Two-stage column generation and applications in container terminal management

Ilaria Vacca, Transport and Mobility Laboratory, EPFL
Matteo Salani, Transport and Mobility Laboratory, EPFL
Michel Bierlaire, Transport and Mobility Laboratory, EPFL

Conference paper STRC 2008

STRC | **8th Swiss Transport Research Conference**
Monte Verità / Ascona, October 15-17, 2008

Two-stage column generation and applications in container terminal management

Ilaria Vacca
TRANSP-OR, EPFL
Lausanne

Phone: +41 21 693 2432
Fax: +41 21 693 8060
ilaria.vacca@epfl.ch

Matteo Salani
TRANSP-OR, EPFL
Lausanne

Phone: +41 21 693 2432
Fax: +41 21 693 8060
matteo.salani@epfl.ch

Michel Bierlaire
TRANSP-OR, EPFL
Lausanne

Phone: +41 21 693 2537
Fax: +41 21 693 8060
michel.bierlaire@epfl.ch

Abstract

Seaport container terminals are source of many interesting large-scale optimization problems, which arise in the management of operations at several decision levels.

In this paper we firstly illustrate a recently proposed framework, called two-stage column generation (Salani and Vacca, 2008), which is suitable for a particular class of large-scale problems, where the structure of the original formulation, namely the presence of a combinatorial number of decision variables, does not allow for straightforward reformulation. The main idea of their approach is to start from a meaningful subset of original variables, for which the problem remains feasible, apply the DW decomposition to this subset of variables, solve the resulting reformulation with standard column generation and then iteratively perform the explicit pricing on original compact-formulation variables, retracing back the reformulation and using complementary-slackness conditions.

The two-stage column generation framework is suitable to be applied to the Tactical Berth Allocation Problem (TBAP) with Quay Crane Assignment, an integrated decision problem which occurs in the management of the quayside resources in container terminals. A compact formulation of TBAP has been recently presented by Giallombardo et al. (2008); in this paper, we propose a Dantzig-Wolfe reformulation of the problem and we show that standard column generation cannot cope with the pricing induced by such reformulation, because of the particular structure of the problem. It makes sense therefore to use two-stage column generation and we illustrate the application of the method to TBAP. Conclusions and future research directions are discussed.

Keywords

column generation, dantzig-wolfe decomposition, complementary slackness, container terminals, berth allocation, quay crane assignment

1 Introduction

Maritime transport has been constantly developing in Europe and worldwide in the last decades, with containerized sea-freight transport gaining a leading role in the exchange of goods (UNCTAD, 2007). This trend is confirmed by the figures in Table 1, which reports the volume of container traffic in TEUs (Twenty feet Equivalent Unit) of the busiest container seaports in the World and in Europe over the last three years.

World		2005	2006	2007
1	Singapore	23,190,000	24,800,000 (+6.94%)	27,932,000 (+12.63%)
2	Shanghai	18,084,000	21,700,000 (+20.00%)	26,150,000 (+20.51%)
3	Hong Kong	22,602,000	23,230,000 (+2.78%)	23,881,000 (+2.80%)
Europe		2005	2006	2007
1	Rotterdam	9,287,000	9,690,000 (+4.34%)	10,790,000 (+11.35%)
2	Hamburg	8,087,550	8,861,804 (+9.57%)	9,900,000 (+11.72%)
3	Antwerp	6,482,030	7,018,799 (+8.28%)	8,176,614 (+16.50%)

Table 1: *Busiest ports worldwide and in Europe (TEUs traffic).*

Seaport container terminals represent nowadays a key node in the intermodal freight transport network and the optimal management of their operations is crucial for efficiently transshipping goods among different transport modes (road, rail, sea).

We can identify many optimization decision problems which arise in a container terminal (Vis and de Koster, 2003; Steenken et al., 2004; Stahlbock and Voss, 2008). The *Berth Allocation Problem* (BAP) consists in allocating ships to berths (discrete case) or to berthing segments along the quay (continuous case); constraints are given by ship's length, berth's depth, time windows and priorities assigned to the ships, favorite berthing areas, etc. The *Quay Crane Assignment Problem* (QCAP) consists in assigning a certain number of quay cranes to ships calling at the port, according to constraints on the available resources as well as time windows on the ship's arrival and stay at the port. The *Quay Crane Scheduling Problem* (QCSP) refers to the scheduling of loading and unloading moves of quay cranes on the vessels, in order to perform pre-assigned tasks on set of containers as fastest as possible; issues related to interference among quay cranes, precedence and operational constraints must be taken into account. The management of *yard operations* involves as well several decision problems: the design of storage policies at the block and bay level according to the specific features of the container (size, weight, destination, import/export etc.); the allocation, routing and scheduling of yard cranes; the design of re-marshalling policies for export containers, etc. Terminal managers are faced day by day with these problems, just to cite a few of them, and decisions have to be made at different planning levels (strategic, tactical and operational).

The standard approach used by terminal planners in solving these problems and making decisions is typically hierarchical; however, recent trends (Vacca et al., 2008) suggest that an integrated planning of related operations allows the terminal to better exploit its limited resources and to improve both the control and the efficiency on its performance.

Among others, a challenging problem is represented by the integration of the berth allocation problem with the quay crane assignment. The BAP and the QCAP are strictly correlated, yet solved hierarchically by the terminal planners. Firstly, the QCAP is solved according to the amount of quay crane hours requested by the vessels calling at the port over a given time horizon, in addition to time windows constraints on the total time spent by vessels at the terminal and limits on the available resources (namely, the total number of available quay cranes). Once a quay crane profile has been assigned to each ship, i.e. the number of quay cranes assigned per working shift to each vessel is known, the berth allocation problem can be solved, using as input the ship's handling time, which is determined indeed by the QC assignment profile. It clearly makes sense to solve these two problems simultaneously, as an integrated planning allows the terminal to better optimize its resources and increase its efficiency, although this integration, as expected, leads also to an increased complexity of the problem.

The integrated problem has been mostly studied at the operational decision level by Park and Kim (2003), Meisel and Bierwirth (2006), Imai et al. (2008), Meisel and Bierwirth (n.d.); only recently Giallombardo et al. (2008) introduced the Tactical Berth Allocation Problem (TBAP) with Quay Cranes Assignment, where authors address the problem from a tactical point of view. The intent is to support the terminal in its negotiation process with shipping lines, when the two parts discuss contracts which usually bound the amount of QC hours assigned to each ship for the following months. In particular, the proposed model aims to allow terminal managers to assign the right value to the QC profiles proposed to shipping lines, taking into account their monetary value as well as their impact on the terminal productivity. In addition to profile evaluation, the combined solution of these two problems at the tactical level optimizes the utilization of terminal resources. The authors present two mixed integer programming formulations, quadratic and linear respectively, which aim, on the one hand, to maximize the total value of chosen QC assignment profiles, and, on the other hand, to minimize the quadratic yard-related costs arising in a transshipment context. Computational results obtained through a commercial software show that the problem is hardly solvable even on small instances.

In this paper, we propose a Dantzig-Wolfe reformulation of TBAP, using as compact formulation the mixed integer linear program presented by Giallombardo et al. (2008). Our reformulation introduces the concept of *berth sequence*, which represents a feasible schedule of ships allocated to a berth, combined with a QC assignment for each scheduled ship. We show that a standard column generation approach is not efficient in this case because the resulting pricing problem, an Elementary Shortest Path Problem with Resource Constraints (ESPPRC), is unmanageable due to the huge size of its underlying network. However, we have remarked that the original problem has a specific structure, namely a combinatorial number of compact-formulation variables, which makes it suitable to be solved via *two-stage column generation*, a novel framework recently proposed by Salani and Vacca (2008) for this particular class of large-scale problems. We therefore apply the two-stage approach to TBAP, obtaining a more tractable pricing problem.

In the remainder of this paper we illustrate the general two-stage column generation approach in section 2 and we show its application to TBAP in section 3. Finally, we draw our conclusions and discuss future research directions in section 4.

2 Two-stage column generation

In this section we illustrate the two-stage column generation methodology, a novel framework introduced by Salani and Vacca (2008) for solving large-scale optimization problems.

Column generation has been traditionally embedded in branch-and-price schemes to compute good quality lower bounds for combinatorial problems reformulated through Dantzig-Wolfe (DW) decomposition (Dantzig and Wolfe, 1960). For further details, we address the reader to Nemhauser and Wolsey (1988) and Desaulniers et al. (2005).

In this context, Salani and Vacca (2008) have identified a particular class of problems in which the structure of the original formulation does not allow for straightforward reformulation.

Specifically, the presence in the original formulation of a large, possibly combinatorial, number of decision variables renders the pricing problem of the associated column generation scheme unmanageable. Applications with such characteristics are found in several domains: routing problems (Nakao and Nagamochi, 2007), scheduling (Xu and Chiu, 2001) and, as already mentioned, container terminals (Giallombardo et al., 2008). In particular, authors focus on a routing application, the Discrete Split Delivery Vehicle Routing Problem with Time Windows (DSD-VRPTW), which is used as a running example to illustrate the two-stage column generation scheme.

The basic idea of the framework is simple: they propose to consider a subset of original variables of the so called *compact* formulation, to solve the so called *extensive* formulation via standard column generation and to generate profitable original variables computing their reduced cost, by complementary-slackness conditions, in the same spirit of basic column generation. At the end of this procedure, a smaller subset of original variables has been possibly used and a subset of sub-optimal variables identified. Since the method addresses linear programs only, i.e. provides dual bounds, it is embedded in a branch-and-price scheme (Barnhart et al., 1998).

In the following sections we illustrate the general methodology in details.

2.1 Dantzig-Wolfe reformulation for integer programs

Consider the following integer linear program, the *original* or *compact formulation* (CF):

$$
\begin{aligned}
z_{IP} = \min \quad & c^T x && (1)\\
s.t. \quad & Ax \geq b, && (2)\\
& Dx \geq d, && (3)\\
& x \in \mathbb{Z}^n_+. && (4)
\end{aligned}
$$

It is assumed that conditions to use a standard column generation approach hold, namely the linear relaxation of the Dantzig-Wolfe reformulation of (1)–(4) provides better bounds than the linear relaxation of the original problem, because of the special structure of constraints $\{x \in \mathbb{Z}^n_+ : Dx \geq d\}$ which can be easily convexified.

Furthermore, we remark that the framework applies to problems which present a combinatorial number of compact integer variables x. This prevents to use the standard column generation approach, as the complexity of the resulting subproblem is unmanageable, as it will be shown later on.

Let $P = conv\{x \in \mathbb{Z}^n_+ : Dx \geq d\} \neq \emptyset$ be a bounded polyhedron. Each $x \in P$ can be represented as a convex combination of extreme points $\{p_q\}_{\{q \in Q\}}$ of P:

$$x = \sum_{q \in Q} p_q \lambda_q, \ \sum_{q \in Q} \lambda_q = 1, \ \ \lambda \in \mathbb{R}^{|Q|}_+. \tag{5}$$

The equivalent *extensive formulation* (EF) of (1)–(4) is:

$$z_{IP} = \min \quad \sum_{q \in Q} c_q \lambda_q \tag{6}$$

$$s.t. \quad \sum_{q \in Q} A_q \lambda_q \geq b, \tag{7}$$

$$\sum_{q \in Q} \lambda_q = 1, \tag{8}$$

$$\lambda \geq 0, \tag{9}$$

$$x = \sum_{q \in Q} p_q \lambda_q, \tag{10}$$

$$x \in \mathbb{Z}^n_+. \tag{11}$$

where $c_q = c^T p_q$ and $A_q = Ap_q \ \ \forall q \in Q$.

If we relax the integrality of x in (11), constraints (10) also become redundant. The resulting *master problem* (MP) is:

$$z_{MP} = \min \quad \sum_{q \in Q} c_q \lambda_q \tag{12}$$

$$s.t. \quad \sum_{q \in Q} A_q \lambda_q \geq b, \tag{13}$$

$$\sum_{q \in Q} \lambda_q = 1, \tag{14}$$

$$\lambda \geq 0. \tag{15}$$

In column generation we repeatedly solve a *restricted master problem* on a subset of variables λ, which otherwise would be an exponential number. At each iteration we add profitable variables not yet in the formulation, if any, by solving the *pricing subproblem*:

$$\min_{q \in Q} \{\tilde{c}_q := c_q - \pi A_q - \pi_0\} \tag{16}$$

where $\pi \geq 0$ is the dual vector associated to constraints (13), $\pi_0 \in \mathbb{R}$ is the dual variable associated to the convexity constraint (14) and $\tilde{c}_q$ is the reduced cost of variable λ_q.

The resulting pricing is an integer linear program, which eventually exhibits the same computational complexity of the original compact problem. Its complexity is affected by (i) the structure of constraints $\{x \in \mathbb{Z}^n_+ : Dx \geq d\}$, i.e. the nature of extreme points of Q, (ii) the number of decision variables in the compact formulation.

According to the assumptions, variables of the compact formulation are combinatorially many and this results in an unmanageable pricing, which mainly motivates the two-stage column generation approach.

2.2 Two-stage column generation: general framework

Let X be the set of compact formulation variables, $|X| = n$. The basic idea of this approach consists in starting with a subset $\bar{X} \subset X$, $|\bar{X}| = \bar{n}$ such that (CF) is feasible and iteratively add profitable variables in $\hat{X} := X \setminus \bar{X}$ using reduced costs arguments as in the standard column generation procedure. At each iteration, the resulting master problem is optimally solved using again column generation. The clear benefit of this approach is that the the associated pricing is solved over a smaller set of variables, i.e. the dimension of the vector p_q representing an extreme point is smaller. Furthermore, not all the variables $x_i \in \hat{X}$ will eventually need to be added.

Without loss of generality and for simplicity of notation we can assume that $x = [\,\bar{x} \,|\, \hat{x}\,]$, $c = [\,\bar{c} \,|\, \hat{c}\,]$, $A = [\,\bar{A} \,|\, \hat{A}\,]$ and $D = [\,\bar{D} \,|\, \hat{D}\,]$.

The *partial compact formulation* (PCF) is defined as follows:

$$\bar{z}_{IP} = \min \quad \bar{c}^T \bar{x} \tag{17}$$

$$s.t. \quad \bar{A}\bar{x} \geq b, \tag{18}$$

$$\bar{D}\bar{x} \geq d, \tag{19}$$

$$\bar{x} \in \mathbb{Z}_+^{\bar{n}}. \tag{20}$$

We remark that $\bar{z}_{IP} \geq z_{IP}$.

Let $\bar{P} = conv\{\bar{x} \in \mathbb{Z}_+^{\bar{n}} \,|\, \bar{D}\bar{x} \geq d\} \neq \emptyset$ be bounded. We can represent each $\bar{x} \in \bar{P}$ as a convex combination of extreme points $\{p_q\}_{\{q \in \bar{Q}\}}$ of $\bar{P}$:

$$\bar{x} = \sum_{q \in \bar{Q}} p_q \lambda_q, \;\; \sum_{q \in \bar{Q}} \lambda_q = 1, \;\; \lambda \in \mathbb{R}_+^{|\bar{Q}|} \tag{21}$$

By substituting $\bar{c}_q = \bar{c}^T p_q$ and $\bar{A}_q = \bar{A} p_q \;\forall q \in \bar{Q}$, we can write the equivalent *partial extensive formulation* (PEF), as seen before in (6)–(11); subsequently, by relaxing the integrality constraints on $\bar{x}$, the *partial master problem* (PMP) can be defined as follows:

$$\bar{z}_{MP} = \min \quad \sum_{q \in \bar{Q}} \bar{c}_q \lambda_q \tag{22}$$

$$s.t. \quad \sum_{q \in \bar{Q}} \bar{A}_q \lambda_q \geq b, \tag{23}$$

$$\sum_{q \in \bar{Q}} \lambda_q = 1, \tag{24}$$

$$\lambda \geq 0. \tag{25}$$

The resulting pricing subproblem:

$$\min_{q \in \bar{Q}} \{\tilde{c}_q := \bar{c}_q - \pi \bar{A}_q - \pi_0\} \tag{26}$$

is now solvable (due to the lower size of $\bar{Q}$) and column generation can be efficiently applied.

The two-stage column generation approach can be briefly outlined as follows:

Algorithm 1: Two-stage column generation

input set $\bar{X}$
repeat
  repeat
    CG1: generate extensive variables λ for partial master problem (PMP)
  until *optimal partial master problem (PMP)* ;
  CG2: generate compact variables x for partial compact formulation (PCF)
until *optimal master problem (MP)* ;

On the one hand, in (CG1) standard column generation applies; in particular, the dual optimal vector π is known at every iteration and thus reduced costs $\tilde{c}_q := \bar{c}_q - \pi\bar{A}_q - \pi_0$ of λ variables can be directly estimated.

On the other hand, in (CG2) we need to know the reduced costs of variables $x_i \in \hat{X}$ in order identify the profitable ones to be added to the partial compact formulation, if any.

The main issue of this approach, indeed, is represented by the computation of the reduced cost of the compact formulation variables, given an optimal solution of the linear relaxation of the extensive formulation. Some attempts in this direction have been proposed with another aim, namely the variable elimination based on reduced cost: for linear integer programs, non-negative variables with a reduced cost greater than the duality gap cannot be positive in any optimal solution, i.e. can be fixed to 0 or eliminated.

Walker (1969) illustrates a method which can be applied if the pricing problem can be solved as a pure linear program. As observed by Irnich et al. (2007), in the case of shortest path sub-problems, this results in SPP on acyclic networks. Poggi de Aragão and Uchoa (2003) propose to keep the so-called coupling constraints in the master problem formulation; however, Irnich et al. (2007) observe that there exists a feasible solution for such a master problem where the reduced costs associated to the coupling constraints are all zero, and the authors raise theoretical and algorithmic reasons for not using such approach. Irnich et al. (2007) propose to estimate the reduced cost of a variable by the smallest reduced cost of a column in which the variable is taken with positive value; however, this method cannot be directly applied to the two-stage framework since a correct estimation would imply the minimization over the entire set of original variables. Finally, Salani and Vacca (2008) suggest in their paper a general methodology to compute the reduced cost of compact-formulation variables, which is mainly based on complementary slackness conditions. Their approach is illustrated in the following subsection.

2.3 Reduced costs of compact variables by complementary slackness

The dual problem of the linear relaxation of the original compact formulation (1)–(4) is:

$$z_{LR} = \max \quad b^T\alpha + d^T\beta \tag{27}$$

$$s.t. \quad \alpha A + \beta D \leq c, \tag{28}$$

$$\alpha,\ \beta \geq 0. \tag{29}$$

Reduced costs of compact variables x are defined as:

$$\tilde{c} = c - \alpha A - \beta D \tag{30}$$

where α is the dual vector associated to constraints $Ax \geq b$ in (2) and β is the dual vector associated to constraints $Dx \geq d$ in (3).

At each completed cycle of (CG1), the partial master problem is optimally solved and the optimal objective function $\bar{z}^*_{MP}$ as well as the optimal values of primal vector λ^* and dual vector π^* are known.

Since π is the dual vector associated to covering constraints (13) in the master problem, it corresponds to dual vector α associated to the same covering constraints (2) in the compact formulation. Similarly for the partial master problem and the partial compact formulation.

Nevertheless, in order to estimate reduced costs (30), we still need to know the dual vector β associated to constraints $Dx \geq d$ of the subproblem in (3).

Salani and Vacca (2008) propose a method to compute reduced costs of compact variables $x \in \hat{X}$ starting from the optimal solution λ^* of a partial master problem and using complementary slackness arguments.

Firstly, given λ^*, it is possible to uniquely reconstruct the equivalent optimal solution $\bar{x}^*$ of the linear relaxation of the partial compact problem. This is trivial when the compact formulation is known and the extensive formulation is obtained through Dantzig-Wolfe reformulation, but this is also possible when only an extensive formulation is given (Villeneuve et al., 2005).

Furthermore, the solution $x = [\,\bar{x}^* \,|\, 0\,]$ is feasible for the linear relaxation of original compact problem (1)–(4), as $\bar{X} \subset X$. Additionally, this solution is trivially optimal for the following constrained problem:

$$
\begin{aligned}
\min \quad & c^T x && (31)\\
s.t. \quad & Ax \geq b, && (32)\\
& Dx \geq d, && (33)\\
& \bar{x} = \bar{x}^*, && (34)\\
& \hat{x} = 0, && (35)\\
& x \geq 0. && (36)
\end{aligned}
$$

The dual of the constrained problem is given by:

$$
\begin{aligned}
\max \quad & b^T\alpha + d^T\beta + \bar{x}^*\gamma && (37)\\
s.t. \quad & \alpha A + \beta D + \gamma I + \delta I \leq c, && (38)\\
& \alpha,\ \beta \geq 0, && (39)\\
& \gamma \in \mathbb{R}^{\bar{n}}, && (40)\\
& \delta \in \mathbb{R}^{\hat{n}}. && (41)
\end{aligned}
$$

Given the optimal solution $x^* = [\,\bar{x}^* \,|\, 0\,]$ of the primal constrained problem, the associated dual optimal solution $[\,\alpha^*,\ \beta^*, \gamma^*, \delta^*\,]$ can be computed using complementary slackness conditions:

$$
\begin{aligned}
\alpha\,(Ax^* - b) &= 0 && (42)\\
\beta\,(Dx^* - d) &= 0 && (43)\\
\gamma\,(\bar{x}^* - \bar{x}^*) &= 0 && (44)\\
\delta\,(\hat{x}^* - 0) &= 0 && (45)
\end{aligned}
$$

in addition to dual feasibility:

$$\alpha A + \beta D + \gamma I + \delta I \leq c \quad (46)$$

$$\alpha,\ \beta \geq 0 \quad (47)$$

Conditions (44) and (45) just state that variables γ and δ are free, since (34) and (35) are equality constraints. Furthermore, values for vector α are fixed, since, as remarked, it corresponds to dual vector π obtained solving the (partial) master problem, i.e. $\alpha^* = \pi^*$. Therefore, conditions (42)–(47) reduce to:

$$\beta\,(Dx^* - d) = 0 \quad (48)$$

$$\pi^* A + \beta D + \gamma I + \delta I \leq c \quad (49)$$

$$\beta \geq 0 \quad (50)$$

The dual vector δ represents the reduced costs of not-yet-added variables $\hat{x}$ that we want to determine.

$$[\gamma,\ \delta]^T = \tilde{c} = c - \alpha A - \beta D$$

The system of linear equations (48)–(50) can be transformed in a linear program using a trivial objective function and solved to optimality, e.g. with the simplex method. To illustrate the framework and for the sake of simplicity the following objective is proposed:

$$\max \mathbf{1^T}\delta,$$

as the analysis and the comparison of different objective functions goes beyond the scope of the work.

Authors finally remark that, if the solution of (48)–(50) is such that $\gamma^* \geq 0$ and $\delta^* \geq 0$, then $[\,\alpha^*,\ \beta^*\,]$ is also a feasible solution of the dual (27)–(29) of the original compact formulation and therefore $(b^T\alpha^* + d^T\beta^*)$ is a valid lower bound to (1)–(4). Furthermore, since all reduced costs are positive, (CG2) stops.

3 Application of two-stage column generation to TBAP

In this section we provide a Dantzig-Wolfe reformulation of TBAP and we illustrate how the two-stage column generation framework can be applied to this problem.

We refer to the compact TBAP formulation introduced by Giallombardo et al. (2008). Given $n = |N|$ ships with time windows on the arrival time at the terminal, $m = |M|$ berths with time windows on availability, a planning time horizon discretized in $|H|$ time steps, a set P_i of feasible QC assignment profiles defined for every ship $i \in N$, and the maximum number of quay cranes available in the terminal, the aim is to find a feasible assignment of ships to berths, a feasible scheduling os ship in every berth and to assign a quay crane profile to every ship, in order to maximize the total value of selected profiles as well as minimize the transshipment-related quadratic yard costs. Computational results show that the problem is difficult to solve as-it-is already on small instances. We have therefore reformulated the problem via Dantzig-Wolfe decomposition. For easiness of illustration, in the following we consider a simplified objective function which only maximizes the total value of selected profiles, ignoring at first the quadratic term; this simplified problem will represent our original compact formulation. As shown in Giallombardo et al. (2008) the quadratic term can be linearized and included in the formulation to match the proposed framework.

We firstly introduce the concept of *berth sequence*, which represents a sequentially ordered (sub)set of ships in a berth with an assigned QC profile. Let Ω^k be the set of all feasible sequences r for berth $k \in M$, including the empty sequence, which means that berth k is not used. Let z_r^k be the decision variable which is 1 if sequence $r \in \Omega^k$ is used by berth k and 0 otherwise. The extensive TBAP formulation is the following:

$$\max \quad \sum_{k \in M} \sum_{r \in \Omega^k} v_r^k z_r^k \tag{51}$$

$$\text{s.t.} \quad \sum_{k \in M} \sum_{r \in \Omega^k} y_{ir}^k z_r^k = 1 \qquad \forall i \in N, \tag{52}$$

$$\sum_{k \in M} \sum_{r \in \Omega^k} q_r^{hk} z_r^k \leq Q^h \qquad \forall h \in H, \tag{53}$$

$$\sum_{r \in \Omega^k} z_r^k = 1 \qquad \forall k \in M, \tag{54}$$

$$z_r^k \in \{0, 1\} \qquad \forall r \in \Omega^k, \forall k \in M. \tag{55}$$

where: v_r^k is the value of sequence $r \in \Omega^k$; y_{ir}^k is 1 if ship i in sequence $r \in \Omega^k$ and 0 otherwise; q_r^{hk} is the number of QCs used by sequence $r \in \Omega^k$ at time step h; Q^h is the number of QCs available at time step h.

The value v_r^k of a sequence $r \in \Omega^k$ is given by the sum of the values of the profiles assigned to ships served by the sequence:

$$v_r^k = \sum_{i \in N} \sum_{p \in P_i} v_i^p \lambda_{ir}^{pk} \tag{56}$$

where v_i^p is the value of profile $p \in P_i$ and λ_{ir}^{pk} is a parameter which is 1 if profile $p \in P_i$ is in sequence $r \in \Omega^k$ and 0 otherwise.

Equations (52) are ship-cover constraints, (53) are QCs capacity constraints and (54) are convexity constraints. The objective function (51) maximizes the total value of chosen sequences.

When identical requirements hold on the subsets i.e. $\Omega^k = \Omega \ \forall k$, we can define the decision variable z_r, which is equal to 1 if sequence $r \in \Omega$ is chosen and 0 otherwise, and the extensive formulation becomes:

$$\max \quad \sum_{r \in \Omega} v_r z_r \tag{57}$$

$$\text{s.t.} \quad \sum_{r \in \Omega} y_{ir} z_r = 1 \qquad \forall i \in N, \tag{58}$$

$$\sum_{r \in \Omega} q_r^h z_r \leq Q^h \qquad \forall h \in H, \tag{59}$$

$$\sum_{r \in \Omega} z_r = m, \tag{60}$$

$$z_r \in \{0, 1\} \qquad \forall r \in \Omega. \tag{61}$$

The linear relaxation of (57)-(61), called Master Problem (MP), has a huge number of columns (variables), as it is defined on the space of all feasible sequences Ω. We define the Restricted Master Problem (RMP) on a subset $\Omega' \subset \Omega$ of columns and we solve it by column generation.

Let $[\pi, \mu, \pi_0]$ be an optimal dual solution to a RMP, where $\pi \in \mathbb{R}^n$ is the dual vector associated to ship-cover constraints (58), $\mu \geq 0$ is the dual vector associated to capacity constraints (59) and $\pi_0 \in \mathbb{R}$ is the dual variable associated to the aggregated convexity constraint (60).

The reduced cost of a sequence r is:

$$\tilde{v}_r = v_r - \sum_{i \in N} \pi_i y_{ir} - \sum_{h \in H} \mu_h q_r^h - \pi_0 \tag{62}$$

where π_i represents the dual price of serving ship i in sequence r and μ_h represents the dual price of using an additional quay crane at time step h.

The *pricing problem* (or *subproblem*):

$$\max_{r \in \Omega \setminus \Omega'} \{\tilde{v}_r\} = \max_{r \in \Omega \setminus \Omega'} \{v_r - \sum_{i \in N} \pi_i y_{ir} - \sum_{h \in H} \mu_h q_r^h\} - \pi_0 \tag{63}$$

identifies a column r^* with maximum reduced cost. If $\tilde{v}_{r^*} > 0$, we have identified a new column to enter the basis; if $\tilde{v}_{r^*} \leq 0$, we have proven that the current solution of RMP is also optimal for MP.

In the pricing problem, several decisions have to be made:

(i) whether ship i is in the sequence or not; this decision is represented by cost component y_{ir};

(ii) whether profile p is used by ship i or not; this decision, represented by λ_{ir}^p, is implicitly included in the pricing problem through cost component $v_r = \sum_{i \in N} \sum_{p \in P_i} v_i^p \lambda_{ir}^p$;

(iii) the order of ships in the sequence; this decision is implicitly represented by cost component q_r^h.

By defining a network $G(\tilde{N}, A)$, which has one node for every ship $i \in N$, for every profile $p \in P_i$ and for every time step $h \in H$, and whose arcs have transit time equals to the length of the profile, we can reduce the pricing problem to an Elementary Shortest Path Problem with Resource Constraints (ESPPRC). The size of this network grows polynomially with the number of ships, working shifts and QC profiles; however, in the worst case (unbounded time windows) the network is complete and has $\sim |\tilde{N}|^2$ arcs. Consequently, since ESPPRC is a NP-Hard combinatorial problem, its solution on such a big network is impractical.

At this point, two-stage column generation is suitable to by applied. We consider a partial compact formulation defined over a subset $P_i' \subset P_i$ of quay crane profiles for every ship $i \in N$ and we reformulate the problem using Dantzig-Wolfe. We assume w.l.o.g that $P_i = [\,\bar{P}_i \,|\, \hat{P}_i\,]$ and that $\bar{\Omega}$ is the set of lower dimension extreme points. The resulting partial master problem is the following:

$$\max \quad \sum_{r \in \bar{\Omega}} v_r z_r \tag{64}$$

$$\text{s.t.} \quad \sum_{r \in \bar{\Omega}} y_{ir} z_r = 1 \qquad \forall i \in N, \tag{65}$$

$$\sum_{r \in \bar{\Omega}} q_r^h z_r \leq Q^h \qquad \forall h \in H, \tag{66}$$

$$\sum_{r \in \bar{\Omega}} z_r = m, \tag{67}$$

$$z_r \in \{0, 1\} \qquad \forall r \in \bar{\Omega}. \tag{68}$$

We repeatedly solve the partial master problem (CG1), adding at each iteration profitable compact formulation variables in $\hat{P}_i$ set (CG2). Specifically, among the quay cranes profiles not yet considered, we select the subset of profiles with strictly positive reduced cost and we iterate the entire process. In this sense the column generation algorithm has two stages: firstly, berth sequences are built considering a subset of QC profiles and, subsequently, promising QC profiles are added to the model.

4 Conclusions

In this paper, we consider a Dantzig-Wolfe reformulation of TBAP, a problem which integrates two hard combinatorial problems arising in the management of container terminals. We start from a mixed integer compact formulation presented by Giallombardo et al. (2008) and obtain an extensive formulation based on the concept of *berth sequence*, which represents a feasible allocation of ships to a berth and a combined assignment of QCs to ships. We remark that the proposed standard reformulation leads to a pricing problem which is hard to solve, namely an Elementary Shortest Path Problem with Resource Constraints (ESPPRC) on a network with a potentially combinatorial number of nodes. After a review of the two-stage column generation method, recently proposed by Salani and Vacca (2008), we discuss its application to the TBAP showing the benefit of the approach in reducing the complexity of the pricing problem.

References

Barnhart, C., Johnson, E. L., Nemhauser, G. L., Savelsbergh, M. and Vance, P. H. (1998). Branch-and-price: Column generation for solving huge integer programs, *Operations Research* **46**(3): 316–329.

Dantzig, G. B. and Wolfe, P. (1960). Decomposition principle for linear programs, *Operations Research* **8**: 101–111.

Desaulniers, G., Desrosiers, J. and Solomon, M. (eds) (2005). *Column Generation*, GERAD 25th Anniversary Series, Springer.

Giallombardo, G., Moccia, L., Salani, M. and Vacca, I. (2008). The tactical berth allocation problem with quay crane assignment and transshipment-related quadratic yard costs, *Proceedings of the European Transport Conference (ETC)*.

Imai, A., Chen, H. C., Nishimura, E. and Papadimitriou, S. (2008). The simultaneous berth and quay crane allocation problem, *Transportation Research Part E: Logistics and Transportation Review* **44**: 900–920.

Irnich, S., Desaulniers, G., Desrosiers, J. and Hadjar, A. (2007). Path reduced costs for eliminating arcs, *Technical Report G-2007-79*, Les cahiers du GERAD, HEC Montréal.

Meisel, F. and Bierwirth, C. (2006). Integration of berth allocation and crane assignment to improve the resource utilization at a seaport container terminal, *Operations Research Proceedings 2005*, Springer, pp. 105–110.

Meisel, F. and Bierwirth, C. (n.d.). Heuristics for the integration of crane productivity in the berth allocation problem, *Transportation Research Part E: Logistics and Transportation Review* **In Press, Corrected Proof**: doi:10.1016/j.tre.2008.03.001.

Nakao, Y. and Nagamochi, H. (2007). A dp-based heuristic algorithm for the discrete split delivery vehicle routing problem, *Journal of Advanced Mechanical Design, Systems, and Manufacturing* **1**(2): 217–226.

Nemhauser, G. L. and Wolsey, L. A. (1988). *Integer and Combinatorial Optimization*, John Wiley & Sons, New York, N. Y.

Park, Y. M. and Kim, K. H. (2003). A scheduling method for berth and quay cranes, *OR Spectrum* **25**: 1–23.

Poggi de Aragão, M. and Uchoa, E. (2003). Integer program reformulation for robust branch-and-price algorithms, *Proceedings of Mathematical Programming in Rio: A conference in honour of Nelson Maculan*, pp. 56–61.

Salani, M. and Vacca, I. (2008). Two-stage column generation and applications, *Technical Report TRANSP-OR 081002*, Transport and Mobility Laboratory, EPFL.

Stahlbock, R. and Voss, S. (2008). Operations research at container terminals: a literature update, *OR Spectrum* **30**: 1–52.

Steenken, D., Voss, S. and Stahlbock, R. (2004). Container terminal operation and operations research - a classification and literature review, *OR Spectrum* **26**: 3–49.

UNCTAD (2007). Review of maritime transport, *Technical report*, United Nations, New York and Geneva.

Vacca, I., Bierlaire, M. and Salani, M. (2008). Optimization at container terminals: Status, trends and perspectives (revised version), *Technical Report TRANSP-OR 080528*, Transport and Mobility Laboratory, EPFL.

Villeneuve, D., Desrosiers, J., Lübbecke, M. E. and Soumis, F. (2005). On compact formulations for integer programs solved by column generation, *Annals of Operations Research* **139**(1): 375–388.

Vis, I. F. A. and de Koster, R. (2003). Transshipment of containers at a container terminal: An overview, *European Journal of Operational Research* **147**: 1–16.

Walker, W. E. (1969). A method for obtaining the optimal dual solution to a linear program using the Dantzig-Wolfe decomposition, *Operations Research* **17**: 368–370.

Xu, J. and Chiu, S. Y. (2001). Effective heuristic procedures for a field technician scheduling problem, *Journal of Heuristics* **7**(5): 495–509.