Technische Universität Berlin

FAKULTÄT II
MATHEMATIK UND
NATURWISSENSCHAFTEN
Institut für Mathematik

A Fast Algorithm for Near Cost Optimal Line Plans

by

Michael R. Bussieck Thomas Lindner
Marco E. Lübbecke

No. 2003/43

Mathematical Methods of Operations Research manuscript No.
(will be inserted by the editor)

A Fast Algorithm for Near Cost Optimal Line Plans

Michael R. Bussieck, Thomas Lindner, Marco E. Lübbecke

[1] GAMS Development Corporation, 1217 Potomac St. NW, Washington DC 20007, e-mail: `MBussieck@gams.com`
[2] Siemens Transportation Systems, Ackerstraße 22, D-38126 Braunschweig, Germany,
e-mail: `Thomas.Lindner@ts.siemens.de`
[3] Technische Universität Berlin, Institut für Mathematik, Sekr. MA 6-1
Straße des 17. Juni 136, D-10623 Berlin, Germany, e-mail: `m.luebbecke@math.tu-berlin.de`

November 11, 2003

Abstract We consider the design of line plans in public transport at a minimal total cost. Both, linear and nonlinear integer programming are adequate and intuitive modeling approaches for this problem. We present a heuristic variable fixing procedure which builds on problem knowledge from both techniques. We derive and compare lower bounds from different linearizations in order to assess the quality of our solutions. The involved integer linear programs are strengthened by means of problem specific valid inequalities. Computational results with practical data from the Dutch Railways indicate that our algorithm gives excellent solutions within minutes of computation time.

Key words Integer programming, nonlinear integer programming, cutting planes, public transport, line planning

MSC (2000) 90C10, 90C30, 90C57, 90B06

1 Introduction

In a public transportation network, we refer to a *line* as a route, or itinerary, together with the specification of how often the route is operated. The line planning problem is the design of lines which meet several operational constraints, most notably the passengers' demand for transportation. The resulting line plan, or route map, is visually known to anyone who ever traveled by bus, tram, subway, or train. Two objectives have been considered in the literature, and these reflect both sides of the story, namely service versus cost aspects. While travelers demand for convenient, ideally *direct* connections [2, 3], transportation companies are forced to make most efficient use of their resources [2, 6, 8, 11], not least for the reason of market deregulations. Minimization of the total cost is the goal of our investigation as well.

One main issue of our paper is to further the integration of two areas in mathematical programming all too often regarded separately, *viz.* nonlinear and mixed integer optimization. While the latter is a well established technique in the planning of public rail transportation issues [4], the track towards the topic via the former is practically unbeaten. In fact, mixed integer nonlinear models and algorithms appear not to belong to the "traditional" operations research active areas. This contrasts the tenor of the OR/MS Today Special Issue on innovative education: Nonlinear optimization deserves more emphasis [9].

In the cost optimal line planning context, Goossens et al. [8] demonstrate how to use lower bounds from a binary linear program [6] in a branch-and-cut algorithm. While the quality of lower bounds is excellent, primal solutions of matching quality cannot be found in reasonable time. In fact, the emphasis of previous exact approaches is on lower bounds. Most recently, Fischetti and Lodi [7] developed a general strategy for obtaining feasible solutions to mixed integer programs, and also make some computational progress on our particular problem. The main purpose of this paper is to justify that an appropriate combination of two models is suited to yield very high quality solutions within very short computation time limits. After all, this—and only this—is relevant to practitioners.

Our paper delivers updated material from [2] which has already been built on by other authors [8]. We contribute linear and nonlinear integer programs, on which we base our variable fixing algorithm. Our successful use of a nonlinear model is a new and surprising development in the design of line plans.

2 Cost Optimal Line Planning

In this section we formally describe the cost optimal line planning problem and introduce our notation. We do not repeat the practical background, and we refrain from comparing cost minimization with the direct traveler approach; see e.g., [3,6] for details on both issues.

The underlying rail network is represented by a graph $G=(V,E)$. Edges $e \in E$ model physical links which connect stations given by the set V of vertices, cf. Figure 1 for an example. Only a predefined subset $\mathcal{R}$ of paths, and sometimes cycles, in G qualifies as possible itineraries along which trains can be operated. The design of a line $l=(r,\varphi)$ requires the choice of a route $r \in \mathcal{R}$ and a frequency $\varphi \in \mathcal{F} \subset \mathbb{Z}_+$, i.e., the number of times the route is operated per hour. Unless otherwise stated, we assume that $\mathcal{F}$ is an interval of (small) integers, including zero. The length of route $r \in \mathcal{R}$ in kilometers is denoted by km_r. A feasible set of lines, or line plan, must also obey lower and upper bounds $F_e^{\min}$ and $F_e^{\max}$ on the line frequency requirement, i.e., on the total number of trains which traverse a link $e \in \mathcal{E}$ per hour. In addition, for each edge $e \in \mathcal{E}$ we are given its traffic load L_e, that is, the total number of seats per hour which are to be provided on the link. Our study is driven by the assumption that the cost of a line plan essentially consist of personnel cost and the cost of rolling stock. This motivates making the number of coaches of each train a decision variable for every particular line. This number must range within given bounds $\underline{c}$ and $\overline{c}$. It is important to note that we consider a single type of coach only, with a single seat capacity cap. Therefore, it is more natural to express the traffic load on an edge in terms of coaches, rather than in terms of of seats. That is, we define $\Lambda_e := \lceil L_e/cap \rceil$, the number of coaches to be provided on link e.

We denote by C_t^{fix} and C_c^{fix}, respectively, the hourly fixed capital and personnel cost incurred per train/per car. The respective operational or variable cost, that is, the per kilometer cost are denoted by C_t^{var} and C_c^{var}. The cost optimal line planning problem is to design a feasible line plan at minimum total cost. This problem is strongly $\mathcal{NP}$-hard [2].

3 Primal Solutions from a Nonlinear Model

Perhaps the most intuitive model formulation for our problem is a nonlinear mixed integer program (MINLP). Such a formulation was already given in Claessens' master's thesis [5] but was later dropped for the publication [6] because all presented results came out of a linearized model. We develop new ideas to overcome some of the technical obstacles which motivates us to revisit the MINLP approach.

Nonlinear mixed integer programming is a comparatively recent area with the first general purpose algorithms appearing in the early 1990s (`GAMS/DICOPT` [10]). Recently, several implementations emerged (branch-and-bound, outer approximation, extended cutting plane) due to more *stable* nonlinear programming codes (`FILTERBB`, `SBB`, `AlphaECP`, `mittlp`, `BARON` (global), other MINLP solvers embedded in modeling languages like `LINGO` and `MINOPT`[1]). An increase of interest in MINLP models and their solution is reflected in the growing number (absolute and relative) of entries at the NEOS[2] site. Within six months (09/01–02/02), their MINLP section got about 2,500 user submissions (private communication).

---

[1] Solver information available on the web under `http://www.gamsworld.org/minlp/links.htm`, `http://at8.abo.fi/~hasku/mittlp/`, and `http://www.lindo.com/table/lgosolvet.html`

[2] URL `http://www-fp.mcs.anl.gov/otc/Guide/SoftwareGuide/`

Figure 1 The Netherlands' InterRegio network with 86 vertices and 113 edges.

Practically relevant nonlinear models have some drawbacks compared to their linear counterparts: There is practically no proof of global optimality due to the non-convexity of practical models; the starting point is critical; in order to prevent algorithms from failing, problem specific knowledge is essential for a *good* model formulation, and thus, for algorithmic guidance. We have chosen our nonlinear approach in view of these considerations.

For each route $r \in \mathcal{R}$ we introduce a variable $x_r \in \mathbb{Z}_+$ which reflects its frequency and an integer variable $\underline{c} \le y_r \le \overline{c}$ for the number of coaches per train. The requirement that the line plan provide sufficient capacity for each link in the network then naturally writes down as

$$\sum_{r \in \mathcal{R}, r \ni e} x_r \cdot y_r \;\ge\; \Lambda_e \qquad \forall e \in E \;. \tag{1}$$

Similar quadratic terms appear in the objective function. However, for the actual number of trains per hour needed to operate a line at a given frequency we have to count more carefully. It depends on the total duration of a route, including its minimum turn-around time needed e.g., for cleaning the train, maintenance, and changing the crew. A division of this duration in minutes by 60 yields a fractional estimate T_r of the number of trains needed to operate $r \in \mathcal{R}$ *once* per hour. The overall requirement on that route is then $\lceil T_r \cdot x_r \rceil$ trains. Although undefined derivatives of such discontinuous terms are handled by solvers, modelers are advised against using them. Instead, defining variables $z_r \in \mathbb{Z}_+$, $r \in \mathcal{R}$ via

$$T_r \cdot x_r \;\le\; z_r \qquad \forall r \in \mathcal{R} \tag{2}$$

we theoretically can get over this problem. Because of our minimization objective there is no need for an explicit upper bound on these variables. By means of a translation we eliminate the non-trivial lower bounds on y_r. That is, new variables y_r count the number of *additional* coaches in excess of the minimal number $\underline{c}$ on route $r \in \mathcal{R}$. The complete MINLP model reads as follows:

$$
\begin{aligned}
\text{minimize} \sum_{r\in\mathcal{R}} & z_r\cdot(C_t^{\text{fix}}+(\underline{c}+y_r)\cdot C_c^{\text{fix}})+km_r\cdot x_r\cdot(C_t^{\text{var}}+(\underline{c}+y_r)\cdot C_c^{\text{var}}) \\
\text{subject to} \quad & \sum_{r\in\mathcal{R},r\ni e} x_r \geq F_e^{\min} && \forall e\in E \\
& \sum_{r\in\mathcal{R},r\ni e} x_r \leq F_e^{\max} && \forall e\in E \\
& \sum_{r\in\mathcal{R},r\ni e} x_r\cdot(\underline{c}+y_r) \geq \Lambda_e && \forall e\in E \\
& T_r\cdot x_r \leq z_r && \forall r\in\mathcal{R} \\
& y_r \leq \overline{c}-\underline{c} && \forall r\in\mathcal{R} \\
& x_r \in \mathcal{F} && \forall r\in\mathcal{R} \\
& y_r \in \mathbb{Z}_+ && \forall r\in\mathcal{R} \\
& z_r \in \mathbb{Z}_+ && \forall r\in\mathcal{R}
\end{aligned}
\tag{NLP}
$$

Solving this model leaves us with the dilemma of local versus global optimality. Implementations of global optimization algorithms like BARON do not find useful bounds and suffer from the same difficulties as linear branch-and-bound codes for finding integer solutions for real-world sized instances. Even large scale local optimization solvers perform poorly on these instances. Our first remedy to make this model useful for the cost optimal line planing problem simplifies the calculation of the number of trains required. Since

$$\lceil T_r\cdot x_r\rceil \quad\leq\quad \lceil T_r\rceil\cdot x_r \tag{3}$$

holds for $x_r\in\mathbb{Z}_+$ we may substitute z_r for the right hand side of (3), thereby getting rid of the additonal variables as well as their defining inequalities. This modification affects the fixed cost term in the objective function only. For our data sets, the effect is mild since we are not given a C_t^{fix}. Still, the level of error induced by the overestimation clearly depends on the instance. With hourly frequencies, i.e., $\mathcal{F}=\{0,1\}$, which are common in German and Dutch InterCity networks this approach is exact. For our instances, quite a lot of inequalities (3) are strict, but comparing the exact cost with the modified cost for all connections in our solutions, we obtain an average relative error of at most 3%. This is explained by the fact that, in our experiments, the best found solutions have almost all lines with a frequency of one anyway. Note that an obvious advantage of this formulation is its small number of variables.

4 Lower Bounds from Linearizations

Since (NLP) is a non convex model, the "lower bound" obtained from solving its continuous relaxation with available local optimization solvers is mathematically useless. As a remedy to this shortcoming we present three linearizations. The first essentially is a proposal by Claessens et al. [6] with an enormous number of variables. We introduce two new models which make up for this defect.

4.1 A Binary Linear Program

The fact that all nonlinear terms in (NLP) are of the form $x_r y_r$ and all variables are integer suggests a linearization by considering only the meaningful values of the respective variable products. More precisely, introduce a binary variable $z_{r,\varphi,\gamma}$ which assumes a value of one if and only if route $r\in\mathcal{R}$ is operated with frequency $\varphi\in\mathcal{F}$ using $\gamma\in\Gamma:=\{\underline{c},\ldots,\overline{c}\}$ coaches [6]. With an additional constraint which ensures that exactly one combination is selected we replace $x_r y_r$ by $\sum_{\varphi\in\mathcal{F}}\sum_{\gamma\in\Gamma}\varphi\cdot\gamma\cdot z_{r,\varphi,\gamma}$. Note that this substitution even eliminates the discontinuous term from the objective function. The model reads

$$
\begin{aligned}
\text{minimize} \quad & \sum_{r\in\mathcal{R}}\sum_{\varphi\in\mathcal{F}}\sum_{\gamma\in\Gamma}\Big(\lceil T_r\cdot\varphi\rceil\cdot(C_t^{\text{fix}}+\gamma\cdot C_c^{\text{fix}})+km_r\cdot\varphi\cdot(C_t^{\text{var}}+\gamma\cdot C_c^{\text{var}})\Big)\cdot z_{r,\varphi,\gamma} \\
\text{subject to} \quad & \sum_{r\in\mathcal{R},r\ni e}\ \sum_{\varphi\in\mathcal{F}}\sum_{\gamma\in\Gamma}\varphi\cdot z_{r,\varphi,\gamma}\geq F_e^{\min} && \forall e\in E \\
& \sum_{r\in\mathcal{R},r\ni e}\ \sum_{\varphi\in\mathcal{F}}\sum_{\gamma\in\Gamma}\varphi\cdot z_{r,\varphi,\gamma}\leq F_e^{\max} && \forall e\in E \\
& \sum_{r\in\mathcal{R},r\ni e}\ \sum_{\varphi\in\mathcal{F}}\sum_{\gamma\in\Gamma}\varphi\cdot\gamma\cdot z_{r,\varphi,\gamma}\geq \Lambda_e && \forall e\in E \\
& \sum_{\varphi\in\mathcal{F}}\sum_{\gamma\in\Gamma}z_{r,\varphi,\gamma}\leq 1 && \forall r\in\mathcal{R} \\
& z_{r,\varphi,\gamma}\in\{0,1\} && \forall r\in\mathcal{R},\ \varphi\in\mathcal{F},\ \gamma\in\Gamma
\end{aligned}
\tag{BLP}
$$

Model (BLP) is more flexible than (NLP) with respect to important operational constraints which result in holes in the domains of variables. For instance, $\mathcal{F}=\{1,2,4\}$ is a reasonable choice where explicitly $3\notin\mathcal{F}$. Moreover, a technical particularity of the rolling stock used by the Dutch Railways does not allow for trains made up of exactly five coaches, while $\Gamma=\{3,4,6,7,\dots\}$ is allowed. Such restrictions entail additional constraints in (NLP) but are easily incorporated in (BLP) simply by dropping the respective variables.

4.2 An Integer Linear Program Based on Frequencies

A drawback of (BLP) clearly is the comparatively large number of variables. Using integer variables, we are able to introduce a more compact new formulation. Similar to (NLP) we state the model in the formulation with additional coaches. Binary variables $x_{r,\varphi}$ indicate whether a route/frequency combination is chosen, and if so, variables $y_{r,\varphi}$ reflect the number of coaches more than $\underline{c}$.

$$
\begin{aligned}
\text{minimize} \quad & \sum_{r\in\mathcal{R}}\sum_{\varphi\in\mathcal{F}}\lceil T_r\cdot\varphi\rceil\cdot(x_{r,\varphi}\cdot C_t^{\text{fix}}+(\underline{c}\cdot x_{r,\varphi}+y_{r,\varphi})\cdot C_c^{\text{fix}}) \\
& \qquad +km_r\cdot\varphi\cdot(x_{r,\varphi}\cdot C_t^{\text{var}}+(\underline{c}\cdot x_{r,\varphi}+y_{r,\varphi})\cdot C_c^{\text{var}}) \\
\text{subject to} \quad & \sum_{r\in\mathcal{R},r\ni e}\ \sum_{\varphi\in\mathcal{F}}\varphi\cdot x_{r,\varphi}\geq F_e^{\min} && \forall e\in E \\
& \sum_{r\in\mathcal{R},r\ni e}\ \sum_{\varphi\in\mathcal{F}}\varphi\cdot x_{r,\varphi}\leq F_e^{\max} && \forall e\in E \\
& \sum_{r\in\mathcal{R},r\ni e}\ \sum_{\varphi\in\mathcal{F}}\varphi\cdot(\underline{c}\cdot x_{r,\varphi}+y_{r,\varphi})\geq \Lambda_e && \forall e\in E \\
& y_{r,\varphi}-(\overline{c}-\underline{c})\cdot x_{r,\varphi}\leq 0 && \forall r\in\mathcal{R},\ \varphi\in\mathcal{F} \\
& \sum_{\varphi\in\mathcal{F}}x_{r,\varphi}\leq 1 && \forall r\in\mathcal{R} \\
& x_{r,\varphi}\in\{0,1\} && \forall r\in\mathcal{R},\ \varphi\in\mathcal{F} \\
& y_{r,\varphi}\in\mathbb{Z}_+ && \forall r\in\mathcal{R},\ \varphi\in\mathcal{F}
\end{aligned}
\tag{ILP}
$$

4.3 A Third Linearization Based on Capacities

Similarly, one might be tempted to come up with a third linearization. Here, a binary variable $x_{r,\gamma}$ selects the combination of a route and the corresponding capacity while an integer variable $y_{r,\gamma}$ indicates the frequency of the respective combination. Even though the constraints of this model are similar to (ILP) their structure is slightly more simple. Most notably, we can state the bounds on the line frequency requirement with constraints having 0/1 coefficients only. This happens at the cost of the discontinuous term reappearing in the objective function, and the necessary workaround in spirit of Equation (2). We may drop the ceiling when using this model for the purpose of lower bounding only.

$$
\begin{aligned}
\text{minimize} \quad & \sum_{r\in\mathcal{R}}\sum_{\gamma\in\Gamma} \lceil T_r \cdot y_{r,\gamma}\rceil \cdot (C_t^{\text{fix}} + \gamma \cdot C_c^{\text{fix}}) + km_r \cdot y_{r,\gamma} \cdot (C_t^{\text{var}} + \gamma \cdot C_c^{\text{var}}) \\
\text{subject to} \quad & \sum_{r\in\mathcal{R}, r\ni e}\sum_{\gamma\in\Gamma} y_{r,\gamma} \geq F_e^{\min} && \forall e \in E \\
& \sum_{r\in\mathcal{R}, r\ni e}\sum_{\gamma\in\Gamma} y_{r,\gamma} \leq F_e^{\max} && \forall e \in E \\
& \sum_{r\in\mathcal{R}, r\ni e}\sum_{\gamma\in\Gamma} \gamma \cdot y_{r,\gamma} \geq \Lambda_e && \forall e \in E \\
& y_{r,\gamma} - \varphi_{\max} \cdot x_{r,\gamma} \leq 0 && \forall r \in \mathcal{R}, \gamma \in \Gamma \\
& \sum_{\gamma\in\Gamma} x_{r,\gamma} \leq 1 && \forall r \in \mathcal{R} \\
& x_{r,\gamma} \in \{0,1\} && \forall r \in \mathcal{R}, \gamma \in \Gamma \\
& y_{r,\gamma} \in \mathcal{F} && \forall r \in \mathcal{R}, \gamma \in \Gamma
\end{aligned}
\tag{ILP-C}
$$

This model has another interesting property in the case that only one frequency (besides zero) is admissible. Again, we may substitute the discontinuous term in the objective function for $\lceil T_r\rceil \cdot y_{r,\gamma}$ but we can also completely drop the $y_{r,\gamma}$ variables including their defining constraints from the formulation altogether. This is because for $\mathcal{F} = \{0,1\}$, i.e., $\varphi_{\max} = 1$, together with $\sum_{\gamma\in\Gamma} x_{r,\gamma} \leq 1$, $r \in \mathcal{R}$ we have $y_{r,\gamma} \leq 1 \cdot x_{r,\gamma}$ anyway and we substitute $x_{r,\gamma}$ for $y_{r,\gamma}$. The remaining model is (BLP) with fixed $\varphi \equiv 1$! The emphasis with respect to important decisions in this model clearly lies on capacities rather than on frequencies. In a sense, among the three proposals this is the linearization closest to (NLP).

5 Valid Inequalities

In strengthening our linearizations by valid inequalities we aim at improving the quality of lower bounds. We present valid inequalities for (ILP) only. Our argumentation is based on the problem rather than on models, and all results of this section immediately carry over to the other linearizations. We prefer giving the intuition behind our results only. The reason for doing so is to demonstrate that rather fancy looking cutting planes are nothing more (and less) but good problem knowledge. Detailed proofs and some advanced preprocessing techniques can be found in [2].

Let us abbreviate $\Lambda_{E'} := \sum_{e\in E'} \Lambda_e$. We will now generalize the following idea. Let $e, f \in E$ be the only edges incident to $v \in V$, and let $\Lambda_e < \Lambda_f$. When no line via f ends in v, the number of coaches running via e is at least Λ_f.

Proposition 1 *Let $E' \subseteq E$, $f \in E \setminus E'$, $\alpha^r_{E'} = |E' \cap r|$, and $\Lambda_f > \Lambda_{E'}$. Then,*

$$
(\Lambda_f - \Lambda_{E'}) \sum_{r\in\mathcal{R}, \alpha^r_{E'}=0, r\ni f}\ \sum_{\varphi\in\mathcal{F}} x_{r,\varphi} + \sum_{r\in\mathcal{R}}\sum_{\varphi\in\mathcal{F}} \alpha^r_{E'} \cdot \varphi \cdot (\underline{c} \cdot x_{r,\varphi} + y_{r,\varphi}) \ \geq\ \Lambda_f \tag{4}
$$

is a valid inequality for (ILP)*.*

The following two classes are particularly useful in our experiments. A tentative explanation is that they exploit observations on capacities *and* frequencies. Consider an edge $e \in E$ with $\underline{c} \cdot F_e^{\min} < \Lambda_e < \underline{c}(F_e^{\min} + 1)$. Then, the number of trains via e is at least $F_e^{\min} + 1$ or the number of additional coaches of lines via e is at least $\xi := \Lambda_e - \underline{c} \cdot F_e^{\min}$.

Proposition 2 *For every edge $e \in E$ with $\underline{c} \cdot F_e^{\min} < \Lambda_e < \underline{c}(F_e^{\min} + 1)$*

$$
\sum_{r\in\mathcal{R}, r\ni e}\sum_{\varphi\in\mathcal{F}} \xi \cdot \varphi \cdot x_{r,\varphi} + \min\{\xi, \varphi\} \cdot y_{r,\varphi} \ \geq\ \xi \cdot (F_e^{\min} + 1) \tag{5}
$$

is a valid inequality for (ILP)*.*

The line frequency requirement impacts the total number of coaches on an edge due to the lower bound $\underline{c}$ on the number of coaches per train. Suppose the line plan contains a line (r,φ) with $e \in r$. Independently of the particular demand Λ_e there must be at least $\underline{c}(F_e^{\min} - \varphi)$ coaches of other lines via e. This is generalized in the following result which is the most effective in our experience.

Proposition 3 *Let $e \in E$ and $\mathcal{R}' \subset \mathcal{R}$ with $e \in r$ for all $r \in \mathcal{R}'$ and $\sum_{r\in\mathcal{R}'}\sum_{\varphi\in\mathcal{F}} x_{r,\varphi} \le 1$ in any feasible solution to* (ILP)*. Then,*

$$\sum_{r\in\mathcal{R}\setminus\mathcal{R}',r\ni e}\ \sum_{\varphi\in\mathcal{F}} \varphi\cdot(\underline{c}\cdot x_{r,\varphi}+y_{r,\varphi}) \ \ge\ \Lambda_e\left(1-\sum_{r\in\mathcal{R}'}\sum_{\varphi\in\mathcal{F}} x_{r,\varphi}\right) +\ \underline{c}\sum_{r\in\mathcal{R}'}\sum_{\varphi\in\mathcal{F}}(F_e^{\min}-\varphi)\cdot x_{r,\varphi} \tag{6}$$

is a valid inequality for (ILP)*.*

Our final proposal involves edges $e \in E$ with a large demand Λ_e but with a small minimum line frequency requirement $F_e^{\min}$. When the number of trains via e is close to $F_e^{\min}$ the number of coaches in the operated lines must be close to $\overline{c}$. The following is a generalization of this statement. Observe that in our models we count coaches in addition to $\underline{c}$.

Proposition 4 *Let $e \in E$ and (r^*,φ^*) be a line with $e \in r^*$. Define $\zeta := \Lambda_e - \overline{c}(F_e^{\min} - \max\{1, F_e^{\min} - F_e^{\max} + \varphi^*\}) - \varphi^*\underline{c}$. If $\zeta \ge 0$, then*

$$\varphi^*\cdot y_{r^*,\varphi^*} - \left(F_e^{\min} - \sum_{(r,\varphi)\neq(r^*,\varphi^*),r\ni e} \varphi\cdot x_{r,\varphi}\right)\lceil\zeta/\varphi^*\rceil \ \ge\ 0 \tag{7}$$

is a valid inequality for (ILP)*.*

We close this section with a few words on separation, i.e., on identifying valid inequalities which are violated by a given fractional solution to (ILP). We cannot give a general scheme for separating (4) for *all* $E' \subseteq E$ which actually is $\mathcal{NP}$-hard [2]. However, our experiments show that inequalities with $|E'| = 1$ dominate those with larger $|E'|$. The number of inequalities (5) and (7) is polynomial by definition. Also, in (6) we consider the special case of $|\mathcal{R}'| = 1$ only. Hence, for our instances, we check only a polynomial number of inequalities. In particular, we separate valid inequalities of Propositions 1–4 in a *cut-and-branch* fashion, i.e., in the root node of the branch-and-bound tree only, see Algorithm 1. In the remainder of this paper, we refer to this latter model (ILP) augmented by valid inequalities as model (CUT). For a branch-and-cut approach based on (BLP) see [8].

6 Turning Models into an Algorithm

We already remarked that optimally solving our nonlinear model is not an option. Even our simplified version of (NLP) using (3) represents a challenge to existing MINLP codes. The problem is still far from being trivial. This can be deduced from the performance of the until recently only method to solve MINLPs: A linearization at the (local) optimum point of the relaxed MINLP results in a mixed integer program and is solved as a second step in the outer approximation algorithm of `DICOPT`. Solving only the MIP by `CPLEX` 7.5 takes over four hours.

Lagrangian relaxation yields unsatisfactory results on the original MINLP model by Claessens [5]. She proposes an iterative rounding heuristic along the following lines. Promising routes with $x_r y_r$ larger than a given threshold are incorporated in the line plan by bounding $x_r \ge 1$ from below. On the other hand, routes with small $x_r y_r$ value are eliminated from further inspection. When all routes are either deleted or incorporated, all x_r variables are rounded up to the next integer. The resulting program is an integer linear program in the y_r variables. An alternative to this approach is to fix y_r variables instead. In the resulting integer linear program one has fixed capacities on each route and sufficient seat capacity needs to be provided for each link at minimal cost. This problem remains $\mathcal{NP}$-hard even for $F_e^{\min} \equiv F_e^{\max} \equiv 1$, $e \in E$ [2]. All these earlier proposals failed in finding good primal solutions.

For large scale mixed integer programs in general, and for our (integer) linearizations in particular, we cannot hope for an optimal solution within practically relevant time bounds. One would terminate a branch-and-bound algorithm when a predefined computation time limit is reached. A drawback is the risk of not having found *any* primal feasible solution at that time. Modern commercial branch-and-bound codes optionally apply sophisticated strategies and invest quite some time in the root node in order to come up with heuristic solutions. Commonly used variable rounding fails for our instances. The latest CPLEX 8.0 fares better in that it implements various *MIP emphasis* preferences to accomodate, among others, a bias towards finding feasible solutions. Fischetti and Lodi [7] develop a general two-level branching framework which controls feasibility at a higher "strategic" level. Computationally, this method outperforms CPLEX' defaults. A considerable improvement as compared to relying on general strategies is to develop a problem specific algorithm which exploits the models we introduced earlier. This is what we do now.

Fixing *all* lines according to Claessens' procedure is a too restrictive measurement, and makes permanent decisions too early. Our idea is to use model (NLP) to get a hint to promising lines and pass this information to model (ILP), still leaving enough freedom to work on these. At first, we invest in a meaningful starting point for (NLP). To this end, we solve the linear programming relaxation of (ILP). Variables in (NLP) are initialized according to their transformed counterparts in the linearized model's optimal solution. A warm or hot start with a former good solution is possible as well. Then, the continuous relaxation of (NLP) is solved. In the solution, a variable x_{r^*} close to zero is interpreted as r^* being an unpromising route, and all variables $x_{r^*,\varphi}$ or lines involving such a route are eliminated from model (ILP) by fixing these variables to zero. All other variables are left untouched. We then succeed in finding an integer solution to the partially fixed (ILP) model within three minutes.

From this point on, we have a feasible solution and we may invest in either improving it or assessing its quality. Thus, we add valid inequalities to (ILP) as discussed at the end of the preceding section. We remove the variable fixing again, and continue with a standard branch-and-bound on model (ILP). We summarize the whole procedure, we term (FIX), in Algorithm 1. Note that the variable fixing is a heuristic, even though for all our instances we quickly find an integer solution. In one sentence, it is a very careful initialization of branch-and-bound.

Our richness of models allows for several alternatives. Instead of using (ILP) we could run (BLP), or combinations of the models. Also, the final model to be solved need not be a linearization. We could feed our integral solution obtained from the fixed linearization into model (NLP) as a better starting point. The subsequent nonlinear programming based branch-and-bound phase performs as is well-known for linear mixed integer programs, except that "bounds" in each node are obtained from the relaxed nonlinear integer program (NLP).

7 Computational Results

We implemented all our models in the GAMS modeling language [1], version 20.6, using the combination of CONOPT2 and SBB for solving the nonlinear models, and CPLEX 7.5 for the solution of the linear (integer) models. Our code is publicly available.[3] We deem it important that the reader be in a position to get first-hand experience with our techniques, reproduce our results, and verify our conclusions. In the operations research community, usually with reference to proprietary data, these key ingredients of scientific activity are often regarded as being of secondary importance. In contrast, we would like to enable the reader to check claims which we do not explicitly support by numerical results.

Our computational experiments are performed on a 700MHz Linux PC with an execution time limit of three hours. We are provided with four practical problem instances from the Dutch Railways (Nederlandse Spoorwegen), `ic97`, `ic98`, `ir98`, and `ar98`. The names stem from the underlying network (InterCity, InterRegio, and AggloRegio) and the year. Variants of these instances have been circulating in the community for years and, except `ir98`, so far withstood a proven optimal solution, even with massive computational power [2]. Only recently, Bixby optimally solved instance `ic97` in a month's computation time using our model (CUT) (private communication), and we obtained an optimal solution to `ic98` using the same model

[3] URL `http://www.gamsworld.org/minlp/apps/blllop`

Algorithm 1 Combining our models into the variable fixing procedure (FIX)

Solve the LP relaxation of (ILP)
Let x^{ILP}, y^{ILP} denote an optimal fractional solution
Initialize variables x^{NLP}, y^{NLP} of model (NLP):
for all $r \in \mathcal{R}$ **do**

$$x_r^{NLP} = \sum_{\varphi \in \mathcal{F}} \varphi \cdot x_{r,\varphi}^{ILP} \text{ for all } r \in \mathcal{R}$$

$$y_r^{NLP} = \begin{cases} \sum_{\varphi \in \mathcal{F}} \varphi \cdot y_{r,\varphi}^{ILP} / x_r^{NLP} & \text{if } x_r^{NLP} > 0 \\ 0 & \text{otherwise} \end{cases}$$

end for
Solve the continuous relaxation of (NLP)
According to the optimal continuous solution, fix variables in (ILP):
for all $r \in \mathcal{R}$, $\varphi \in \mathcal{F}$ **do**
  if $x_r^{NLP} \leq threshold$ (default: 10^{-5}) **then**
    Add constraint $x_{r,\varphi} = 0$ to (ILP)
  end if
end for
Integrally solve the fixed (ILP)
Drop again constraints $x_{r,\varphi} = 0$ from (ILP)
Sequentially add valid inequalities:
Add (5) to (ILP)
if (4) violated **then**
  Add (4) to (ILP)
else if (6) violated **then**
  Add (6) to (ILP)
else if (7) violated **then**
  Add (7) to (ILP)
end if
Solve LP relaxation of (ILP), and remove all cuts which are not binding in the optimal solution
Warm start (ILP) with the integral solution found
Integrally solve (ILP) by standard branch-and-bound

and comparable computational efforts. Since our comparison is intended to be qualitative in nature we do not present numerical results but illustrate our experiments in Figures 2 through 5. Horizontal solid lines indicate the optimal objective function value, or for `ar98`, the best known upper and lower bounds, respectively. All bounds are obtained using model (CUT).

In brief, as expected, there is no consistently best *all* purpose model. However, when looking at upper and lower bounds separately, the picture changes. Most strikingly, for three out of four instances our algorithm (FIX) provides us with our by far best feasible solutions in about *five minutes* computation time per instance. There is practically no improvement after three hours of additional search with either model, (ILP) or (BLP). A bit surprisingly, the integer solutions found by model (ILP) alone do not benefit from the smaller model size and the related quicker evaluation of branch-and-bound nodes compared to (BLP). In what regards lower bounds, model (CUT) is significantly superior except in one case; we give a cut statistic in Table 1. (ILP) appears to be next best, again with one exception. Interestingly enough, this one exception (with (FIX) unsuccessful, (CUT) and (ILP) inferior to (BLP), and the instance being optimally solvable within practical time bounds) is one and the same instance, namely `ir98`.

Results for model (ILP-C), both in terms of upper and lower bounds are poor, and we refrain from plotting them. The structure of this seemingly intuitive model is too heavily disturbed by the necessity to model the discontinuous term in the objective function via (2). Note that these findings stress the advantage of the other two linearizations. It should be mentioned that Goossens et al. optimally solve a variant of instance `ir98` within three minutes of computation time using (BLP) as the basic model [8]. The effectiveness of their branch-and-cut approach for (BLP) leads to the question how well this algorithm performed with (ILP) as basic model. However, as indicated in [8] the data instances these authors use are slightly different from ours and serious comparisons cannot be drawn. From their experiments it appears that problem specific branching rules, like branching on lines or capacities are not significantly superior over standard variable branching.

We made several experiments in order to identify the most effective combination of our models for the variable fixing algorithm, cf. Table 2. Admittedly, the benefit of model (NLP) is marginal in terms of additional solution quality. Almost equally good solutions are obtained by just solving e.g., the relaxed (ILP) model, eliminate unpromising lines as above, and solve the resulting smaller integer program. However, in our opinion, model (NLP) most naturally reflects the practical background via (1) which may lead to a slightly better "guidance" of the variable fixing. We consider it very important that—with or without (NLP)—we always obtain integer solutions of high quality; this points to the robustness of both approaches, and we value the added flexibility of using several models.

Instance	(4)	(5)	(6)	(7)	plain LP	w/ (4)–(7)	(ILP) root	w/ (4)–(7)
`ic97`	15	4	108	24	3845.47	3915.54	3875.45	3922.63
`ic98`	19	1	134	5	4328.83	4412.83	4376.06	4418.15
`ir98`	20	2	47	0	2230.54	2291.07	2286.28	2299.17
`ar98`	72	15	162	16	5882.90	6009.10	6064.06	6094.40

Table 1 Cut statistic and impact of valid inequalities on the linear programming relaxation of model (ILP). The columns indicate the name of the instance, the respective number of violated inequalities, and the respective objective function values of the LP relaxation, the LP relaxation with violated inequalities added, the LP value in the root node of the branch-and-bound tree (after CPLEX' preprocessing), and the latter with violated inequalities added, in that order.

init	guide	fixed	`ic97`	`ic98`	`ir98`	`ar98`
(ILP)	(NLP)	(ILP)	$\star$ **4088.52**	$\star$ **4521.67**	$\star$ 2385.84	**6336.44**
—	(ILP)	(ILP)	$\star$ **4088.52**	4559.90	$\star$ 2417.97	6368.08
(BLP)	(NLP)	(BLP)	4371.66	4811.93	$\star$ **2378.33**	6326.64
—	(BLP)	(BLP)	4476.51	4639.63	$\star$ 2386.27	6383.59

Table 2 Comparison of different model combinations for our variable fixing strategy. Headings 'init', 'guide', and 'fixed', respectively, refer to the relaxation which is used to find an initial solution for model (NLP), the relaxation for determining which lines to discard, and the partially fixed model to find an integer feasible solution. We report the objective function value of the best integer solution found after three minutes of the 'fixed' run. A $\star$ indicates an optimal solution of the 'fixed' model.

8 Conclusions

Nonlinear integer programming is quite often an adequate and intuitive modeling approach for combinatorial optimization problems. Recently, powerful software became available for actually solving these models, and we have to re-think about such approaches. In this paper we support this claim, and demonstrate how to combine nonlinear techniques with "classical" integer linear programming in order to successfully attack a very hard practical combinatorial optimization problem. The heuristic we present is robust in that it delivers an excellent integer solution for all our instances.

For better or for worse, the success or failure of an approach is linked to the dramatic development of commercial solvers. Not seldom we see that we can learn a lot from a tailored algorithm but an off-the-shelf product is much superior in terms of computation time, and often quality. This is especially true for linear and mixed integer programs. Solvers are *much* more sophisticated than those available when we set out with this research in 1997; in fact, we experimented, among others, with CPLEX versions 2.0 through 8.0 and it is not unlikely that future versions are competitive to our success with the variable fixing algorithm, or put its current usefulness in question altogether. However, we would like to make the point that these solvers also improve *because* of research like ours. Then again, in our opinion, the development of industry standard mixed integer nonlinear solvers is just at its beginning, and the headway of using a nonlinear model may even be amplified in the future.

One must critically ask the question whether an intelligent use of different models and solvers is eligible to be termed an *algorithm*, even though the notion is certainly technically appropriate. Our answer is affirmative. The outcome in terms of the building blocks *model* and *solver* looks simple but beneath the surface a lot of algorithmic understanding is indispensable. On the other hand, this simplicity at a higher level enables the practitioner to obtain reproducible and predictable results in finite time. We would like the notion of a *practical* algorithm.

Besides fundamental investigations in the latter direction open research avenues include the following specific modifications: Lines need not be operated back and forth. This may lead to concatenations of several lines. An additional complication in this setting is whether the resulting line plan is robust as to the subsequent planning stage of train scheduling. Also, rolling stock does not need to be dedicated to specific lines. Column generation, an entirely different algorithmic approach, has been suggested for the maximization of the number of direct travelers in [2]. A continuation of our study would investigate how suitable this approach is for the minimization of operational cost. We are confident that further computational and methodological progress will eventually lead to an integrated treatment of all the mentioned networks.

References

1. A. Brook, D. Kendrick, and A. Meeraus. *GAMS: A User's Guide, Release 2.25*. Scientific Press, San Francisco, CA, 1992.
2. M.R. Bussieck. *Optimal Lines in Public Rail Transport*. PhD thesis, Braunschweig University of Technology, 1998. Available under `http://www.biblio.tu-bs.de/ediss/data/19990121a/19990121a.pdf`.
3. M.R. Bussieck, P. Kreuzer, and U.T. Zimmermann. Optimal lines for railway systems. *European J. Oper. Res.*, 96(1):54–63, 1997.
4. M.R. Bussieck, T. Winter, and U.T. Zimmermann. Discrete optimization in public rail transport. *Math. Programming*, 79(1–3):415–444, 1997.
5. M.T. Claessens. De kost-lijnvoering. Master's thesis, University of Amsterdam, 1994.
6. M.T. Claessens, N.M. van Dijk, and P.J. Zwaneveld. Cost optimal allocation of rail passenger lines. *European J. Oper. Res.*, 110(3):474–489, 1998.
7. M. Fischetti and A. Lodi. Local branching. *Math. Programming*, 98(1–3):23–47, 2003.
8. J.-W. Goossens, S. van Hoesel, and L. Kroon. A branch-and-cut approach for solving line planning problems. METEOR Research Memorandum RM/01/016, University of Maastricht, 2001. To appear in *Transportation Sci.*
9. T.A. Grossman. Reversing tradition. *OR/MS Today*, 28(4):22–25, August 2001.
10. G. R. Kocis and I. E. Grossmann. Computational experience with DICOPT solving MINLP problems in process systems engineering. *Computers and Chemical Engineering*, 13:307–315, 1989.
11. P.J. Zwaneveld. *Railway Planning—Routing of Trains and Allocation of Passenger Lines*. PhD thesis, Rotterdam School of Management, 1997.

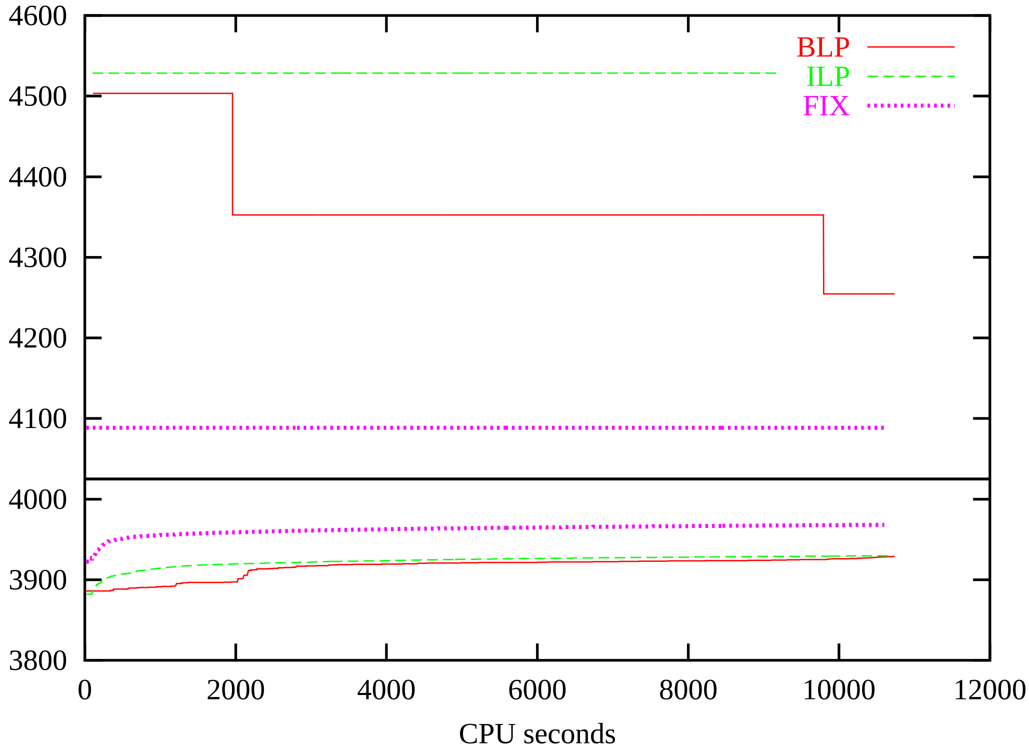

Figure 2 Development of bounds for instance `ic97`

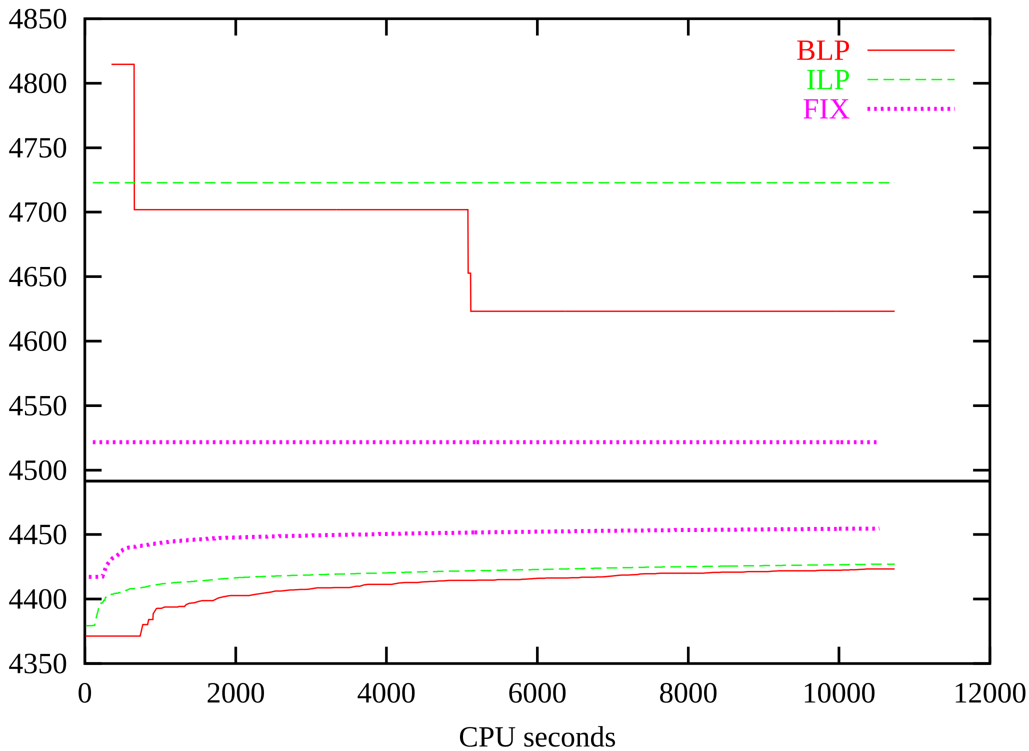

Figure 3 Development of bounds for instance `ic98`

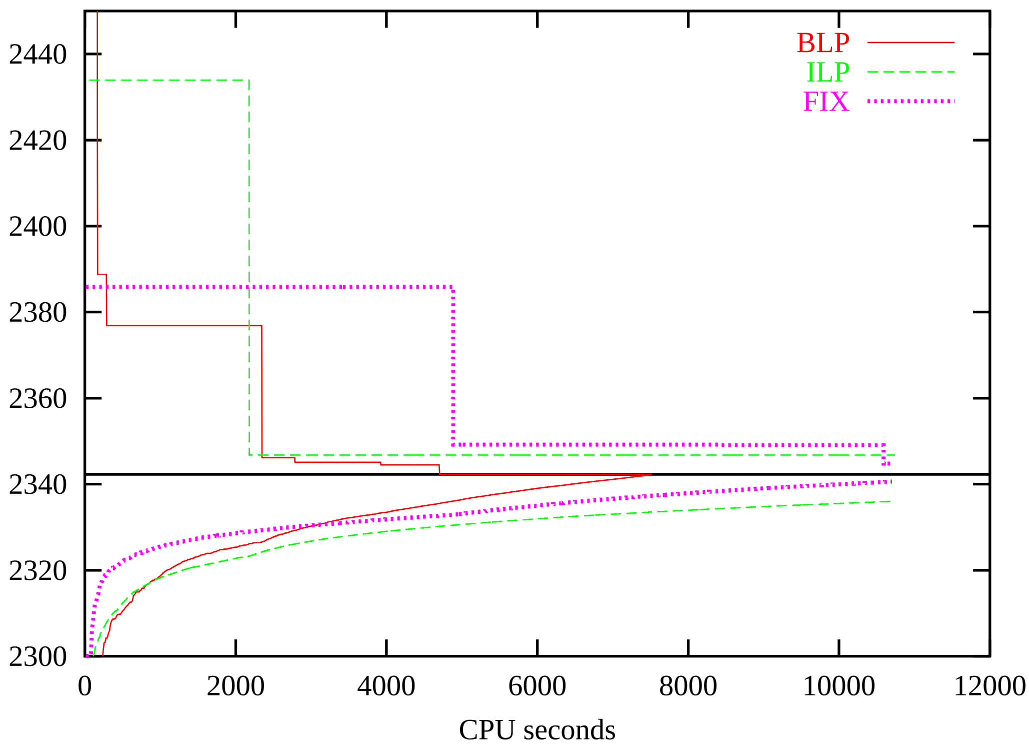

Figure 4 Development of bounds for instance `ir98`

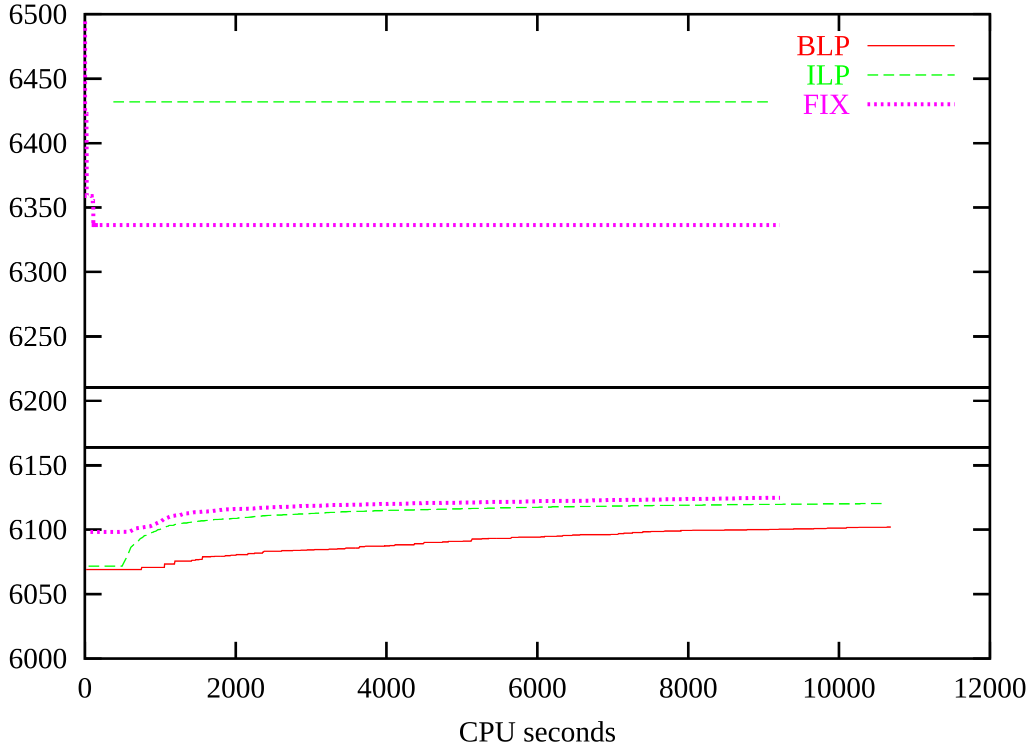

Figure 5 Development of bounds for instance `ar98`

Reports from the group

"Combinatorial Optimization and Graph Algorithms"

of the Department of Mathematics, TU Berlin

2003/43 *Michael R. Bussieck, Thomas Lindner, and Marco E. Lübbecke:* A Fast Algorithm for Near Cost Optimal Line Plans

2003/42 *Marco E. Lübbecke:* Dual Variable Based Fathoming in Dynamic Programs for Column Generation

2003/37 *Sándor P. Fekete, Marco E. Lübbecke, and Henk Meijer:* Minimizing the Stabbing Number of Matchings, Trees, and Triangulations

2003/25 *Daniel Villeneuve, Jacques Desrosiers, Marco E. Lübbecke, and François Soumis:* On Compact Formulations for Integer Programs Solved by Column Generation

2003/24 *Alex Hall, Katharina Langkau, and Martin Skutella:* An FPTAS for Quickest Multicommodity Flows with Inflow-Dependent Transit Times

2003/23 *Sven O. Krumke, Nicole Megow, and Tjark Vredeveld:* How to Whack Moles

2003/22 *Nicole Megow and Andreas S. Schulz:* Scheduling to Minimize Average Completion Time Revisited: Deterministic On-Line Algorithms

2003/16 *Christian Liebchen:* Symmetry for Periodic Railway Timetables

2003/12 *Christian Liebchen:* Finding Short Integral Cycle Bases for Cyclic Timetabling

762/2002 *Ekkehard Köhler and Katharina Langkau and Martin Skutella:* Time-Expanded Graphs for Flow-Dependent Transit Times

761/2002 *Christian Liebchen and Leon Peeters:* On Cyclic Timetabling and Cycles in Graphs

752/2002 *Ekkehard Köhler and Rolf H. Möhring and Martin Skutella:* Traffic Networks and Flows Over Time

739/2002 *Georg Baier and Ekkehard Köhler and Martin Skutella:* On the *k*-splittable Flow Problem

736/2002 *Christian Liebchen and Rolf H. Möhring:* A Case Study in Periodic Timetabling

723/2001 *Berit Johannes:* Scheduling Parallel Jobs to Minimize Makespan

716/2001 *Christian Liebchen:* The Periodic Assignment Problem (PAP) May Be Solved Greedily

711/2001 *Esther M. Arkin, Michael A. Bender, Sándor P. Fekete, Joseph S. B. Mitchell, and Martin Skutella:* The Freeze-Tag Problem: How to Wake Up a Swarm of Robots

710/2001 *Esther M. Arkin, Sándor P. Fekete, and Joseph S. B. Mitchell:* Algorithms for Manufacturing Paperclips and Sheet Metal Structures

705/2000 *Ekkehard Köhler:* Recognizing Graphs without Asteroidal Triples

704/2000 *Ekkehard Köhler:* AT-free, coAT-free Graphs and AT-free Posets

702/2000 *Frederik Stork:* Branch-and-Bound Algorithms for Stochastic Resource-Constrained Project Scheduling

700/2000 *Rolf H. Möhring:* Scheduling under uncertainty: Bounding the makespan distribution

698/2000 *Sándor P. Fekete, Ekkehard Köhler, and Jürgen Teich:* More-dimensional packing with order constraints

697/2000 *Sándor P. Fekete, Ekkehard Köhler, and Jürgen Teich:* Extending partial suborders and implication classes

696/2000 *Sándor P. Fekete, Ekkehard Köhler, and Jürgen Teich:* Optimal FPGA module placement with temporal precedence constraints

695/2000 *Sándor P. Fekete, Henk Meijer, André Rohe, and Walter Tietze:* Solving a "hard" problem to approximate an "easy" one: heuristics for maximum matchings and maximum Traveling Salesman Problems

694/2000 *Esther M. Arkin, Sándor P. Fekete, Ferran Hurtado, Joseph S. B. Mitchell, Marc Noy, Vera Sacristán and Saurabh Sethia:* On the reflexivity of point sets

693/2000 *Frederik Stork and Marc Uetz:* On the representation of resource constraints in project scheduling

691/2000 *Martin Skutella and Marc Uetz:* Scheduling precedence constrained jobs with stochastic processing times on parallel machines

689/2000 *Rolf H. Möhring, Martin Skutella, and Frederik Stork:* Scheduling with AND/OR precedence constraints

685/2000 *Martin Skutella:* Approximating the single source unsplittable min-cost flow problem

684/2000 *Han Hoogeveen, Martin Skutella, and Gerhard J. Woeginger:* Preemptive scheduling with rejection

683/2000 *Martin Skutella:* Convex quadratic and semidefinite programming relaxations in Scheduling

682/2000 *Rolf H. Möhring and Marc Uetz:* Scheduling scarce resources in chemical engineering

681/2000 *Rolf H. Möhring:* Scheduling under uncertainty: optimizing against a randomizing adversary

680/2000 *Rolf H. Möhring, Andreas S. Schulz, Frederik Stork, and Marc Uetz:* Solving project scheduling problems by minimum cut computations (Journal version for the previous Reports 620 and 661)

674/2000 *Esther M. Arkin, Michael A. Bender, Erik D. Demaine, Sándor P. Fekete, Joseph S. B. Mitchell, and Saurabh Sethia:* Optimal covering tours with turn costs

669/2000 *Michael Naatz:* A note on a question of C. D. Savage

667/2000 *Sándor P. Fekete and Henk Meijer:* On geometric maximum weight cliques

666/2000 *Sándor P. Fekete, Joseph S. B. Mitchell, and Karin Weinbrecht:* On the continuous Weber and k-median problems

664/2000 *Rolf H. Möhring, Andreas S. Schulz, Frederik Stork, and Marc Uetz:* On project scheduling with irregular starting time costs

661/2000 *Frederik Stork and Marc Uetz:* Resource-constrained project scheduling: from a Lagrangian relaxation to competitive solutions

Reports may be requested from:

Sekretariat MA 6–1
Fakultt II – Institut fr Mathematik
TU Berlin
Straße des 17. Juni 136
D-10623 Berlin – Germany

e-mail: klink@math.TU-Berlin.DE

Reports are also available in various formats from

`http://www.math.tu-berlin.de/coga/publications/techreports/`

and via anonymous ftp as

`ftp://ftp.math.tu-berlin.de/pub/Preprints/combi/Report-`*number*`-`*year*`.ps`