she is able to take off/land, the propertied visas and the language proficiency. Furthermore, the crew members express their preference for each activity and for the moment (day/time) they want to be scheduled. Activities have a specific service time resulting from the pre-determined start and end time. Pairings and reserve duties are characterized by a required number of crew members for each skill competency performing the task. Pairings are also described by a specific flight time, the required number of experienced crew members of cabin and service personnel, the required number of crew members of cabin and service personnel knowing all the safety regulations of the specific pairing, the aircraft type, the required visa type and required language competency. Pre-assigned activities are characterized by the crew member required to perform the duty. Based on these specifications, the compatibility of a crew member with a specific activity can be determined in advance. Moreover, based on the start and end times, an activity is defined as a morning activity (starting between 6 am and 10 am), a day activity (scheduled between 10 am and 22 pm), a night activity ((partially) scheduled between 22 pm and 6 pm), or an overnight activity (starting before 22 pm and finishing after 6 pm). A complete day (i.e., a time interval of 24 hours starting at midnight) is called idle if no duty or part of a duty is executed during that day, otherwise the day is called working (Caprara et al., 1998a; Caprara et al., 1998b).

Furthermore, the imposed horizontal and vertical rules need to be defined and specified according to the company's policy. The vertical rules constrain the required number of crew members for each skill category, the required number of experienced crew members for cabin and service personnel and the required number of crew members for cabin and service personnel knowing all safety regulations of the activity. The penalty cost of not satisfying the vertical rules amounts to 1,000 for pairings, 100 for reserve duties and 10,000 for pre-assigned activities. In order to minimize the open time adequately, these penalty costs are multiplied by the respective service time of the understaffed activity. Horizontal regulations guard the social quality of a crew's roster line and/or restrict the sequencing of activities. Obviously crew members can only be assigned to one activity at a time and to activities which are compatible with the crew's qualifications. There is a minimum rest time restricting the sequencing of activities of 18 hrs after a day or morning activity and of 22 hrs after overnight and night duties. Furthermore, weekly and monthly restrictions are imposed constraining the attributes of a crew's roster line. An overview of all incorporated horizontal regulations for regular and extra crew members is provided in table 1 together with appropriate minimum and maximum values for each constraint. Moreover, the fairness values and the accounted penalty costs when the characteristics of the roster assigned to regular crew members deviate from these averages are also displayed in the table.

Table 1. Horizontal rules restricting a crew's roster line

Constraints	Minimum value	Maximum value	Fairness value	Fairness Penalty cost
overnight duties/week	0	2	(*)	1
number of pairings/week	1	4	2	10
number of reserve blocks/week	1	3	2	10
number of days off/week	1	2	2	10
flight time/week	(*)	(*)	20	2
service time/week	(*)	(*)	38	2
overnight duties/month	0	4	(*)	2
number of pairings/month	8	16	12	10
number of reserve blocks/month	4	12	8	10
number of days off/month	4	8	6	10
flight time/month	(*)	(*)	(*)	2
service time/month	(*)	(*)	(*)	2
number of consecutive working days	4	8	-	-
number of consecutive days off	1	2	-	-
number of consecutive morning duties	1	2	-	-
number of consecutive night duties	1	2	-	-
number of consecutive overnight duties	0	1	-	-

(-: not applicable; (*): pre-calculated upon the characteristics of the problem instance)

The crew rosters for freelance personnel are restricted by the same maximum parameter values for the distinct constraints whereas all minimum value constraints are removed.

4.1.2 Meta-heuristic parameter settings

Inspired by the ideas of Kolisch and Hartmann (2006), we use the number of created schedules as a stop criterion for our meta-heuristic procedure. For the computational experiments we apply a stop criterion of 20,000 schedules. Within the limits of this stop criterion, we need to carefully design our algorithm such that we find the right balance between diversification and intensification. Hence, we combined the different local search methods discussed in section 3.5 as follows, i.e., first, the single crew roster-based local search optimizes randomly 30% of the crew members' rosters. Next, the activity-based local search and the period-based local search, which optimizes 50% of the identified periods, are applied. Last, the schedule-based local search is carried out. This improvement method re-distributes the complete crew rosters, followed by 70% of all possible three- and two-days combinations. The improvement method is iterated over each single improvement method and over all improvement methods until no improvement can be found. Furthermore, the total population size amounts to 50 containing 30 population elements in $Refset_1$ and 20 solutions in $Refset_2$. Mutation is applied on 30% of all activities and is implemented as the activity-based local search incorporating random costs.

4.2 Structural analysis of the scatter search procedure

Table 2 displays the effect of different meta-heuristic principles implemented in our procedure. More precisely, different strategies and operation modes are implemented for the diversification generation method, the subset generation method, the solution combination method, the mutation operator, the improvement method and the reference set update method. For each strategy we display the solution quality, the percentage deviation from the best performing procedure, the CPU time (in seconds) and its ranking for each step in the procedure. In order to test the effect of the different strategies we start from the best performing heuristic procedure and implement or leave out a certain strategy or characteristic of the procedure. In this way, we can analyze unambiguously the impact of the different strategies in terms of solution quality.

Table 2. The performance of different meta-heuristic strategies

	Solution Quality	% Deviation	CPU time (s)	Ranking
Diversification generation method				
Completely random generation	79,803.20	1.69%	23.74	2
Completely heuristic generation	80,939.09	3.13%	23.92	3
Combined random and heuristic generation	78,480.67	0.00%	25.44	1
Subset generation method				
Completely random selection	79,984.35	1.92%	24.29	3
Controlled random selection	79,290.86	1.03%	24.46	2
Controlled selection based on fitness/diversity value	78,480.67	0.00%	25.44	1
Solution Combination Method				
One point crossover	79,997.84	1.93%	23.50	7
Best tournament selection (social quality)	79,949.21	1.87%	23.30	6
Best tournament selection (overall solution quality)	79,915.44	1.83%	23.62	4
Random tournament selection	79,860.72	1.76%	24.10	3
Path relinking	79,941.27	1.86%	31.18	5
Restricted path relinking	78,886.45	0.52%	23.54	2
Hybrid solution combination method	78,480.67	0.00%	25.44	1
Improvement method				
No improvement method	131,979.70	68.17%	27.48	6
No single crew roster-based local search	80,985.19	3.19%	15.33	4
No period-based local search	80,751.80	2.89%	31.46	3
No activity-based local search	79,691.01	1.54%	24.16	2
No schedule-based local search	84,515.52	7.69%	31.26	5
All local search mechanisms	78,480.67	0.00%	25.44	1
Non-iterated local search	82,847.10	5.56%	23.45	4
Iterated local search per local search mechanism	80,081.54	2.04%	29.37	2
Iterated local search over local search mechanisms	80,487.00	2.56%	25.98	3
Combined Iterated local search	78,480.67	0.00%	25.44	1
Mutation				
With mutation	78,480.67	0.00%	25.44	1
Without mutation	79,507.67	1.31%	22.05	2
Update method				
Static update	79,893.11	1.80%	24.13	2
Dynamic update	78,480.67	0.00%	25.44	1

For the *diversification generation method* we observe that a more diverse generation of initial population elements outperforms the procedure initializing solely heuristic population elements ('Completely heuristic generation'). This beneficial effect of diversity can be found back not only in the random generation of some population elements but also in the diverse manner solutions are generated (i.e., random and heuristic generation).

The results for the *subset generation method* indicate that the higher the intelligence in the selection system of population elements, the better the computational performance. The strategy 'Controlled selection based on fitness/diversity value' which controls the choice of type of subset (leading to diversification and/or intensification) and selects population elements based on their fitness/diversity value outperforms both other strategies ('Completely random selection' and 'Controlled random selection'). The probabilities for the types of subsets, i.e., $Refset_1 \times Refset_1$, $Refset_1 \times Refset_2$, or $Refset_2 \times Refset_2$, were defined as 45%, 45%, and 10% respectively.

The *solution combination method* shows that the one-point crossover operator is outperformed by the more disruptive crossover operator (i.e., the crossover operator based on random tournament selection) and the combination mechanisms incorporating a higher degree of problem-specific information (i.e., the path relinking approaches and the crossover operator based on best tournament selection). The best tournament selection methods indicate a better performance when incorporating all objective function information than focusing only on the social quality of the individual crew rosters. The restricted path relinking approach outperforms all other algorithms exploring only one combination method. However, hybridizing different solution combination methods yield even better results. In the algorithm 'Hybrid solution combination method', we apply the three best combination methods (i.e., the restricted path relinking, the random tournament selection and the best tournament selection) with a certain probability (i.e., respectively 50%, 25%, and 25%) in order to construct a new solution point.

In analyzing the *improvement method* we examine the role of each local search mechanism and the role of iterating the improvement method.

The beneficial effect of the local search methods can be clearly discerned when leaving the improvement method out of the algorithm ('No improvement method'). In order to establish the contribution of the different local search methods, we omitted each improvement method leading to four different versions of the algorithm, i.e., 'No single crew roster-based local search', 'No period-based local search', 'No activity-based local search', and 'No schedule-based local search'). Excluding the schedule-based local search leads to the worst results in terms of solution quality. This implies that this singular local search has the highest contribution in the improvement method when improving a new solution point. The contribution of the period-based and the single crew roster-based local search are fairly similar. The activity-based local search, however, has the smallest contribution probably due to the small size of its neighbourhood.

Furthermore, we tested how the improvement method is best iterated, i.e., no iterations are applied (‘Non-iterated local search’), each local search method is separately iterated until no improvement is made (‘Iterated local search per local search mechanism’), all local search methods are iterated until no improvement is made (‘Iterated local search over local search mechanisms’), and the combination of the latter two (‘Combined iterated local search’). The table reveals the beneficial effect of iterating the improvement method.

Incorporating the *mutation* operator in the algorithm leads to a better performance in terms of solution quality.

The table reveals that *updating the reference set* dynamically (‘Dynamic update’) achieves better results then the static update method (‘Static update’) in terms of solution quality.

4.3 Comparison of exact and heuristic procedures

In this section we compare the performance of the proposed scatter search procedure with variable neighbourhood search using 20,000 evaluated schedules as a stop criterion and with the branch-and-price procedure of Gamache et al. (1999) truncated at 600 seconds. The latter solution method is based on mathematical programming solving the linear relaxation of the generalized set partitioning problem by column generation and many NP-hard shortest path subproblems, both tailored to the specific problem formulation as described in section 2.2. Moreover, we have enhanced the column generation procedure by incorporating Lagrangean dual pruning in order to alleviate the “tailing-off effect” to terminate the column generation method sooner (Vanderbeck and Wolsey, 1996; Barnhart et al., 1998; Van den Akker et al., 2002). If the resulting linear programming relaxation is fractional, 0-1 branching is applied to allow or prohibit the assignment of a specific activity to a particular crew member. The results are displayed in table 3.

In the upper part of table 3 the computational results of the scatter search (‘SS’), the variable neighbourhood search (‘VNS’) and the branch-and-price procedure (‘B&P’) in terms of solution quality and CPU time are displayed. In order to give insight in which dimension of the objective function the scatter search performs better than the two other procedures, we display the three different objectives of the objective function, i.e., crew preferences (‘Crew Preferences’), fairness or impartiality among the crew members (‘Fairness’), and the operational costs due to understaffed activities and the employment of extra or freelance personnel (‘Operational costs’). Furthermore, we indicate the percentage both procedures deviate from the scatter search procedure in terms of overall solution quality.

In the middle and bottom part of table 3, a pairwise comparison is made of the scatter search with the variable neighbourhood search and the branch-and-price procedure respectively. We report how the scatter search performs worse (‘SS < VNS’ and ‘SS < B&P’), equal (‘SS = VNS’ and ‘SS = B&P’), and better (‘SS > VNS’ and ‘SS > B&P’) using two performance parameters, i.e., the percentage of problem instances (‘% Problem Instances’) and the deviation between the scatter search procedure and the respective procedure in terms of solution quality (‘% Deviation’).

Table 3. The performance comparison of different procedures with the scatter search methodology

Overall			
	SS	VNS	B&P
Solution Quality	78,480.67	80,402.41	102,701.42
- Crew Preferences	1,830.77	1,816.99	1,882.48
- Fairness	709.90	703.92	808.94
- Operational Costs	75,940.00	77,881.50	100,010.00
CPU time (s)	25.44	20.52	510.27
% Deviation	0.00%	2.45%	30.86%
SS vs VNS			
	SS < VNS	SS = VNS	SS > VNS
% Problem Instances	26.50%	1.00%	72.50%
% Deviation	0.43%	0.00%	5.69%
SS vs B&P			
	SS < B&P	SS = B&P	SS > B&P
% Problem Instances	33.50%	6.50%	60.00%
% Deviation	0.83%	0.00%	88.88%

The upper part of the table reveals that the scatter search outperforms the variable neighbourhood search and the branch-and-price procedure by 2.45% and 30.86% respectively. Within the stop criterion of 600 seconds, the branch-and-price procedure is able to solve 40% of the problem instances to optimality. The other problem instances obtain an integer solution either during the exploitation of the search tree, either due to the heuristic tree search described in Gamache et al. (1999), or due to the single pass heuristic exploited by Moore et al. (1978) and Byrne (1988) initializing the column generation procedure. Examining the different components of the objective function, we observe that the branch-and-price procedure performs worse for all three objective function dimensions compared with the other two procedures. When comparing the scatter search with the variable neighbourhood search, the table reveals that the scatter search obtains solutions with much lower operational costs whereas the crew members’ preferences and fairness among the crew members are slightly higher. This lower social quality is due to the objective function structure and the defined penalty costs enabling the operational costs as the dimension with the relative highest potential for improvement.

Despite the fact that the variable neighbourhood search is implemented with the same stop criterion, the scatter search outperforms the variable neighbourhood search for 72.5% of the problem instances and by 5.69% in terms of solution quality. This indicates the beneficial effect of implementing problem specific solution combination methods over generating a new solution point totally ad random. For the problem instances the scatter search is performing worse, the deviation in solution quality is rather small, i.e., 0.43%.

Furthermore, the table reveals that the scatter search outperforms the branch-and-price procedure for 60% of the problem instances and by 88.88% in terms of solution quality. This huge deviation is partly due to the inability of solving the linear programming relaxation within the stop criterion of 600 seconds for some of these problem files and, hence, no integer solution can be calculated using the partial branch-and-bound tree exploration. The obtained solutions result then out of the single pass heuristic. The branch-and-price procedure achieves better results in terms of solution quality for 33.5% of the problem instances. Despite this non-negligible amount of problem instances, the deviation in solution quality our scatter search procedure is performing worse is rather small, i.e., 0.83%.

Based on these computational comparisons (in terms of solution quality and CPU time), we can conclude the scatter search surpasses the implemented variable neighbourhood search and the branch-and-price procedure.

5 Conclusion

In this paper, a scatter search procedure has been proposed for an airline crew rostering problem which rosters regular crew members and decides if extra and/or freelance personnel should be employed. Moreover, the procedure determines the rosters of the engaged extra and freelance workforce minimizing the operational costs. The crews' lines-of-work are constructed based on a personalized rostering approach ensuring fairness between the regular crew members taking the crew members' preferences into account. We have discussed and investigated all aspects leading to a successful and effective procedure able to solve real-world crew rostering problems. Moreover, we have compared the performance of our algorithm with a variable neighbourhood search and an exact branch-and-price procedure showing the efficient performance of the proposed algorithm.

For future research, the integration of crew pairing and crew rostering into one problem would turn out beneficial over the decomposition approach. This integration would provide a higher flexibility making it easier to maximize the social quality of the schedule and minimize the operational costs. The further development of high-quality meta-heuristic approaches will become an important factor in this step since increasing the problem size, flexibility and complexity will be harder to solve for exact procedures. Furthermore, the construction of robust schedules will become a main issue in future research. Generally, schedules need to be maintained and repaired until the moments of operations resulting in delays and cancellations which lead to dissatisfied customers, higher operational costs, and a lower social quality for the crew members. Procedures need to be developed which can uptake this uncertainty and variability upfront leaving out the costly ad hoc adjustments.

Acknowledgements

We acknowledge the support given to this project by the 'Bijzonder Onderzoeksfonds (BOF), Ghent University, Belgium' under the contract number B/04668/01.

References

Aickelin, U., and Dowsland, K.A., 2000, "Exploiting problem structure in a genetic algorithm approach to a nurse rostering problem", Journal of Scheduling, 3, 139-153.

Anbil, R., Gelman, E., Patty, B., and Tanga, R., 1991, "Recent Advances in Crew-Pairing Optimization at American Airlines", Interfaces, 21, 62-74.

Anbil., R., Barnhart, C., Hatay, L., Johnson, E.L., and Ramakrishnan, V.S., 1992, "Crew-Pairing Optimization at American Airlines Decision Technologies", in Ciriani, T.A., and Leachman, R.C., (eds.), Optimization in Industrial Environments, Wiley, 7-24.

Ahuja, R., Egun, O., Orlin J., and Punnen, A., 2002, "A survey of very large-scale neighbourhood search techniques", Discrete Applied Mathematics, 123, 75-102.

Andersson, E., Housos, E., Kohl, N., and Wedelin, D., 1998, "Crew Pairing Optimization", in Yu, G., (ed.), Operations Research in the Airline Industry, Kluwer, Norwell, 228-258.

Barnhart, C., Johnson, E.L., Anbil, R., and Hatay, L., 1994, "A Column-Generation Technique for the Longhaul Crew-Assignment Problem", in Ciriani, T.A., and Leachman, R.C., (eds.), Optimization in Industrial Environments, Wiley, New York, 7-24.

Barnhart, C., Hatay, L., and Johnson, E.L., 1995, "Deadhead Selection for the Long-Haul Crew Pairing Problem", Operations Research, 43, 491-495.

Barnhart, C., and Shenoir, R.G., 1996, "A Alternate Model and Solution Approach for the Longhaul Crew Pairing Problem", Technical Report, MIT.

Barnhart, C., Johnson, E.L., Nemhauser, G.L., Savelsbergh, M.W., and Vance, P.H., 1998, “Branch-and-Price: Column Generation for solving huge Integer Programs”, Operations Research, 46, 316-329.

Barutt, J., and Hull, T., 1990, “Airline Crew Scheduling – Supercomputers and Algorithms”, SIAM News, 23, 20-22.

Beasley, J., and Chu, P., 1996, “A genetic algorithm for the set covering problem”, European Journal of Operational Research, 94, 392-404.

Buhr, J., 1978, “Four methods for monthly crew assignment – a comparison of efficiency”, 1978 AGIFORS Symposium Proceedings, 18, 403-430.

Burke, E.K., Cowling, P., De Causmacker, P., and Vanden Berghe, G., 2001, “A memetic approach to the nurse rostering problem”, Applied Intelligence, 3, 199-214.

Byrne, J., 1988, “A Preferential Bidding System for Technical Aircrew”, AGIFORS Symposium Proceeding, 28, 87-99.

Campbell, K.W., Durfee, R.B., and Hines, G.S., 1997, “FedEx Generates Bid Lines Using Simulated Annealing”, Interfaces, 27, 1-16.

Cappanera, P., and Gallo, G., 2004, “A Multicommodity Flow Approach to the Crew Rostering Problem”, Operations Research, 52, 583-596.

Caprara, A., Toth, P., and Vigo D., 1998a, “Modeling and Solving the Crew Rostering Problem”, Operations Research, 46, 820-830.

Caprara, A., Focacci, F., Lamma, E., Mello, P., Milano, M., Toth, P., and Vigo, D., 1998b, “Integrating Constraint Logic Programming and Operations Research Techniques for the Crew Rostering Problem”, Software – Practice and Experience, 28, 49-76.

Caprara, A., Monaci, M., and Toth, P., 2003, “Models and algorithms for a staff scheduling problem”, Mathematical Programming – Series B, 98, 445-476.

Christou, I.T., Zakarian, A., Liu, J., and Carter, H., 1999, “A Two-Phase Genetic Algorithm for Large-Scale Bidline-Generation Problems at Delta Airlines”, Interfaces, 29, 51-65.

Desaulniers, G., Desrosiers, J., Dumas, Y., Marc, S., Rioux, B., Solomon, M.M., and Soumis, F., 1997, “Crew Pairing at Air France”, European Journal of Operational Research, 97, 245-259.

Desrochers, M., and Soumis, F., 1989, “A Column Generation Approach to the Urban Transit Crew Scheduling Problem”, Transportation Science, 23, 1-13.

Desrosiers, J., Dumas, Y., Solomon, M., and Soumis, F., 1995, “Time Constrained Routing and Scheduling”, Handbooks in Operations Research and Management Science, Network Science, 8, 35-139.

Dias, T.M., Ferber, D.F., de Souza, C.C., and Moura, A.V., 2003, “Constructing nurse schedules at large hospitals”, International Transactions in Operational Research, 10, 245-265.

El Moudani, W., Cosenza, C.A.N., de Coligny, M., and Mora-Camino, F., 2001, "A Bi-Criterion Approach for the Airlines Crew Rostering Problem", Lecture Notes in Computer Science, Proceedings of the First International Conference on Evolutionary Multi-Criterion Optimization, 1993, 486-500.

Ernst, A.T., Jiang, H., Krishnamoorthy, M., and Sier, D., 2004, "Staff scheduling and rostering: A review of applications, methods and models", European Journal of Operational Research, 153, 3-27.

Fahle, T., Junker, U., Karisch, S.E., Kohl, N., Sellmann, M., and Vaaben, B., 2002, "Constraint Programming Based Column Generation for Crew Assignment", Journal of Heuristics, 8, 59-81.

Freling, R., Lentink, R., and Wagelmans, A., 2001, "A Decision Support System for Crew Planning in Passenger Transportation Using a Flexible Branch-and-Price Algorithm", In Voss, S., and Daduna, J.R., (eds.), Computer-Aided Scheduling of Public Transport, Lecture Notes in Economics and Mathematical Systems, 505, 73-90.

Gamache, M., and Soumis, F., 1998, "A Method for Optimally Solving the Rostering Problem," in Yu, G., (ed.), Operations Research in the Airline Industry, Kluwer, Boston, 124-157.

Gamache, M., Soumis, F., Villeneuve, D., Desrosiers, J., and Gélinas, E., 1998, "The Preferential Bidding System at Air Canada", Transportation Science, 32, 246-255.

Gamache, M., Soumis, F., Marquis, G., and Desrosiers, J., 1999, "A Column Generation Approach for large-scale Aircrew Rostering Problems", Operations Research, 47, 247-263.

Gamache, M., Hertz, A., Ouellet, J.O., 2007, A graph coloring model for a feasibility problem in monthly crew scheduling with preferential bidding", Computers & Operations Research, 34, 2384-2395.

Gershkoff, I., 1989, "Optimizing Flight Crew Schedules", Interfaces, 19, 24-43.

Giaferri, C., Hamon, J.P., and Lengline, J.G., 1982, "Automatic Monthly Assignment of Medium-Haul Cabin Crew", AGIFORS Symposium Proceeding, 22, 69-95.

Glover, F., 1998, "A Template for Scatter Search and Path Relinking", Lecture Notes in Computer Science, 1363, 13-54.

Glover, F., and Laguna, M., 2000, "Fundamentals of Scatter Search and Path Relinking", Control and Cybernetics, 3, 653-684.

Graves, G.W., McBride, R.D., Gershkoff, I., Anderson, D., and Mahidhara, D., 1993, "Flight Crew Scheduling", Management Science, 39, 736-745.

Hansen, P., and Mladenović, N., 2001, "Variable neighbourhood search: Principles and applications", European Journal of Operational Research, 130, 449-467.

Hjorring, C.A., and Hansen, J., 1999, Column Generation with a Rule Modelling Language for Airline Crew Pairing", In Proceedings of the 34th Annual Conference of the Operational Research Society of New Zealand, Hamilton, New Zealand, 133-142.

Hoffman, K.L., and Padberg, M., 1993, "Solving airline crew scheduling problems by branch-and-cut", Management Science, 39, 657-682.

Housos, E., and Elmroth, T., 1997, "Automatic Optimization of Subproblems in Scheduling Airline Crews", Interfaces, 27, 68-77.

Ioachim, I., Desrosiers, J., Dumas, Y., Solomon, M.M., and Villeneuve, D., 1995, "A request clustering algorithm for door-to-door handicapped transportation", Transportation Science, 29, 63-79.

Irnich, S., and Desaulniers, G., 2004, "Shortest Path Problems with Resource Constraints", Les Cahiers du Gérad, G-2004-11.

Jarrah, A.I.Z., and Diamond, J.T., 1997, "The Problem of Generating Crew Bidlines", Interfaces, 27, 49-64.

Kapsalis, A., Rayward-Smith, V.J., and Smith, G.D., 1993, "Solving the graphical Steiner tree problem using genetic algorithms", Journal of the Operational Research Society, 44, 397-406.

Kerati, S., El Moudani, W., de Coligny, M., and Mora-Camino, F., 2002, "A Heuristic Genetic Algorithm for the Airline Crew Scheduling Problem", EU/ME, the European Chapter on Meta-Heuristics, EURO working group.

Kohl, N., and Karisch, S.E., 2000, "Integrating Operations Research and Constraint Programming Techniques in Crew Scheduling", In Proceedings of the 40th Annual AGIFORS Symposium, Istanbul, Turkey.

Kohl, N., and Karisch, S.E., 2004, "Airline Crew Rostering: Problem Types, Modeling and Optimization", Annals of Operations Research, 127, 223-257.

Kolisch, R., and Hartmann, H., 2006, "Experimental investigation of heuristics for resource-constrained project scheduling: An update", European Journal of Operational Research, 174, 23-37.

Kuhn, H., 1955, "The Hungarian method for the assignment problem", Naval Research Logistics, 2, 83-97

Lasry, A.D., McInnis, D., Soumis, F., Desrosiers, J., Solomon, M.M., 2000, "Air Transat Uses ALTITUDE to Manage its Aircraft Routing, Crew Pairing and Work Assignment", Interfaces, 30, 41-53.

Lavoie, S., Minoux, M., and Odier, E., 1998, "A New Approach for Crew Pairing Problems by Column Generation with an Application to Air Transportation", European Journal of Operational Research, 35, 45-58.

Levine, D., 1996, "Application of a Hybrid Genetic Algorithm to Airline Crew Scheduling", Computers & Operations Research, 23, 547-558.

Maenhout, B., and Vanhoucke, M., 2005, "An Electromagnetism meta-heuristic for the nurse scheduling problem", to appear in Journal of Heuristics

Makri, A., and Klabjan, D., 2004, "A New Pricing Scheme for Airline Crew Scheduling", INFORMS Journal on Computing, 16, 56-67.

Marti, R., Laguna, M., and Glover, F., 2006, "Principles of scatter search", European Journal of Operational Research, 169, 359-372.

Medard, C.P., and Sawhney, N., 2004, "Airline crew scheduling: From planning to operations", Carmen Systems AB, Göteborg, Sweden, Carmen Research and Technology Report CRTR-0406.

Moore, R., Evans, J., and Noo, H., 1978, "Computerized tailored blocking", AGIFORS Symposium Proceedings, 18, 343-361.

Nicoletti, B., 1975, "Automatic Crew Rostering", Transportation Science, 9, 33-42.

Reeves, C.R., 1995, "A Genetic Algorithm for Flowshop Sequencing", Computers & Operations Research, 22, 5-13.

Ryan, D.M., 1992, "The solution of massive generalized set partitioning problems in air crew rostering", Journal of the Operational Research Society, 43, 459-467.

Sarra, D., 1988, "The automatic assignment model", 1988 AGIFORS Symposium Proceedings, 28, 23-37.

Sellmann, M., Zevoudakes, K., Stamatopoulos, P., and Fahle, T., 2002, "Crew Assignment via Constraint Programming: Integrating Column Generation via Heuristic Tree Search", Annals of Operations Research, 115, 207-225.

Thiel, M.P., 2004, "Team-oriented Airline Crew Rostering for Cockpit Personnel", Proceedings of the 9th International Conference on Computer-Aided Scheduling of Public Transport (CASPT 2004), 1-26.

Tingley, G.A., 1979, "Still another solution method for the monthly aircrew assignment problem", 1979 AGIFORS Symposium Proceedings, 19, 143-203.

Vance, P.H., Barnhart, C., Johnson, E.L., and Nemhauser, G.L., 1997, "Airline Crew Scheduling: A New Formulation and Decomposition Algorithm", Operations Research, 45, 188-200.

Van den Akker, M., Hoogeveen, H., and Van de Velde, S., 2002, "Combining Column Generation and Lagrangean Relaxation to Solve a Single-Machine Common Due Date Problem", Informs Journal on Computing, 14, 37-51.

Vanderbeck, F., and Wolsey, L.A., 1996, "An exact algorithm for IP column generation", Operations Research Letters, 19, 151-159.

Yu, G., 1998, Operations Research in Airline Industry, Kluwer Academic Publishers, Dordrecht, 1998.