A combinatorial relaxation methodology for berth allocation problems

Evellyn Cavalcante* Johan Oppen* Phillipe Samer* Sebastián Urrutia †

July 1, 2016

Abstract

We investigate an optimization problem in container ports, for which previous models based on generalized set partitioning formulations have been studied. We describe two combinatorial relaxations based on computing maximum weighted matchings in suitable graphs. Besides dual bounds, those also provide a variable reduction technique and a matheuristic. Computational results indicate that the algorithm is an efficient alternative capable of finding the known optimal solutions on 83% of the benchmark instances. Even in the most difficult set of instances, solutions within 4% of the optimal value are provided, with an average of 30% of the time required by methods previously described in the literature.

Keywords: dual bounds, preprocessing, matheuristics, matchings, set partitioning, port operations, maritime logistics.

1 Introduction

In this work, we investigate a discrete optimization problem in maritime logistics. The Berth Allocation and Quay Crane Assignment Problem (BACAP) aims to allocate berthing position/time, and a number of quay cranes (QCs) for arriving vessels in a seaport container terminal. Feasible assignments in the BACAP need to fulfill requirements on desired berthing period and position, and an agreement on the QCs availability.

The Berth Allocation Problem (BAP) was introduced by Hoffarth and Voß [1994] and is one of the main problems in the operations of a container terminal [Stahlbock and Voß, 2008]. The Quay Crane Assignment Problem (QCAP) is usually solved in an integrated manner to the BAP, since it is computationally more tractable [Bierwirth and Meisel, 2010]. We investigate in this work the BACAP, a combination of both problems which was first studied by Park and Kim [2003]. The authors propose integer programming models to the problem and a Lagrangian heuristic. A criticism to this work is that, for simplicity, the

* Molde University College – P.O.box 2110, N-6402, Molde, Norway. (evellyn.cavalcante@gmail.com, johan.oppen@himolde.no, phillipes@gmail.com)
† Departamento de Ciência da Computação, Universidade Federal de Minas Gerais – Av. Antônio Carlos, 6627, Belo Horizonte, MG, 31270-010, Brazil. (surrutia@dcc.ufmg.br)

Acknowledgement: this work was carried out within the Port-Ship Coordinated Planning project, supported by the Norwegian Research Council under project number 227084/O70.
productivity rate of transporting a ship’s cargo is considered proportional to the number of cranes assigned to it, which is not true in practice [Meisel and Bierwirth, 2009].

In the work of Meisel and Bierwirth [2009], the problem is formulated with a compact model which includes the interference effects of the number of cranes in productivity, and obtains optimal solutions to small instances of 20 vessels. For large instances (with 30 and 40 vessels), primal limits are generated using different heuristics. In particular, the method based on squeaky wheel optimization is considered the state of the art among available heuristics for the BACAP. Considering a different application modeling, Vacca et al. [2013] introduce an alternative formulation with an exponential number of variables and a sophisticated branch and price algorithm.

More recently, BACAP variations have been formulated as a generalization of the set partitioning problem [Çağatay Iris et al., 2015, Türkoğulları et al., 2014]. In such formulations, each variable represents a feasible allocation/assignment for a single vessel, and its cost is a linear combination of deviations from the desired berthing period and the allocation of cranes. Çağatay Iris et al. [2015] solve the corresponding model after generating all variables a priori. Key features of this approach are a set of solution bounds and effective techniques for reducing the number of variables. The authors achieve satisfactory results for the same instances introduced by Meisel and Bierwirth [2009]. The work of Türkoğulları et al. [2014] follows the same strategy, but assume different application modeling and instance set.

Our contribution in this work consists of two combinatorial relaxations to the generalized set partitioning formulation of the BACAP used by Çağatay Iris et al. [2015]. Besides showing that matchings in two suitable graphs provide dual bounds to the problem, we describe two ways to employ these relaxations in solution algorithms for the problem following the variable enumeration approach. First, we derive a preprocessing method based on the removal of columns that imply a suboptimal solution, as it is done in the work of Çağatay Iris et al. [2015]. Finally, we present a matheuristic capable of finding the known optimal solutions on 83% of the benchmark instances. Even in the most difficult instance set, solutions within 4% of the optimal value are provided, with an average of 30% of the time required by the method of Çağatay Iris et al. [2015].

The paper is organized as follows. In Section 2 we describe the generalized set partitioning formulation for the BACAP. Section 3 describes two combinatorial relaxations for the problem, and its extension into a variable reduction technique and a matheuristic. Computational results are presented in the Section 4, and final remarks are compiled in Section 5.

2 Generalized set partitioning formulation

For completeness, we begin with a compilation of application characteristics that are assumed in this modeling of the BACAP.

1. The planning horizon is discretized into equal sized periods.
2. The quay crane relocating time is negligible, since the length of each period is much larger than such displacement time.
3. The quay length is discretized into equal-sized sections.
4. Each section of the quay is occupied by at most one vessel in each time period.
5. Each crane can be assigned to at most one vessel, in each period of time.

6. Each vessel has a minimum and maximum number of cranes that must be assigned to its service.

7. The number of quay cranes assigned to a vessel does not change during its service.

8. The service on a vessel by quay cranes starts upon its berthing at the terminal and is not interrupted until it departs.

Several problem variations have also been studied, to model different applications features. Particular attention has been devoted to questions such as assigning a time-variant number of quay cranes to a vessel [ Çağatay Iris et al., 2015], the assignment of specific quay cranes [Türkoğulları et al., 2014], and the continuous model of the quay space [Li et al., 2015]. The precise classification and different taxonomies of the related work is beyond the scope of this paper, and the interested reader is referred to the careful surveys of Stahlbock and Voß [2008], Bierwirth and Meisel [2010], and Bierwirth and Meisel [2015].

We describe next the generalized set partitioning problem (GSPP) model for the BACAP presented by Çağatay Iris et al. [2015]. Let $V$ be the set of vessels, $T$ be the set of time slots in the horizon, $L$ be the set of berthing positions in the quay, and $K$ be the number of available QCs.

Also define the set of berthing time/position combinations $P = T \times L$, and $\Omega$ as the complete set of feasible allocations: each element in $\Omega$ assigns a position in the quay, time slots and a number of cranes to serve a single vessel. Note that $|\Omega| \leq (|V| \times |P| \times K)$ since feasible assignments respect each vessel requirements in a problem instance.

Decision variables $y \in \mathbb{B}^{\Omega}$ indicate which individual assignments are used in the solution. The cost $c_j$ of an individual assignment $y_j$, for $j \in \Omega$, consists of a linear combination of QC usage in that allocation and deviations from parameters on the desired position on the quay, and expected starting and finishing times for the service. For the sake of generality in our contribution, we omit the exact expression for the cost $c_j$, which varies slightly in related work.

Finally, the coefficient matrices are as follows. $A \in \mathbb{B}^{V \times \Omega}$ associates each column with a single vessel: $a_{ij}$ is equal to one if column $j$ refers to an assignment for vessel $i$; otherwise, it is equal to zero. $B \in \mathbb{B}^{P \times \Omega}$ represents berths as combinations of time intervals and quay positions: $b_{pj}$ is one iff the given pair of (time, space) positions corresponding to $p \in P$ is used in the assignment $y_j$. An element of $Q \in \mathbb{Z}^{T \times \Omega}$ determines how many QCs are used by $y_j$ in time period $t$. Then, the BACAP formulation described by Çağatay Iris et al. [2015] corresponds to

$$z = \min \left\{ \sum_{j \in \Omega} c_j y_j : y \in \mathcal{P}_{\text{gspp}} \cap \mathbb{B}^{\Omega} \right\},$$

(1)
where $P_{gspp}$ denotes the polyhedral region defined by:

\[
\sum_{j \in \Omega} a_{ij} y_j = 1 \quad \forall i \in V \tag{2}
\]

\[
\sum_{j \in \Omega} b_{pj} y_j \leq 1 \quad \forall p \in P \tag{3}
\]

\[
\sum_{j \in \Omega} q_{tj} y_j \leq K \quad \forall t \in T \tag{4}
\]

\[
y_j \leq 1 \quad \forall j \in \Omega \tag{5}
\]

\[
y_j \geq 0 \quad \forall j \in \Omega \tag{6}
\]

Set partition constraints (2) ensure that all vessels are served by exactly one assignment, while set packing in (3) forbid overlapping of vessel assignments in each single time/space slot. Inequalities (4) guarantee that QCs availability in the terminal is respected. Constraints (5) and (6) correspond to the linear relaxation of the binary programming formulation.

The description of $P_{gspp}$ contains exponentially many variables with respect to the input size. On the one hand, this suggests the use of column generation algorithms, such as Saadaoui et al. [2015] describe for the BAP subproblem. On the other hand, our work and that of Çagatay Iris et al. [2015] and Türkoğulları et al. [2014] present successful computational experiments using algorithms building on the a priori enumeration of variables, followed by the solution of the resulting model by an integer linear programming solver. A key point in this approach, as we describe in Section 4, is the effectiveness of variable reduction techniques, such as those described by Çagatay Iris et al. [2015] and the one we introduce in the next section. Even the largest benchmark instances of the BACAP used by Çagatay Iris et al. [2015] and Meisel and Bierwirth [2009] can be solved using the memory amount of a current workstation.

3 Combinatorial relaxations and algorithms

In the following, we denote two assignments for different vessels as compatible if they have no overlap in berthing time and space. Equivalently, representing the two assignments in a Cartesian plane (with time and space coordinates), they are compatible iff the corresponding rectangles do not intersect each other.

We start by introducing in Section 3.1 two combinatorial relaxations for the GSPP formulation of the BACAP, which yield dual bounds to the optimal value of the objective $z$ in (1). Then, those results are extended into a preprocessing method in Section 3.2, and a matheuristic to find approximate solutions to the problem, in Section 3.3.

3.1 Matchings in two interesting graphs

We construct next two simple, undirected graphs, representing a subset of the enumerated assignments. Let $\Omega_i \subseteq \Omega$ be the subset of assignments corresponding to a given vessel $i \in V$. 
We define the graph $G_1(V, E_1)$, with a vertex for each vessel. The set $E_1$ includes an edge $(i, j)$ if the individual assignments of best cost for vessels $i$ and $j$ are not compatible with each other. Let $c'_j$ denote the minimum cost assignment for vessel $j$; that is, $c'_j = \min\{c(y_j) : y_j \in \Omega_j\}$. Analogously, let $c''_j$ be the second minimum cost assignment for $j$. The cost $c_1(i, j)$ of an edge in $G_1$ is defined by the least difference among such costs, for the corresponding vessels $i$ and $j$. That is: $c_1(i, j) \triangleq \min\{(c'_i - c'_j), (c''_i - c''_j)\}$. Then, the following bound on the cost of any feasible solution holds.

**Theorem 1.** Let $M \subseteq E_1$ denote a maximum weighted matching in $G_1$, and $w(M) = \sum_{e \in M} c_1(e)$ be its weight. Then $LB_1 \triangleq w(M) + \sum_{j \in V} c'_j$ is a lower bound to the optimal value $z$ in (1).

**Proof.** The selection of the best individual assignments for each vessel corresponds to relaxing constraints (3) and (4). Therefore, this is a trivial lower bound to the cost of any feasible solution, and amounts to $\sum_{j \in V} c'_j$.

Starting with the trivial selection of best individual assignments, the weight of an edge $(i, j) \in E_1$ corresponds to the minimum cost increase due to exchanging one such assignment for the second best. Clearly, this new pair of assignments for vessels $i$ and $j$ can still be infeasible, but the sum of their costs is a lower bound to the cost of any compatible assignment for these vessels.

Note that we cannot imply that the accumulated costs of edges incident to a same vertex are necessary, because the graph does not provide information about which of the extremes of an edge takes the second best assignment; it is even possible that, following such an exchange, other edges might not exist. However, one can consider any matching in $G_1$, corresponding to disjoint pairs of vessels whose best assignments are not compatible. Therefore, the weight of any matching is a required cost increase over $\sum_{j \in V} c'_j$, implied by the pairwise overlap of the corresponding individual assignments. In particular, a maximum weighted matching corresponds to the strongest such bound in $G_1$.

Our second dual bound strengthens the information on the cost of compatible assignments between pairs of vessels. Let $G_2(V, E_2)$ denote a complete graph, with a vertex for each vessel. Define the cost $c_2(i, j)$ of an edge in $E_2$ as the cheapest compatible assignments for vessels $i$ and $j$, that is: $c_2(i, j) \triangleq \min\{c(y_i) + c(y_j) : y_i \in \Omega_i, y_j \in \Omega_j, y_i$ and $y_j$ are compatible}. Then, we have the following result.

**Theorem 2.** Let $M \subseteq E_2$ be a maximum weighted matching in $G_2$. Then, $LB_2 \triangleq \sum_{e \in M} c_2(e)$ is a lower bound to the optimal value $z$ in (1).

**Proof.** The weight of a single edge $(i, j) \in E_2$ is the sum of the minimum cost assignments for vessels $i$ and $j$, maintaining their non-overlapping constraints. A selection of edges not sharing a vertex (i.e. a matching) thus corresponds to pairing up vessels and determining their best compatible assignments, which is required in any solution satisfying (3). Therefore, the weight of any matching in $G_2$ is a lower bound to the cost of a feasible solution, since this clearly relaxes constraints regarding the overlap of unpaired vessels. A maximum weighted matching thus provides the strongest such bound in $G_2$.

Although this result holds for any number of vessels, it would be unnecessarily weaker for odd $|V|$, since some vertex would not be covered by the matching, thus not contributing to
the lower bound. To circumvent this, we simply add to \( G_2 \) an artificial vertex \( s \), with edges to every other vertex \( i \), with costs \( c_2(s, i) = \min\{c(y_i) : y_i \in \Omega_i\} \).

We conclude with a result on the relative strength of the bounds obtained in the two relaxations. Note that, in the simple case where all the best individual assignments are pairwise compatible with each other, we verify: (i) the graph \( G_1 \) has no edges, and the bound \( LB_1 \) corresponds to trivial bound \( \sum_{j \in V} c'_j \); (ii) any perfect matching \( M \) in the graph \( G_2 \) has maximum weight, amounting to the sum of costs of the best individual assignments. Therefore, the bounds are equal: \( LB_2 = \sum_{(i, j) \in M} c_2(i, j) = \sum_{(i, j) \in M} (c'_i + c'_j) = \sum_{u \in V} c'_u = LB_1 \). We show below that the second bound can never be weaker than the first. Specific cases where it is stronger are intuitive: it suffices to have a pair of vertices for which there are no compatible assignments employing the best allocation for one of them.

**Theorem 3.** The lower bound attained from graph \( G_2(V, E_2) \) is stronger than that from graph \( G_1(V, E_1) \); i.e. for any given problem instance, \( LB_2 \geq LB_1 \) holds.

**Proof.** We assume without loss of generality that the number of vertices \( |V| \) is even, since, as described above, we suggest including an artificial vertex in the corresponding graph instances with odd \( |V| \) to get a stronger bound. Suppose there were an instance where \( LB_1 > LB_2 \). We build a matching in \( G_2 \) with cost at least \( LB_1 \), showing that the hypothesis is absurd.

For any edge \((i, j) \in E_1\), we can compare the cost functions in \( G_1 \) and \( G_2 \); recall that the latter graph is complete. By definition, \( c_1(i, j) \) corresponds to the minimum cost increase implied by the overlap of the best individual assignments for \( i \) and \( j \), while \( c_2(i, j) \) corresponds to the actual sum of the costs of the best compatible assignments. It follows that:

\[
c_2(i, j) \geq c'_i + c'_j + c_1(i, j)
\]  

(7)

Let \( M_1 \subseteq E_1 \) be a maximum weighted matching in \( G_1 \), which thus yields the lower bound \( LB_1 \) from that graph. We define the analogous set of edges in \( G_2 \) as \( M_2 = \{(i, j) \in E_2 : \text{there exists the edge } (i, j) \in M_1\} \). The set \( M_2 \) is a matching in \( G_2 \), by construction.

We distinguish two cases. If \( M_1 \) is perfect, then \( M_2 \) is perfect as well since both graphs have the same vertex set. We can infer about their weights:

\[
w(M_2) = \sum_{(i, j) \in M_2} c_2(i, j) \geq \sum_{(i, j) \in M_2} (c'_i + c'_j + c_1(i, j)) = \sum_{(i, j) \in M_1} c_1(i, j) + \sum_{u \in V} c'_u = LB_1,
\]

(8)

where the first inequality holds by (7), and the second equality is true because the matchings are perfect.

If \( M_1 \) is not perfect, there are pairs of vertices \((x, y)\) not covered by \( M_1 \). By hypothesis, \( M_1 \) has maximum weight; hence \((x, y) \notin E_1\), i.e. the individual assignments of least cost for \( x \) and \( y \) are compatible. Therefore, the edge in \( G_2 \) corresponding to each such pair \((x, y)\) has cost \( c_2(x, y) = c'_x + c'_y \). We can extend \( M_2 \) to a perfect matching \( M'_2 \) in \( G_2 \) by arbitrarily connecting pairs of vertices not yet covered by \( M_2 \). Let \( C \) denote the set of edges selected this way, such that \( M'_2 = M_2 \cup C \), and \( \sum_{(x, y) \in C} c_2(x, y) = \sum_{(x, y) \in C} c'_x + c'_y \). Then, analogously to
the previous case, we have:

\[
w(M'_2) = \sum_{(i,j) \in M_2} c_2(i,j) + \sum_{(x,y) \in C} c_2(x,y) \\
\geq \sum_{(i,j) \in M_2} (c'_i + c'_j + c_1(i,j)) + \sum_{(x,y) \in C} c_2(x,y) \\
= \sum_{(i,j) \in M_2} c_1(i,j) + \sum_{u \in V} c'_u \\
= LB_1,
\]

where the last equalities hold because \(M'_2\) covers all vertices.

Therefore, the matchings in \(G_2\) built in both cases (8) and (9) have weight at least \(LB_1\), providing a lower bound on \(LB_2\), which is defined as the maximum weight of a matching in \(G_2\). Since we start with a general input instance, the hypothesis that \(LB_1 > LB_2\) could hold is absurd, and we always verify that \(LB_2 \geq LB_1\).

### 3.2 Preprocessing method for variable reduction

The previous results yield dual, lower bounds on the optimal value \(z\) of problem (1), and can also be extended to a preprocessing method. The goal is to fix at null value (or, equivalently, reduce) a number of decision variables in the resulting model after the enumeration of feasible assignments, while preserving the optimal, exact solution.

It is worth remarking that, since this technique is applied prior to the model optimization, such a proposal can be integrated with any approach based on enumerating the variables of the GSPP formulation and solving the resulting model with an integer linear programming algorithm. This strategy has already been adopted by [Çağatay Iris et al., 2015]. In that work, the authors use lower bounds implied by probing the selection of a single assignment or a pair of assignments for two different vessels. In Section 4, we compare their method with the one we propose in this paper.

The next result assumes that an upper bound to \(z\) is available. In the computational experiments presented in this paper, we use the solution value provided by the warm starting heuristic described by [Çağatay Iris et al., 2015].

First, we temporarily assume that a given assignment \(y_k \in \Omega_k\) is fixed in the solution. We define the complete graph \(G_{2,k}(V \setminus \{k\}, E_{2,k})\). The corresponding edge costs \(c_{2,k}\) regard the best compatible assignments for two vessels, which are also compatible with \(y_k\). That is: \(c_{2,k}(i,j) = \min\{c(y_i) + c(y_j) : y_i \in \Omega_i, y_j \in \Omega_j, y_i \text{ and } y_j \text{ are compatible with each other and with } y_k\}\). Finally, we evaluate the increase on the lower bound \(LB_2\) from Theorem (2) implied by fixing the assignment \(y_k\): if the new lower bound exceeds a known upper bound, we conclude that this assignment can not be part of an optimal solution. The result is summarized as follows.

**Proposition 1.** Let \(LB_{2,k}\) denote the lower bound from Theorem 2 determined over \(G_{2,k}\). Given any upper bound \(UB\) to \(z\) in (1), if \(c(y_k) + LB_{2,k} > UB\), then there is no optimal solution which includes the assignment \(y_k \in \Omega_k\), and the corresponding variable/column can be removed from the model.
Therefore, we have an iterative algorithm for removing unnecessary variables in the model, while preserving all optimal solutions of the problem. For each feasible assignment in the GSPP formulation, one need only to evaluate the new lower bound as depicted above.

Note that an analogous method can be derived from Theorem 1. Nevertheless, it follows immediately from Theorem 3 that it cannot be stronger, i.e. it cannot remove a column which the result in Proposition 1 does not.

3.3 Column ranking matheuristic

We introduce next an algorithm belonging to the class of matheuristics, or model-based heuristics, which integrate heuristics and exact methods of mathematical programming [Ribeiro and Maniezzo, 2015, Maniezzo et al., 2010]. Specifically, we employ the combinatorial relaxation information to obtain a reduced model, which is optimized next with an integer linear programming solver. We remark that previous strategies in the matheuristics literature include solving a reduced model: e.g. after the heuristic removal of variables [Stefanello et al., 2015, Fanjul-Peyro and Ruiz, 2011], or to partially optimize components of a previous solution [Lalla-Ruiz and Voss, 2016].

First, note that every solution to the formulation consists of only $|V| < \Omega$ columns. One could wonder if there would be an ideal criterion for classifying columns, such that high quality solutions could be consistently achieved using only a fraction of the best ranked columns. In this context, we discard the optimality certificate, and seek a high quality solution in reduced computation time.

The core of the method we propose is Algorithm 1, which ranks and selects a subset $\Omega_F$ of columns from the GSPP model. The selection builds on the combinatorial bound $LB_{2,k}$ from Proposition 1 (denoted by $\Delta$, computed on the loop starting at line 3). The set of selected columns $\Omega_F$ corresponds to a subset of the polyhedron $P_{gspp}$, and optimizing over it provides an upper bound $\bar{z} \geq z$ to the original problem (1). As we indicate in our computational results, this bound can match the optimal value for benchmark instances of the BACAP even when using a relatively small fraction of columns.

The algorithm parameters are as follows.

$\sigma$ : the percentage of the best ranked columns which should be included in the final model;

$\mu$ : the minimum number of columns (when available) corresponding to each vessel, which the algorithm should ensure in the final model.

The latter parameter $\mu$ ensures that each vessel has a number of assignment options (as selected in the loop starting at line 13), while $\sigma$ controls the selection of columns among those implying the best dual bounds (loop starting at line 8). It would be natural to consider algorithm variations, e.g. selecting an exact number of columns, or performing a statistical study of the parameters. Both tasks could be approached in future work.

We add that, since we give up on the optimality certificate, it is not necessary to reach a null duality gap in the resulting model to end the algorithm. Such a strategy may be
interesting to meet runtime requirements in a BACAP application.

Algorithm 1: column ranking

| Input | : initial set of columns Ω, number of vessels | | \[V\] |
| Output | : set \(Ω_F\) of columns selected for the final model |
| Parameters | : minimum percentage \(σ\) of the best ranked columns, minimum number of columns per vessel \(μ\) |

\(Ω_F \leftarrow \emptyset\)

\(Δ(k) \leftarrow 0\) for each \(y_k \in Ω\)

\textbf{foreach} \(y_k \in Ω\) \textbf{do}

\(G_{2,k}\) be the graph of compatible assignments, defined in Proposition 1

\(\text{Let } M\) be a maximum weighted matching in \(G_{2,k}\)

\(//\) lower bound implied by using this column; see Prop. 1

\(Δ(k) \leftarrow c(y_k) + \sum_{e \in M} c_2(e)\)

\(\) Let \(L\) denote the list of columns in Ω, ordered by increasing values of \(Δ\)

\textbf{while} \(|Ω_F|/|Ω| < σ\) \textbf{do}

\(\) Let \(d\) be the least \(Δ\) value of a column in \(L\)

\textbf{foreach} column \(y_k \in L\) \textbf{with} \(Δ(k) = d\) \textbf{do}

\(Ω_F \leftarrow Ω_F \cup \{y_k\}\)

\(L \leftarrow L \setminus \{y_k\}\)

\textbf{foreach} vessel \(i = 1, \ldots, |V|\) \textbf{do}

\textbf{while} \(|\{y \in Ω_F : y\ is\ an\ assignment\ for\ i\}| < μ\) \textbf{and} \(L\ contains\ a\ column\ referring\ to\ i\) \textbf{do}

\(\) Let \(y_i \in L\) be a column referring to \(i\), of least \(Δ\)

\(Ω_F \leftarrow Ω_F \cup \{y_i\}\)

\(L \leftarrow L \setminus \{y_i\}\)

\textbf{return} \(Ω_F\)

4 Computational results

The goals of the computational evaluation we present are twofold: to compare the impact of the preprocessing algorithm we describe with those previously available in the literature, and to assess the strength of the solutions provided by the matheuristic we introduce. We consider the same BACAP benchmark instances used by Meisel and Bierwirth [2009] and Çağatay Iris et al. [2015] to evaluate the efficiency of the proposed algorithms. There are ten instances of three different sizes, with 20 (small), 30 (medium) and 40 (large) vessels.

All algorithms were implemented in C++ language using Gurobi solver version 6.5. To compute maximum weighted matchings, with the blossom shrinking algorithm by Edmonds [Edmonds, 1965], we use the efficient implementation available in the open source Library for Efficient Modeling and Optimization in Networks (LEMON) [Dezső et al., 2011]. The time complexity of that implementation is \(O(mn \log n)\) in the worst case, where \(n\) is the number of graph vertices and \(m\) is the number of edges.

The solver runtime used in the matheuristic experiments was limited to 1800 seconds. The experiments corresponding to the techniques proposed by [Çağatay Iris et al., 2015] were not time limited. It is important to highlight that all results we present concerning the work by [Çağatay Iris et al., 2015] were evaluated with our own implementation of their algorithms. Moreover, the numbers concerning their results may have slight variations when compared to the original work, which we have concluded to be explained by numeric precision matters.
All experiments were run in a machine with an Intel Core i7 4790K (4.00 GHz) CPU and 16GB of RAM.

Table 1 compares the performance of the variable reduction technique proposed in this work and those by [Çagatay Iris et al., 2015]. We present the final number of columns after each reduction phase. $|\Omega|$ indicates the original number of columns after the enumeration of feasible assignments. $|\Omega_1|$ is the final number of columns after two preprocessing methods proposed by [Çagatay Iris et al., 2015], which are based on simple column removal tests considering comparisons with an upper bound. The comparisons are made with each column cost itself, or by fixing it and choosing the best assignments of others vessels, ignoring feasibility conditions.

| Instance | ID | $|\Omega|$ | $|\Omega_1|$ | $|\Omega_2|$ | $R_I(\%)$ | $|\Omega_3|$ | $R_C(\%)$ |
|----------|----|----------|----------|----------|-----------|----------|-----------|
| 20       | 1  | 316278   | 102657   | 60114    | 80.99     | 54582    | 82.74     |
|          | 2  | 338438   | 19202    | 1349     | 96.28     | 11318    | 92.75     |
|          | 3  | 252618   | 94083    | 77457    | 69.34     | 60941    | 75.88     |
|          | 4  | 390624   | 106970   | 74730    | 80.87     | 73199    | 81.26     |
|          | 5  | 275238   | 33857    | 23689    | 91.39     | 18852    | 93.15     |
|          | 6  | 305485   | 28133    | 1349     | 96.28     | 1345     | 99.60     |
|          | 7  | 252618   | 94083    | 77457    | 69.34     | 60941    | 75.88     |
|          | 8  | 350703   | 61200    | 41789    | 92.62     | 40448    | 88.76     |
|          | 9  | 286683   | 91561    | 65926    | 77.00     | 64592    | 77.47     |
|          | 10 | 359320   | 112444   | 68.71    | 83589     | 76.74     |

The column number after the main reduction methods are as follows. $|\Omega_2|$ is the final number of columns after applying the two variable reduction methods proposed by [Çagatay Iris et al., 2015] over the $|\Omega_1|$ set. Finally, $|\Omega_3|$ stands for the total number of columns.

$R_I = \frac{|\Omega - |\Omega_2|}{|\Omega|}$; $R_C = \frac{|\Omega - |\Omega_3|}{|\Omega|}$; $R_I_{|V|}$ stands for the arithmetic mean of $R_I$ with $|V|$ vessels; $R_{C|V|}$ denotes the arithmetic mean of $R_C$ with $|V|$ vessels.

The column number after the main reduction methods are as follows. $|\Omega_2|$ is the final number of columns after applying the two variable reduction methods proposed by [Çagatay Iris et al., 2015] over the $|\Omega_1|$ set. Finally, $|\Omega_3|$ stands for the total number of columns.
removing the successive application of the two techniques regarded in Ω₁, the first variable probing method (trivial one) of [Çağatay Iris et al., 2015] and, at last, the combinatorial reduction we propose.

This table also presents final percentage results, \( R_I \) and \( R_C \), of the preprocessing methods considering \( |Ω_2| \) and \( |Ω_3| \), respectively. In average, the variable reduction is higher for the small and medium instances. The reduction can be as effective as 99.60%, for instance 2. It can be noticed that the reduction is slightly larger after applying the combinatorial relaxation technique (see results in bold) for all instances except the number 2. Another fact worth highlighting in these results is that the combinatorial reduction technique is more significant, in average, for the larger instances (increase of 2.14 percentage points) if compared to the techniques proposed by [Çağatay Iris et al., 2015], than for the smaller (increase of 1.95 pp) or medium ones (increase of 1.75 pp).

Table 2: Results concerning quality of solutions and runtime

| Instance | Matheuristic results |Runtime | Çağatay Iris et al. [2015] | Efficiency |
|----------|----------------------|--------|--------------------------|------------|
| \(|V|\)  | ID \( \bar{z} \)  \( \Omega_F(\%) \) \( \bar{z} \) Gap\(_{OPT}(\%) \) | Matheuristic \( T_C \) (s) \( T_{CF} \) (s) \( T_I \) (s) \( T_{IF} \) (s) \( T_E(\%) \) |  |
| 20      | 1 89,00 60,84 89,00 0,00 32,44 56,65 14,84 64,28 88,13 |  |
|         | 2 56,20 100,00 56,20 0,00 1,01 1,09 0,09 0,16 68,12 |  |
|         | 3 85,70 59,09 85,70 0,00 9,63 73,19 0,79 114,73 63,79 |  |
|         | 4 81,80 47,86 81,80 0,00 8,58 39,24 0,16 60,22 65,16 |  |
|         | 5 59,20 99,13 59,20 0,00 2,62 7,99 0,16 7,59 105,27 |  |
|         | 6 59,20 100,00 59,20 0,00 1,53 6,73 0,16 5,09 122,22 |  |
|         | 7 75,20 94,32 75,20 0,00 6,42 21,33 2,56 21,41 107,08 |  |
|         | 8 61,40 73,20 61,40 0,00 5,39 22,10 0,95 54,32 83,67 |  |
|         | 9 79,00 52,22 79,00 0,00 16,37 45,45 3,77 54,32 83,67 |  |
|         | 10 101,00 45,56 101,00 0,00 26,05 76,27 3,13 214,43 35,57 |  |

\( \Omega_F = \frac{\Omega_M}{\Omega_2}; \Omega_M \) denotes the number of variables after the matheuristic filter; \( T_E = \frac{T_{CF}}{T_{IF}} \); \( \bar{T}_{E(V)} \) corresponds to the geometric mean of \( T_E \) for instances with \(|V| \) vessels.

We executed different experiments to evaluate the matheuristic algorithm. Several param-
eter combinations were tested: $\sigma \in \{10, 20, 25, 33\}$ and $\mu \in \{1000, 2000\}$. The configuration $(\sigma = 10, \mu = 2000)$ achieved the best results, which are reported in Table 2. For each instance, we present the optimal solution value $z$, the column percentage results after matheuristic filter ($\Omega_F$), followed by the cost $\bar{z}$ and the total GAP$_{OPT}$ (from the optimal value $z$) of the best solution found through the matheuristic. Next we present the partial runtime of applying our variable reduction technique ($T_C$) and the one by [Çağatay Iris et al., 2015] ($T_I$), and the total runtime for solving the mathematical model using the Gurobi solver: $T_{CF}$ and $T_{IF}$, respectively. Lastly, we show the runtime improvement of our proposal ($T_E$) compared to that by [Çağatay Iris et al., 2015].

It can be seen from these results that the proposed methodology is able to find the known optimal solution in 83% of instances, while the gap is below 4% for those solutions which are not optimal. Note that the runtime efficiency of our methodology is inferior to the literature in half of the instances with 20 vessels and in one of the medium instances. Nevertheless, these results do not have a significant impact because the corresponding execution times are at most 44s, and the time difference between the methods does not exceed 2s. For example, while instance 2 is solved instantaneously by the technique from the literature, it spends 1.09s using the matheuristic. Therefore, our proposal loses on the average efficiency for small instances.

On the other hand, observing the results for medium and large instances in the benchmark, it is clear that the matheuristic results are significantly superior to those from literature. The known optimal solution is found in all instances with 30 vessels, spending on average 59% of the time spent by the literature method. For more difficult instances, with 40 vessels, the optimal solution is obtained in half of the cases, while very good solutions can be achieved using about 30% of the time consumed by the method presented by [Çağatay Iris et al., 2015].

Considering instance 23, for example, the optimal solution is obtained 8 times faster by the matheuristic algorithm. In instance 28, the matheuristic methodology consumes less than 6% of literature algorithm runtime to find a solution with a gap less than 3% of the optimal value. These results suggest that the column classification using the combinatorial bound with the matheuristic is a good selection criterion: the model size decreases drastically, the solution runtime reduces consistently in instances with more challenging dimensions, while keeping solutions up to 4% of the optimum.

5 Concluding remarks

We investigate in this work a maritime logistics problem, formulated in integer linear programming as a generalization of the set partitioning problem. The task is to obtain an individual allocation of each vessel scheduled to arrive on a container terminal, assigning them a physical space, a period of time and a number of quay cranes for their service.

We introduce combinatorial relaxations of the formulation, solving weighted matching problems in two suitable graphs representing a subset of feasible allocations. Besides intrinsically interesting for the dual bounds they provide, the relaxations are also extended into a variable reduction technique and a column ranking matheuristic. The latter algorithm builds a reduced model using only the best ranked columns, according to an enhanced combinatorial bound.

Computational assessments of the proposed methodology and comparisons with the methods described as state of the art indicate its effectiveness. The matheuristic could reach the
known optimal solutions on 83% of the benchmark instances of the problem; otherwise, the solutions were within 4% of the optimal value, while using an average of 30% of the time required by methods recently introduced in the literature.

References

Christian Bierwirth and Frank Meisel. A survey of berth allocation and quay crane scheduling problems in container terminals. *European Journal of Operational Research*, 202(3):615–627, 2010. ISSN 0377-2217. doi: http://dx.doi.org/10.1016/j.ejor.2009.05.031. URL http://www.sciencedirect.com/science/article/pii/S0377221709003579.

Christian Bierwirth and Frank Meisel. A follow-up survey of berth allocation and quay crane scheduling problems in container terminals. *European Journal of Operational Research*, 244(3):675–689, 2015. ISSN 0377-2217. doi: http://dx.doi.org/10.1016/j.ejor.2014.12.030. URL http://www.sciencedirect.com/science/article/pii/S0377221714010480.

Çağatay Iris, Dario Pacino, Stefan Ropke, and Allan Larsen. Integrated berth allocation and quay crane assignment problem: Set partitioning models and computational results. *Transportation Research Part E: Logistics and Transportation Review*, 81:75–97, 2015. ISSN 1366-5545. doi: http://dx.doi.org/10.1016/j.tre.2015.06.008. URL http://www.sciencedirect.com/science/article/pii/S1366554515001301.

Balázs Dezső, Alpár Jüttner, and Péter Kovács. LEMON – an Open Source C++ Graph Template Library. *Electronic Notes in Theoretical Computer Science*, 264(5):23–45, 2011. ISSN 1571-0661. doi: 10.1016/j.entcs.2011.06.003.

Jack Edmonds. Paths, trees, and flowers. *Canadian Journal of Mathematics*, 17:449–467, 1965. URL http://cms.math.ca/10.4153/CJM-1965-045-4.

Luis Fanjul-Peyro and Rubén Ruiz. Size-reduction heuristics for the unrelated parallel machines scheduling problem. *Computers & Operations Research*, 38(1):301–309, 2011. ISSN 0305-0548. doi: http://dx.doi.org/10.1016/j.cor.2010.05.005. URL http://www.sciencedirect.com/science/article/pii/S0305054810001188. Project Management and Scheduling.

Lutz Hoffarth and Stefan Voß. *Operations Research Proceedings 1993*, chapter Berth allocation in a container terminal — Development of a decision support system (in German), pages 89–95. Springer Berlin Heidelberg, Berlin, Heidelberg, 1994. ISBN 978-3-642-78910-6. doi: 10.1007/978-3-642-78910-6_28. URL http://dx.doi.org/10.1007/978-3-642-78910-6_28.

Eduardo Lalla-Ruiz and Stefan Voß. Popmusic as a matheuristic for the berth allocation problem. *Annals of Mathematics and Artificial Intelligence*, 76(1):173–189, 2016. ISSN 1573-7470. doi: 10.1007/s10472-014-9444-4. URL http://dx.doi.org/10.1007/s10472-014-9444-4.

Feng Li, Jiuh-Biüng Sheu, and Zi-You Gao. Solving the continuous berth allocation and specific quay crane assignment problems with quay crane coverage range. *Transportation Science*, 49(4):968–989, 2015. doi: 10.1287/trsc.2015.0619. URL http://dx.doi.org/10.1287/trsc.2015.0619.
Vittorio Maniezzo, Thomas Stützle, and Stefan Voß. *Matheuristics: hybridizing metaheuristics and mathematical programming*, volume 10 of *Annals of Information Systems*. Springer US, 2010. ISBN ISBN: 978-1-4419-1305-0.

Frank Meisel and Christian Bierwirth. Heuristics for the integration of crane productivity in the berth allocation problem. *Transportation Research Part E: Logistics and Transportation Review*, 45(1):196 – 209, 2009. ISSN 1366-5545. doi: http://dx.doi.org/10.1016/j.tre.2008.03.001. URL http://www.sciencedirect.com/science/article/pii/S1366554508000768.

Young-Man Park and Hwan Kap Kim. A scheduling method for berth and quay cranes. *OR Spectrum*, 25(1):1–23, 2003. ISSN 1436-6304. doi: 10.1007/s00291-002-0109-z. URL http://dx.doi.org/10.1007/s00291-002-0109-z.

Celso C. Ribeiro and Vittorio Maniezzo. Preface to the special issue on matheuristics: Model-based metaheuristics. *International Transactions in Operational Research*, 22(1):1–1, 2015. ISSN 1475-3995. doi: 10.1111/itor.12142. URL http://dx.doi.org/10.1111/itor.12142.

Yousra Saadaoui, Nitish Umang, and Emma Frejinger. A column generation framework for berth scheduling at port terminals. Technical Report CIRRELT-2015-15, Université de Montréal, 2015. URL http://www.cirrelt.ca/DocumentsTravail/CIRRELT-2015-15.pdf.

R. Stahlbock and S. Voß. Operations research at container terminals: A literature update. *OR Spectrum*, 30(1):1–52, 2008. doi: 10.1007/s00291-007-0100-9. URL www.scopus.com/inward/record.uri?eid=2-s2.0-36248971841&partnerID=40&md5=4588095e77c11800fb.

Fernando Stefanello, Olinto C. B. de Araújo, and Felipe M. Müller. Matheuristics for the capacitated p-median problem. *International Transactions in Operational Research*, 22(1):149–167, 2015. ISSN 1475-3995. doi: 10.1111/itor.12103. URL http://dx.doi.org/10.1111/itor.12103.

Y.B. Türkoğulları, Z.C. Taşkin, N. Aras, and İ.K. Altunel. Optimal berth allocation and time-invariant quay crane assignment in container terminals. *European Journal of Operational Research*, 235(1):88 – 101, 2014. ISSN 0377-2217. doi: http://dx.doi.org/10.1016/j.ejor.2013.10.015. URL http://www.sciencedirect.com/science/article/pii/S0377221713008382.

Ilaria Vacca, Matteo Salani, and Michel Bierlaire. An exact algorithm for the integrated planning of berth allocation and quay crane assignment. *Transportation Science*, 47(2):148–161, 2013. doi: 10.1287/trsc.1120.0428. URL http://pubsonline.informs.org/doi/abs/10.1287/trsc.1120.0428.