New Heuristics to Solve the "CSOP" Railway Timetabling Problem

L. Ingolotti[1], A. Lova[2], F. Barber[1], P. Tormos[2], M. A. Salido[1], and M. Abril[1]

[1] DSIC, Polytechnic University of Valencia, Spain
{lingolotti, fbarber, msalido, mabril}@dsic.upv.es
[2] DEIOAC, Polytechnic University of Valencia, Spain
{allova, ptormos}@eio.upv.es

Abstract. The efficient use of infrastructures is a hard requirement for railway companies. Thus, the scheduling of trains should aim toward optimality, which is an NP-hard problem. The paper presents a friendly and flexible computer-based decision support system for railway timetabling. It implements an efficient method, based on meta-heuristic techniques, which provides railway timetables that satisfy a realistic set of constraints and, that optimize a multi-criteria objective function.

Key Words: Constraint Satisfaction, Decision Support, Planning and Scheduling

1 Introduction

The main motivations for this work are the need of obtaining an automatic timetabling process for railway companies, the hard requirement of using efficiently railway infrastructures and the challenge that this process implies for the application and research of techniques in the Artificial Intelligence field. The literature of the 1960s, 1970s, and 1980s relating to rail optimization was relatively limited. Compared to the airline and bus industries, optimization was generally overlooked in favor of simulation or heuristic-based methods. However, Cordeau et al. [1] point out greater competition, privatization, deregulation, and increasing computer speed as reasons for the more prevalent use of optimization techniques in the railway industry. Our review of the methods and models that have been published indicates that the majority of authors use models that are based on the Periodic Event Scheduling Problem (PESP) introduced by Serafini and Ukovich [7]. The PESP considers the problem of scheduling as a set of periodically recurring events under periodic time-window constraints. The model generates disjunctive constraints that may cause the exponential growth of the computational complexity of the problem depending on its size.

Schrijver and Steenbeek [5] have developed `CADANS`, a constraint programming-based algorithm to find a feasible timetable for a set of PESP constraints. The scenario considered by this tool is different from the scenario that we used; therefore, the results are not easily comparable. Nachtigall and Voget [4] also use PESP constraints to model the cyclic behavior of timetables and to consider the minimization of passenger waiting times as the objective function. Their solving procedure starts with a solution that is obtained in a way similar to the one that timetable designers in railway companies use. This initial timetable is then improved by using a genetic algorithm. In our problem, the waiting time for connections is not taken into account because we only consider the timetabling optimization for a single railway line.

The train scheduling problem can also be modeled as a special case of the job-shop scheduling problem (Silva de Oliveira [8], Walker et al. [11]), where train trips are considered *jobs* that are scheduled on tracks that are regarded as *resources*. The majority of these works consider the scheduling of new trains on an empty network. However, infrastructure management railway companies usually also require the optimization of new trains on a line where many trains are already in circulation (that is, trains that have a fixed timetable). With this main objective, Lova et al. [2] propose a scheduling method based on *reference stations* where the priority of trains, in the case of conflict, changes from one iteration to another during the solving process.

2 Problem Specification

A railway line is composed by an ordered sequence of locations $\{l_0, l_1, ..., l_\mathrm{n}\}$ which are linked by single or double track. The problem consists in scheduling a set of new trains ($\mathrm{T_{new}}$), taking into account that the railway line may be occupied by other trains in circulation ($\mathrm{T_C}$) whose timetables cannot be modified. The trains may belong to different operator types and their respective journeys may also be different from each other.

The timetable assigned to each new train must be feasible, i.e. it must fulfill the set of constraints defined in Section 2.1. Besides feasibility, two additional goals are pursued in this work: computational efficiency and optimality, which is measured according to the objective function defined in Section 2.2.

2.1 Feasibility of a Solution - Set of Constraints: CONS

The problem of obtaining a feasible and optimal railway timetabling can be defined as a Constraint Satisfaction and Optimization Problem (CSOP). Considering $\mathrm{T} = \mathrm{T_{new}} \cup \mathrm{T_C}$, variables arr_i^t and dep_i^t represent the arrival/departure times of each train $t \in \mathrm{T}$ from/to each station l_i of its journey. For each train $t \in \mathrm{T_C}$, these variables are already instantiated, so that it is not necessary consider any constraint among them. A feasible railway timetable must satisfy a set of constraints `CONS`, which are defined by the Spanish Manager of Railway Infrastructure (ADIF).

Set of Constraints[3]

- *Interval for the Initial Departure/Arrival Time* : given the intervals $[\mathrm{I}^t_{\mathrm{L}}, \mathrm{I}^t_{\mathrm{U}}]$ and $[\mathrm{F}^t_{\mathrm{L}}, \mathrm{F}^t_{\mathrm{U}}]$, each train $t \in \mathrm{T}_{\mathrm{new}}$ should leave/arrive its initial/final station $l^t_0/l^t_{n_t}$ at a time $dep^t_0/arr^t_{n_t}$ such that,

$$\mathrm{I}^t_{\mathrm{L}} \leq dep^t_0 \leq \mathrm{I}^t_{\mathrm{U}} \quad . \tag{1}$$

$$\mathrm{F}^t_{\mathrm{L}} \leq arr^t_{n_t} \leq \mathrm{F}^t_{\mathrm{U}} \quad . \tag{2}$$

- *Running Time*: for each track section $l^t_i \rightarrow l^t_{i+1}$, and for each train $t \in \mathrm{T}_{\mathrm{new}}$ is given a running time $\Delta^t_{i\rightarrow(i+1)}$ such that

$$arr^t_{i+1} = dep^t_i + \Delta^t_{i\rightarrow(i+1)} \quad . \tag{3}$$

- *Commercial Stop*: each train $t \in \mathrm{T}_{\mathrm{new}}$ is required to remain in a station l^t_i of its journey at least C^t_i time units

$$dep^t_i \geq arr^t_i + \mathrm{C}^t_i \quad . \tag{4}$$

- *Headway Time*: if two trains, $\{t_i, t_j\} \subseteq \mathrm{T}$, traveling in the same direction leave the same location l^t_k towards the location l^t_{k+1}, they are required to have a difference in their departure times of at least φ^{d}_k and a difference in their arrival times of at least φ^{a}_k. Particularly, when the blocking type in the track section is *Automatic* then $\varphi^{\mathrm{d}}_k = \varphi^{\mathrm{a}}_k$ and their values are determined by the user. When the blocking type of the track section is *Manual*, the second train must wait in the station until the first train arrives to the next open station.

$$|dep^{t_i}_k - dep^{t_j}_k| \geq \varphi^{\mathrm{d}}_k \quad . \tag{5}$$

$$|arr^{t_i}_{k+1} - arr^{t_j}_{k+1}| \geq \varphi^{\mathrm{a}}_k \quad . \tag{6}$$

- *Crossing*: A single-track section cannot be occupied at the same time by two trains going in the opposite directions. Considering that: T_{D} and T_{U} are the set of trains that travel in *down* and *up* direction respectively, $t \in \mathrm{T}_{\mathrm{D}}$, $t' \in \mathrm{T}_{\mathrm{U}}$ and $i \rightarrow j$ is a track section (down direction), this constraint is modeled by the following expression.

$$dep^{t'}_j > arr^t_j \vee dep^t_i > arr^t_i \quad . \tag{7}$$

- *Reception/Expedition Time*: Crossing operations are performed in stations and they require to keep temporal safety margins (Reception and Expedition Time) between the involved trains. The difference between the arrival times/departure and arrival times of any two trains, $\{t', t\} \subseteq \mathrm{T} \wedge \{t', t\} \nsubseteq \mathrm{T}_{\mathrm{D}} \wedge \{t', t\} \nsubseteq \mathrm{T}_{\mathrm{U}}$, in a same station l is defined by the expressions below,

[3] In this work are described the main problem constraints, the rest of constraints that have been considered are detailed in [10].

where R_t is the reception time specified for the train that arrives to l first and E_t is the expedition time specified for t.

$$arr_l^{t'} \geq arr_l^t \rightarrow arr_l^{t'} - arr_l^t \geq R_t \ . \tag{8}$$

$$|dep_l^{t'} - arr_l^t| \geq E_t \ . \tag{9}$$

- *Overtaking on the track section*: overtaking must be avoided between any two trains, $\{i, j\} \nsubseteq T_C \wedge (\{i, j\} \subseteq T_D \vee \{i, j\} \subseteq T_U)$, on any common track sections, $k \rightarrow (k+1)$, of their journeys:

$$(arr_{k+1}^i \geq arr_{k+1}^j) \leftrightarrow (dep_k^i \geq dep_k^j) \ . \tag{10}$$

- *Frequency*: The user can specify a given frequency among a subset of trains of T_{new}. Each subset of trains G must be composed by trains that have a same journey and that travel in a same direction. For each group, a different frequency may be specified, which must belong to an interval $[F_L^G, F_U^G]$, where the lower and upper bound may be the same. In this case a fixed frequency is specified for the group. For each station where a given frequency must be considered, the following expression must be satisfied

$$\{t', t\} \subset G \wedge dep_i^{t'} > dep_i^t \wedge (\nexists t")_G (dep_i^t < dep_i^{t"} < dep_i^{t'})$$
$$\rightarrow F_L^G \leq dep_i^{t'} - dep_i^t \leq F_U^G \ . \tag{11}$$

2.2 Optimality of a Solution - Objective function

We have defined the *Minimum Total Running Time* (Γ_{opt}^t) of each train $t \in T_{new}$, as the minimum time required by t to complete its journey, satisfying all the problem constraints in `CONS` but only taking into account the set of trains in circulation T_C; the trains $T_{new} \setminus t$ are ignored. We have two criteria to measure the quality of each solution: (*i*) the average delay (δ) of the new trains with respect the *Minimum Total Running Time*, and (*ii*) the deviation (σ) between the average delay of trains going in up direction (δ_U), and the average delay of trains going in down direction (δ_D). These values are computed according the following expressions:

$$\delta_t = \frac{arr_{n_t}^t - dep_0^t - \Gamma_{opt}^t}{\Gamma_{opt}^t}; \ \delta_U = \frac{\sum_{t \in T_U \cap T_{new}} \delta_t}{|T_U \cap T_{new}|}; \ \delta_D = \frac{\sum_{t \in T_D \cap T_{new}} \delta_t}{|T_D \cap T_{new}|}; \ \delta = \frac{\delta_U + \delta_D}{|T_{new}|}$$

The average delay deviation of each set of trains, up and down direction, with respect to the average delay of all new trains, is:

$$\sigma = \sqrt{\frac{(\delta_U - \delta)^2 + (\delta_D - \delta)^2}{2}}$$

A given weight, ω_{delay} and ω_{eq}, is assigned to each criterion respectively. The weight assigned to each one is determined by the railway planner. Finally, considering that *TTABLE* is one problem solution and therefore is the set of timetables for each new train, the objective function of this problem is formulated as:

$$f(\mathit{TTABLE}) = \mathrm{MIN}(\omega_{delay} \times \delta + \omega_{eq} \times \sigma) \tag{12}$$

3 A Scheduling Order-Based Method (SOBM)

In order to obtain an optimal solution (timetables) in our CSOP, the process obtains successively different solutions, keeping the best one obtained each time. The search process of each solution is heuristically guided (Section 3.1), and a pruning process is applied to reduce the search time (Section 3.3). The process obtains new solutions iteratively until a given end condition is fulfilled (number of iterations, a given time interval, a minimum cost for the objective function, etc.). Each solution is obtained by repeating the following sequence of steps: select a train, assign a timetable to the selected train in a given track section of its journey and evaluate the partial solution to decide wether to continue or to prune, until a valid timetable had been assigned to every train in $\mathrm{T}_{\mathrm{new}}$; or until the current partial solution is discarded by the pruning process.

3.1 Heuristic Decision for the Priority Assignment

The disjunctive constraints are the main cause of the complexity of this problem. These constraints are due to the competition of two trains for the same resource, i.e. track section or track in a given station. When there is a conflict between two trains for the same resource, and one of them is a train in circulation, the train in circulation will always have higher priority, and the new train will always be delayed. However when the conflict occurs between two new trains, we use a heuristic based on the selection order of the trains to determine which of the two new trains will have higher priority on this resource. Each solution will be determined by this priorities assignment.

We consider the problem as a search tree whose root node (*initial node*) represents the empty timetabling. For each node where no successor is possible, there is an artificial terminal node (*final node*). Let $\mathrm{T}_{\mathrm{open}}$ be the set of new trains whose timetables have not yet been completed ($\mathrm{T}_{\mathrm{open}} \subseteq \mathrm{T}_{\mathrm{new}}$). Each intermediate node is composed of a pair (t_i, s_j), which indicates that a feasible timetable must be found for the train $t_i \in \mathrm{T}_{\mathrm{open}}$ in its track section s_j. When the timetable of a train t_i is completed, this train is eliminated from the set $\mathrm{T}_{\mathrm{open}}$. Each level of the search tree indicates which part of the timetable of each train can be generated. The method must determine in each level, which of the nodes will be chosen. The problem consists of finding a path in the search tree, (from the *initial* node to the *final* node), so that the order of priorities established by this path produces the minimum cost according to the objective function.

In this point we describe how is selected each node that will compose the path of one solution. For each node (t_i, s_j) in a given level of the tree, consider that a feasible timetable was assigned to t_i from its initial station l_0^t until the station $l_{j_t}^t$. For each level, we measure the partial delay of each train $t \in \mathrm{T}_{\mathrm{open}}$ according to the following expression:

$$\delta_{\mathrm{partial}}^t = \frac{arr_{j_t}^t - dep_0^t - \Gamma_{\mathrm{opt}}^t}{\Gamma_{\mathrm{opt}}^t} \qquad (13)$$

Given that the minimum partial delay is $\delta_{\min} = \min_{t \in \mathrm{T_{open}}}(\delta^t_{\text{partial}})$, the probability ρ_t, of train t being selected is computed according to the following expression:

$$\rho_t = \frac{(\delta^t_{\text{partial}} - \delta_{\min} + \varepsilon)^\alpha}{\sum_{t \in \mathrm{T_{open}}} (\delta^t_{\text{partial}} - \delta_{\min} + \varepsilon)^\alpha} \tag{14}$$

A node is chosen according to the parameterized Regret-Based Biased Random Sampling (RBRS) [6] and [9], so that the train with higher priority is not necessarily the train chosen, due to the random component of the RBRS method.

If the node (t_i, s_j) is selected, then the next step will consist in setting the timetable for the train t_i in its track section s_j which is described in the next subsection.

3.2 Timetable Generation and Constraint Verification

Once a given node (t_i, s_j) has been selected, its computed a departure time for the train t_i from the location l^t_j ($s_j = l^t_j \rightarrow l^t_{j+1}$) and an initial timetable is assigned in the track section s_j according to the constraints 3 and 4. From this initial timetable, the process proceeds to verify the rest of constraints in `CONS` in the track section s_j.

The constraints are verified taking into account the trains whose journeys include the track section under consideration. The new trains whose journeys include this track section but that have not yet been selected in s_j are not taken into account. Therefore, when the process indicates that train t_i has a conflict with another new train t' (violating constraints such as 7, 10, ... etc), it means that (t', s_j) has been selected before than (t_i, s_j). Thus, if the process detects a conflict between the train t_i and another train, it will delay to t_i, since the other train has been selected previously and therefore it has higher priority than t_i.

It important to remark that the order in which the new trains are selected influences the global scheduling. The order of selection determines the priority among the trains and the way that each conflict between two new trains (crossing, overtaking, capacity in stations, headway time,..., etc) will be solved.

3.3 Heuristic Decision for the Prune Process

Each time the process finishes building the timetable for a train t_i in a track section s_j, it estimates the objective function value corresponding to the current partial scheduling in the best of the cases (δ_{est}). In other words, when a train has not yet assigned a timetable from a given track section of its journey, the process estimates that this train will employ the minimum time possible, sum of running time plus commercial stop ($M^t_{i_t \rightarrow n_t}$), to go from this track section until its destination. The following expression shows how is computed δ_{est}.

$$\delta_{\text{est}} = \frac{\sum_{t \in \mathrm{T_{new}}} \delta^t_{\text{est}}}{|\mathrm{T_{new}}|} \tag{15}$$

$$\delta_{\text{est}}^t = \frac{arr_{i_t}^t - dep_0^t + M_{i_t \to n_t}^t - \Gamma_{\text{opt}}^t}{\Gamma_{\text{opt}}^t} \tag{16}$$

If the estimated cost for the current partial scheduling is greater than the best solution cost (obtained up to point), then the current iteration is aborted and the partial scheduling is discarded. The partial solution is not saved because a problem for a real environment would imply a very high spatial cost.

The method incorporates knowledge about the objective function of the problem in order to estimate when a solution might not be better than a previous one. It is a conservative technique because it does not risk a solution until it is sure that the solution will not be better than the best one obtained up to this point. We increase the method efficiency with this pruning process.

4 Results

The Spanish Manager of Railway Infrastructure (ADIF) provides us with real instances to obtain a realistic evaluation of the proposed heuristic. We describe ten problem instances in Figure 1.a (columns 2 to 9) by means of: the length of the railway line, number of single/double track sections, number of stations,number of trains and track sections (TS) corresponding to all these trains, for trains already in circulation and for new trains, respectively. The results are shown in

Problems	Infrastructure Despcription				In Circulation		New Trains	
	Km	1-Way	2-Way	Stat	Trains	TS	Trains	TS
1	209,1	25	11	22	40	472	53	543
2	129,4	21	0	15	27	302	30	296
3	177,8	37	4	25	11	103	11	146
4	225,8	33	0	23	113	1083	11	152
5	256,1	38	0	28	80	1049	15	235
6	256,1	38	0	28	81	1169	16	159
7	96,7	16	0	13	47	1397	16	180
8	96,7	16	0	13	22	661	40	462
9	298,2	46	0	24	26	330	11	173
10	401,4	37	1	24	0	0	35	499

(a)

Problems	RANDOM		SOBM	
	#of Solutions	Obj%	#of Solutions	Obj%
1	169	8,6	168	5,9
2	611	10,1	608	10
3	2185	21,1	3101	16
4	311	13,2	445	5,5
5	396	19,3	452	17,5
6	424	14,7	521	14,1
7	267	18	263	15,4
8	67	50,9	85	45,5
9	1112	11,5	1129	8,7
10	405	19,2	397	17,9

(b)

Fig. 1: Results obtained with the SOBM method

Figure 1.b. This Table presents the best value of the objective function and the number of feasible solutions that were obtained for each problem (columns 2 and 3 for the RANDOM approach; columns 4 and 5 for the SOBM approach). The algorithm is implemented using C++ running on a Pentium IV 3,6 Ghz. The running time was of 300" for all the problems, and the parameters of the RBRS set to α=1 and ε=0,05. The difference between the RANDOM approach and the SOBM approach is the way that the trains are chosen at each level of the

search tree. In the first approach the trains are chosen randomly, in the second approach, the trains are chosen according to the method explained in Section 3. The results show that a guided heuristic such as SOBM explores regions in a more promising way in less time than the RANDOM approach does and this leads to better solutions.

MOM: A Decision Support System

We have developed a tool, MOM (*Modulo Optimizador de Mallas*), that provides solutions for the timetabling problem whose requirements have been given in Section 2. MOM solves the problem using the SOBM method (Section 3). It is a tool that makes easier to railway planner the task of obtaining feasible and high quality timetables for the trains that must be added to a given railway line.

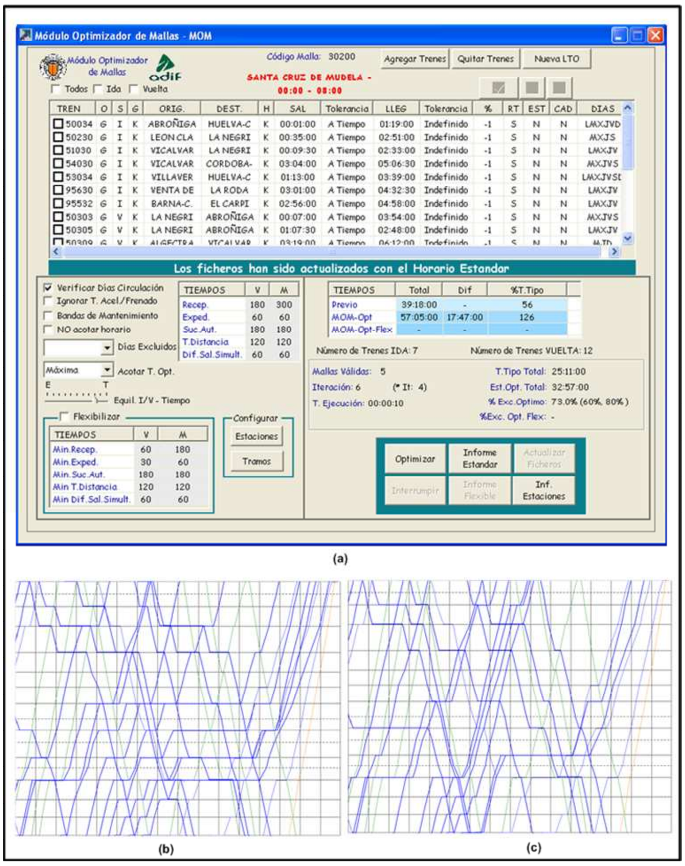

Fig. 2: A graphical example of solutions given by MOM

MOM allows the user to configure each problem instance as far as infrastructure, user requirements, solving process and quality of solutions is concerned.

- MOM interacts with the ADIF Database to obtain data related to the railway infrastructure (tracks in stations and sections, maintenance intervals, closing time, blocking type, etc.) and trains (journeys, commercial stops, etc. However, MOM allows the user modify the obtained data in order to provide the nearest scenario to every particular environment or user requirements, such that the given solutions are useful and significant. Moreover, MOM also allows to parameterize each train: intervals for departure and arrival times, running times in sections, frequency of departures, circulation days, specific safety margins (reception, expedition and headway time)in specific sections, etc.
- MOM deals with hard and soft (i.e. *flexible*) constraints. That is, some given constraints (reception time, expedition time, etc.) can be relaxed within given bounds in order to obtain a better global scheduling (improvement in the objective function value). Any case, every relaxation should be finally validated by the user.
- In order to compute the quality of each solution, the objective function given in Section 2.2 takes into account: (i) the average delay of the trains with respect to the minimum running time, and (ii) the balance between the average delay of trains going in up and down direction respectively. The user specifies to MOM the weight that must be assigned to each criterion in the objective function.

In Figure 2.a is shown the MOM interface. The Figures 2.b and 2.c show two solutions. In the first one, the solving process has satisfied all the constraints without relax the variables domain. In the second solution, the problem was solved using the flexible mode. Certain constraints were relaxed and the average delay was 20% less than the average delay obtained in the first solution.

Conclusions

In this paper, we present a Scheduling Order-Based Method (SOBM) that proposes an efficient and flexible heuristic-driven method for the Train Timetabling Problem, which is modeled as a CSOP. Heuristic-decisions are based on both the knowledge about the problem and the multi criteria objective function.

Several realistic instances of the problem have been verified as well as different traffic conditions and train configurations. The method can be applied to any railway line and does not require a specific configuration in the railway infrastructure. The set of constraints can be modified without affecting the solving process used during the optimization.

The method has been implemented as a friendly and flexible user-aid system so that the railway planner can obtain train timetables under several scenarios and user requirements, in an efficient way and with high quality. It allows the railway planner to employ more time in the analysis tasks and the decision making process that corresponds to the medium and long term planning. To our knowledge, there are no in-use computer-based systems that consider all the shown features. We consider the main contributions to be the efficiency and flexibility of the method taking into account its application to real and complex scenarios. *MOM* is currently being used by the Spanish Manager of Railway Infrastructure (ADIF).

Acknowledgments

This work has been supported by the Research projects TIN2004-06354-C02-01 (MEC, Spain - FEDER) and GV04B/516 (Generalitat Valenciana, Spain). We are grateful to the Spanish Manager of Railway Infrastructure (ADIF) and to Jose Estrada Guijarro for give us the opportunity to interact with a real environment, for provide us with the instances presented in this paper, and for his effort and time to define the problem and evaluate the resulting schedules.

References

1. Cordeau, J., Toth, P. and Vigo, D. A survey of optimization models for train routing and scheduling. Transportation Science 32, 380-446 (1998)
2. Lova, A., Tormos, P., Barber, F., Ingolotti, L., Salido, M.A. and Abril, M. Intelligent Train Scheduling on a High-Loaded Railway Network. ATMOS 2004: Algorithmic Methods and Models for Optimization of Railways. Lecture Notes in Computer Science, LNCS (Springer Verlag). To appear. 2005
3. *MOM: Mdulo Optimizador de Mallas: http://www.dsic.upv.es/users/ia/gps/MOM/*
4. Nachtigall, K. and Voget, S. Minimizing waiting times in integrated fixed interval timetables by upgrading railway tracks 103, 610–627, (1997)
5. Schrijver, A., and Steenbeek, A. Timetable construction for railned. Technical Report, CWI, Amsterdam, The Netherlands (in Dutch), (1994)
6. Schirmer, A. and Riesenberg, S. Parameterized heuristics for project scheduling-Biased random sampling methods. Technical Report 456. Institute fr Betriebswirtschaftslehre der UNIVERSITT KIEL (1997)
7. Serafini, P. Ukovich, W. A mathematical model for periodic scheduling problems. SIAM Journal on Discrete Mathematics 2(4), 550–581, (1989)
8. Silva de Oliveira, E. Solving Single-Track Railway Scheduling Problem Using Constraint Programming. Phd Thesis. Univ. of Leeds, School of Computing (2001)
9. Tormos, P. and Lova, A. A Competitive Heuristic Solution Technique For Resource-Constrained Project Scheduling. Annals Of Operations Research 102, 65–81, (2001)
10. Tormos, P. et al. A Genetic Approach to Train Scheduling on a High-Traffic Railway Line. Algorithmic meThods and Models for Optimization of railwayS ATMOS 2005. Electronic Notes in Theoretical Computer Science. To be appear.
11. Walker, C., Snowdon, J. and Ryan, D. Simultaneous disruption recovery of a train timetable and crew roster in real time. Comput. Oper. Res. 32 (8), 2077–2094, (2005)