2.10 A simulation based approach to optimize berth throughput under uncertainty at marine container terminals

A simulation based approach to optimize berth throughput under uncertainty at marine container terminals

Mihalis M. Golias
University of Memphis
mihalisgolias@yahoo.com

Abstract. Berth scheduling is a critical function at marine container terminals and determining the best berth schedule depends on several factors including the type and function of the port, size of the port, location, nearby competition, and type of contractual agreement between the terminal and the carriers. In this paper we formulate the berth scheduling problem as a bi-objective mixed-integer problem with the objective to maximize customer satisfaction and reliability of the berth schedule under the assumption that vessel handling times are stochastic parameters following a discrete and known probability distribution. A combination of an exact algorithm, a Genetic Algorithms based heuristic and a simulation post-Pareto analysis is proposed as the solution approach to the resulting problem. Based on a number of experiments it is concluded that the proposed berth scheduling policy outperforms the berth scheduling policy where reliability is not considered.

1.0 INTRODUCTION

In this paper, we deal with the discrete space and dynamic vessel arrival berth scheduling problem (DDBSP), which can be formulated as the machine scheduling problem [1, 2]. The DDBSP received and continues to receive increased attention from the research community as it is a problem that marine container terminal operators deal with on a daily basis [3]. In this paper we formulate the DDBSP as a bi-objective mixed-integer problem with the objective to maximize berth throughput and maximize the reliability of the berth schedule, under the assumption that the vessel handling times are stochastic parameters with a known discrete probability distributions. Berth throughput is taken under consideration by the minimization of the total service time for all the vessels. In order to maximize the reliability of the berth schedule, a risk measure is proposed that is dependent on the vessel-to-berth assignment and the distributions of the vessels' handling times.

The remainder of this paper is organized as follows: The next section describes the problem and motivation for considering stochasticity in the vessel handling times, and presents the model formulation. The third section presents the solution algorithm and the fourth section a small number of numerical examples to evaluate the proposed approach. The final section summarizes findings and suggest future research directions.

2.0 MODEL FORMULATION

As has been supported by the literature [4] the competitiveness of a container terminal depends on various factors. It has also been well established by researchers and practitioners that decisions in container terminals are interrelated [5]. More specifically, decisions regarding the BSP are closely related and affect(ed) by the decisions regarding the scheduling and productivity of the quay cranes and the internal transport movers [3, 5]. Combining these problems into one single problem cannot be handled efficiently and the number of assumptions adopted (when researchers tried to partially combine them) portrays an approximation of reality [5]. For that reason the majority of research, to our knowledge, has focused in isolating each problem and assumed deterministic inputs from its related counterpart problems (for example in the berth scheduling problem the assignment of the quay cranes on each vessel is assumed as an input). To that end, the majority of BSP models have not accounted for the stochastic nature of the vessel handling times; a stochasticity that

stems from the fact that quay cranes and internal transport vehicles serving the vessels do not have a deterministic productivity (e.g. random down time of quay cranes, unpredicted congestion in the yard, etc). The only exceptions have been three separate studies by Moorthy and Teo [6], Golias et al. [7], and Zhou and Kang [8]. Unlike the model presented herein though, Zhou and Kang [8] proposed a non-linear model formulation minimizing only the total waiting time, Golias et al. [7] focused on online conceptual formulations, and Moorthy and Teo [6] proposed an approach focusing in the randomness of the vessels arrival times and which is relevant only when a substantial number of vessels arrive periodically (as stated by the authors).

In this paper we propose a linear mixed integer bi-objective formulation where we account for the stochasticity in the vessel handling times and assume they are stochastic variables following different discrete probability distributions. The vessel handling time distributions at each berth can be obtained from historical data (i.e. berth assignment, number of QCs and ITVs, breakdown rates of QCs, utilization of yard, vessel handling volumes etc) using data mining algorithms but in this paper we assume that they are known for all the vessels at all the berths. Based on these distributions a risk measure for the berth schedule is proposed and minimized at the same time with the total service time for all the vessels. We choose to introduce the risk measure in contrast to formulating a stochastic optimization problem as the inherent combinatorial complexity of such a model would make it impossible to construct a meaningful heuristic that would efficiently search through the extremely large set of vessel handling time scenarios.

The two objectives introduced, when conflicting, will cause the improvement of one objective to degrade the performance of the other; thus the terminal operator needs to select a schedule that balances between the two objectives. Berth schedules with a

high berth throughput have a greater degree of risk (i.e. risk of matching the total service time when the stochastic vessel handling times are realized). On the other hand berth schedules with a lesser degree of risk (decreased berth throughputs), provide more confidence to the terminal operator that the resulting assignment will be stable in terms of the handling times for each vessel and thus deviations from the initial schedule will be minimized in case rescheduling is needed. The proposed model formulation and solution algorithm will provide the terminal operator with the berth schedule that optimally balances between the two objectives. In the following subsection we introduce the risk function and the full model formulation, followed by the solution algorithm in section 3.

2.1 Estimation of berth schedule risk

Let $M_{ij} = \{c_{ij}^1, c_{ij}^2, \dots c_{ij}^m\}$ be the set of the m possible handling times of vessel j at berth i. Also let $P(c_{ij}^a)$ be the probability that:

$$c_{ij} = c_{ij}^a \in M_{ij}, \sum_{c_{ij}^a \in M_{ij}} P(c_{ij}^a).$$ Then the

expected handling time at berth i for vessel j is equal to: $E(c_{ij}) = \sum_{c_{ij}^a \in M_{ij}} c_{ij}^a P(c_{ij}^a)$. In this paper we define as the measure of risk of a vessel j served at berth i as:

$$R_{ij} = \sum_{c_{ij}^a \in M_{ij}} \{\max(0, (c_{ij}^a - E(c_{ij}))P(c_{ij}^a)\}.$$ We

demonstrate this notion by a simple example with one vessel and two berths. Table 1 summarizes the data and results for this example. If the vessel is served at berth 1 then there is a 10% probability that the handling time will be 25 hours, an 80% probability that the handling time will be 31 hours and so on. On the other hand if the vessel is served at berth 2 then there is a 70% probability that the handling time will be 22 hours, a 10% that the handling time will be 25 hours and so on. Although serving the vessel at berth 2 has the lowest expected handling time, it also has the highest risk of exceeding the expected handling time.

Table 1. Example of berth schedule risk estimation for one vessel and two berths

m	Berth 1 ($i=1$ $j=1$)			Berth 2 ($i=2$ $j=1$)		
	c_{ij}^m	$P(c_{ij}^m)$	R_{ij}	c_{ij}^m	$P(c_{ij}^m)$	R_{ij}
1	25	10%	-	22	70%	-
2	31	80%	0.0	25	10%	-
3	35	5%	0.19	38	9%	1.07
4	43	5%	0.59	43	11%	1.86
	$E(c_{ij}) = 31$		0.78	$E(c_{ij}) = 26$		2.94

In order to formulate our problem we further define the following:

Nomenclature
Sets

I, J	set of berths and vessels

Decision Variables

$x_{ij} \in \{0,1\}, i \in I, j \in J$	1 if vessel j is served at berth i and zero otherwise
$y_{ab} \in \{0,1\}, a, b \in J$	1 if vessel b is served at the same berth as vessel a as its immediate successor and zero otherwise
$f_j \in \{0,1\}, j \in J$	1 if vessel j is the first vessel to be served at its assigned berth
$l_j \in \{0,1\}, j \in J$	1 if vessel j is the last vessel to be served at its assigned berth
$t_j \in R^+, j \in J$	start time of service for vessel j

Parameters

$c_{ij}^a \in M_{ij}, i \in J, j \in J$	handling time of vessel j at berth i with probability $P(c_{ij}^a)$
$A_j, j \in J$	arrival time of vessel j
$S_i, i \in I$	time berth i becomes available for the first time in the planning horizon

The bi-objective model formulation minimizing the vessel total service time and risk (from now on referred to as RSBM) is formulated as follows:

$$f_1 : \min \left[\sum_{j \in J} t_j + \sum_{i \in I} \sum_{j \in J} (E(c_{ij}) x_{ij}) \right] \quad (1)$$

$$f_2 : \min \sum_{i \in I} \sum_{j \in J} x_{ij} R_{ij} \quad (2)$$

Subject to:

$$\sum_{i \in I} x_{ij} = 1, \forall j \in J \quad (3)$$

$$f_b + \sum_{a \neq b \in J} y_{ab} = 1, \forall b \in J \quad (4)$$

$$l_a + \sum_{b \neq a \in J} y_{ab} = 1, \forall a \in J \quad (5)$$

$$f_a + f_b \leq 3 - x_{ia} - x_{ib}, \forall i \in I, a, b \in J, a \neq b \quad (6)$$

$$l_a + l_b \leq 3 - x_{ia} - x_{ib}, \forall i \in I, a, b \in J, a \neq b \quad (7)$$

$$y_{ab} - 1 \leq x_{ia} - x_{ib} \leq 1 - y_{ab}, \forall i \in I, a, b \in J, a \neq b \quad (8)$$

$$t_j \geq A_j, \forall j \in J \quad (9)$$

$$t_j \geq S_i x_{ij}, \forall i \in I, j \in J \quad (10)$$

$$t_b \geq t_a + \sum_{i \in I} E(c_{ia}) x_{ia} - M(1 - y_{ab}), \quad (11)$$

$$\forall a, b \in J, a \neq b$$

The first objective function (1) minimizes the expected total service time for all the vessels (from now on referred to as the expected cost or EC). The second objective function (2) minimizes the total risk of exceeding the expected value of handling time for the vessels. Constraint set (3) ensures that each vessel will be served once, while constraint set (4) ensures that each vessel will either be served first or be preceded by another vessel. In a similar manner constraint set (5) ensures that each vessel will either be last or it will be served before another vessel. Constraint sets (6) and (7) ensure that only one vessel can be served first and last at each berth. Constraint set (8) ensures that a vessel can be served after another vessel only if both are served at the same berth. Constraint sets (9) and (10) ensure that the vessel service start time will be greater than the vessel arrival or the time that the berth where the vessel is served becomes available for the first time in the planning horizon. Constraint set (11) estimates the start time of service for each vessel.

3.0 SOLUTION ALGORITHM
The RSBM is a bi-objective optimization problem and for both single objective

218

problems (derived once we eliminate one of the two objectives) no exact solution algorithm exist that can be applied and solve them in polynomial time. In order to tackle this issue a new heuristic approach is presented. The proposed heuristic is an improvement of the exact algorithm 2-PPM proposed by Lemerse et al., [9]. The concept of the 2-PPM algorithm was to split the search space into equal and predetermined partitions and then for each partition use the ε-constraint method to find a solution. The algorithm proposed herein follows the same concept of partitioning the search space, but does so in an adaptive manner without having to predefine the size or the number of partitions. Furthermore, instead of the ε-constraint method the weighted approach is used to produce a solution within each partition, resulting in a faster estimation of the PF (PF). Before we present the proposed heuristic we define the following:

Definition 1: Let X is the feasible space of the RSBM and $x \in X$ be a feasible solution. Solution $a \in X$ dominates solution $b \in X$ if:
$\{f_1(a) \le f_1(b), f_2(a) < f_2(b)\}$ or
$\{f_1(a) < f_1(b), f_2(a) \le f_2(b)\}$. Any dominated solutions do not belong in the PF (the set of PF solutions is denoted from now on as Ω).

Definition 3: Let $x_n^{pf} \in \Omega \subset X$ be the nth PF solution of the RSBM.

The proposed heuristic that can provide Ω (from now on called the Bi-objective Berth Scheduling Heuristic or BBSH), can be described using the pseudo-code in figure 1. As the single objective problems (solved at step 2 of the BBSH) are *NP*-Hard, a Genetic Algorithms (GAs) based heuristic, proposed for the DDBSP by Golias et al., [10] is employed as the solution algorithm. The GAs heuristic is briefly described in the following subsection for consistency purposes.

Step 1: Set $Z_{new} = \varnothing$

$\Omega = Z = (x_1^{pf} : \arg\min_{x \in X} f_1(x), x_2^{pf} : \arg\min_{x \in X} f_2(x))$

$\Pi_1 = f_1(x_1^{pf}), \Pi_2 = f_2(x_2^{pf})$

Step 2:
for $i=2: |Z|$

if $\left| \dfrac{f_1(x_i^{pf}) - f_1(x_{i-1}^{pf})}{f_1(x_i^{pf})} \right| > 0.01$ and

$\left| \dfrac{f_2(x_{i-1}^{pf}) - f_2(x_i^{pf})}{f_2(x_{i-1}^{pf})} \right| > 0.01$

$P : x_{|Z|+1}^{pf} = \arg\min_{x \in X} \left(\dfrac{f_1(x)}{\Pi_1} + \dfrac{f_2(x)}{\Pi_2} \right)$

$s.t.$

$f_1(x_{i-1}^{pf}) < f_1(x) < f_1(x_i^{pf})$

$f_2(x_i^{pf}) < f_2(x) < f_2(x_{i-1}^{pf})$

if P is infeasible: $Z = Z \setminus x_{|Z|-1}^{pf}$
else
$Z_{new} = Z_{new} \cup x_{|Z|+1}^{pf}, \Omega = \Omega \cup x_{|Z|+1}^{pf}$
end if

end if
end for
$Z = Z \cup Z_{new}, Z_{new} = \varnothing$
Order and renumber solutions in Z based on tuple $\{f_1(x), f_2(x)\}, x \in Z$
Step 3: If $|Z|>1$ go to step 2 else end

Figure 1. BBSH Pseudo-code

3.1 GAs heuristic

The GAs heuristic consists of four parts: a) the chromosomal representation, b) the chromosomal mutation, c) the fitness evaluation and d) the selection process. For scheduling problems integer chromosomal representation is more adequate, since the classical binary representation can obscure the nature of the search [11]. In this paper, we use an integer chromosomal representation, in order to exploit in full the characteristics of the problem. An illustration of the chromosome structure is given in figure 2 for a small instance of the problem with 6 vessels and 2 berths. As seen in figure 2 the chromosome has twelve cells. The first 6 cells represent the 6 possible service orders at berth1 and the last 6 cells the 6 possible service orders at berth 2. In

the assignment illustrated in figure 2, vessels 2, 4, and 5 are served at berth 1 as the first, second and third vessel respectively, while vessels 1, 3, and 6 are served at berth 2 as the first, second, and third vessel respectively.

Berth	1						2					
Vessel	2	4	5	0	0	0	1	3	6	0	0	0

Figure 2. Illustration of chromosome representation

Four different types of mutation are applied as part of the genetic operations to all the chromosomes at each generation: insert, swap, inversion, and scramble mutations. Since the RSBM is a minimization problem the smaller the values of each objective function are, the higher the fitness value will be. We define the fitness function of a chromosome as: $f_{pt}(x) = F_{max}^{pt} - z_{pt}(x)$, where F_{max}^{pt} is the maximum value of the objective function z_{it}, and f_{it} is the value of the fitness function of objective function i at iteration t of the GA. However, this value is not know in advance and is replaced by the largest value z_{pt} of each objective function at each iteration. Chromosomes whose objective function values do not satisfy the constraints of problem P are replaced by their parents.

3.2 Post Pareto simulation

The algorithm described in the previous subsections will produce a number of non-dominated solutions (i.e. berth schedules of the PF). The next step will be to select one of these solutions as the schedule to be implemented. This follow up step is known as post-Pareto analysis and can be quite a challenging task since, in the absence of subjective or judgmental information, none of the corresponding trade-offs can be said to be better than the others [13]. In the problem studied herein we employ simulation as means to select one schedule from the PF that will be implemented. The simulation entails the use of a simple Monte Carlo procedure that generates random instances of the vessels handling times and estimates an average of the total service time over all the instances. The procedure can be described as follows. Let K be the total number of different handling time instances (i.e. realizations of the vessels handling time) we wish to produce and $CPF_{ij}(\cdot)$ the cumulative distribution function of vessels' j handling time at berth i. This procedure is shown in figure 3.

```
for k=1:K, i=1:|I|, j=1:|J|
    Generate a number from the
    Uniform Distribution [0,1]: u=U(0,1)
    Set cs_{ij}^k = CPF_{ij}(u) , where: cs_{ij}^k : is
    the k^{th} realization of vessels' j
    handling time at berth i
end
```

Figure 3. Monte Carlo Procedure (MCP)

For each one of the solutions in the PF we estimate the mean value of the total service time (from now on referred to as the mean simulated cost or *MSC*) over all the K realizations of the vessel handling times (obtained from the MCP) as:

$$MSC(x_n^{pf}) = \frac{\sum_{k=1}^{K} f_1(cs_{ij}^k, x_n^{pf})}{K}$$. The

solutions with the minimum *MSC* over all the schedules in the PF (from now on referred to as the Pruned PF Solution or PPFS and denoted by x^{ppfs}) is selected as the schedule to be implemented.

4.0 COMPUTATIONAL EXAMPLES

Problems used in the experiments were generated randomly, but systematically. We developed twenty base problem instances, where vessels are served with various handling volumes at a multi-user container terminal (MUT) with five berths, with a planning horizon of one week, and various handling volumes. The range of variables and parameters considered herein were chosen according to [10, 13]. As previously discussed in subsection 2.1 of this paper we assume that vessel handling times are

220

stochastic parameters following a discrete probability distribution. Without loss of generality for the computational examples presented herein we assumed that each vessel at each berth will have a maximum of five possible handling times (i.e. $m=5$). We assumed that the handling time at each berth for each vessel increases randomly (from the minimum handling time) based on a uniform probability distribution with a minimum of 5% and a maximum of 15% (i.e.: if c_{ij}^1 is the minimum handling time of vessel j at berth i then the second through the fifth other possible handling times are estimated as: $c_{ij}^2 = c_{ij}^1 \times U(1.05, 1.15)$,

$c_{ij}^3 = c_{ij}^2 \times U(1.05, 1.15)$, $c_{ij}^4 = c_{ij}^3 \times U(1.05, 1.15)$,

$c_{ij}^5 = c_{ij}^4 \times U(1.05, 1.15)$). In total 5 datasets where created for each of the two different vessel inter-arrival times with different handling volumes.

For each dataset, and out of these five possible handling times at each berth, we randomly assigned the handling time that will receive the highest probability (from now on referred to as the dominant handling time or c_{ij}^d and dominant probability or $P(c_{ij}^d)$ respectively). For each dataset we developed four different cases where we allowed $P(c_{ij}^d)$ to vary based on the values shown in table 2. The probabilities for the four remaining handling times were estimated using the procedure shown in figure 4.

Table 2. Dominant handling time probabilities

Dominant Probability	50%-60%	60%-70%	70%-80%	80%-90%

In total 50 problem instances were created. Concerning the GAs based heuristic the population size was set to 50 chromosomes and the GA-based heuristic is stopped if no improvement is observed for 100 iterations (i.e. new or improved schedules found). Finally, the algorithm was implemented in Matlab and the experiments were performed

on an ASUS desktop personal computer (E5300@2.60GHz) with 6GB memory.

$$Step\_1 : P(c_{ij}^m) = U(P(c_{ij}^d), 1), \forall c_{ij}^m \neq c_{ij}^d \in M_{ij}$$

$$Step\_2 : P(c_{ij}^m) = \frac{1 - P(c_{ij}^d)}{\sum_{c_{ij}^m \neq c_{ij}^d \in M_{ij}} P(c_{ij}^m)} P(c_{ij}^m)$$

Figure 4. Handling Time Probabilities Estimation Procedure

4.1 Evaluation of berth scheduling policy

In this subsection we evaluate the payoff of introducing the second objective function. For each one of the 50 problem instances, previously described, we obtained the PF using the heuristic algorithm presented in section 3. For each schedule in the PF we calculated the MSC over a sample size of $K=10\,000$. As discussed in section 3.2 the schedule to be implemented will be the one with the minimum MSC. To evaluate the effectiveness of the proposed policy we compared the $MSC(x^{ppfs})$ value to the MSC value of the solution in the PF with the minimum EC. The latter is the solution: $x_1^{pf} : \underset{x \in X}{\arg\min} f_1(x)$ (i.e. the solution we would obtain if we did not consider the risk function as the problem would be single objective) and from now on will be referred to as the Nadir PF Schedule or NPFS and denoted by x^{npfs}. Table 3 reports the results from the comparison between $MSC(x^{ppfs})$ and $MSC(x^{npfs})$ values for all the 200 problem instances. The percentages reported in table 3 are estimated as: $\frac{MSC(x^{npfs}) - MSC(x^{ppfs})}{MSC(x^{npfs})}$, and answer the following question: "*On average should we expect a gain in berth throughput if we choose the PPFS over the NPFS and by how much?*". From the results in table 3 we observe that the PPFS always produces a smaller MSC. For example for the first dataset and the first dominant probability (i.e. 50%-60%) the MSC of the NPFS

solution is 12% and 3% larger than the *MSC* of the PPFS solution for the 3 and 5 hours vessel inter-arrival times respectively.

Table 3. *MSC* values difference (in %) between the NPFSs and the PPFSs (3 and 5 hours of vessel inter-arrival)

Dataset	50%-60%	60%-70%	70%-80%	80%-90%
1	12 / 3	19 / 27	9 / 13	12 / 31
2	15 / 31	10 / 49	30 / 57	4 / 34
3	10 / 44	2 / 41	34 / 18	9 / 12
4	44 / 32	7 / 33	12 / 20	8 / 9
5	24 / 14	37 / 0	42 / 29	7 / 63

5.0 CONCLUSIONS

In this paper, we formulated the discrete space and dynamic vessel arrival berth scheduling problem as a bi-objective mixed-integer problem with the objective to maximize the berth throughput and the reliability of the berth schedule, under the assumption that vessel handling times are stochastic parameters with known discrete probability distributions. In order to maximize the reliability of the berth schedule, a measure of risk was proposed dependent on the vessel-to-berth assignment. In order to solve the resulting problem, a combination of an exact and a GAs based heuristic were proposed and a number of simulation experiments were performed. Based on results from these experiments it was concluded that considering risk (from the inherent stochasticity of the vessel handling times) in berth scheduling can provide schedules with improved berth throughput when the vessel handling times are realized. Future research is focusing in: a) in a model formulation where the mutual impact of the vessels' stochastic handling times are considered, and b) evaluation of the proposed framework in terms of the robustness, dominance, and expected loss of the final schedule over all the Pareto points.

6.0 REFERENCES

[1] Imai, A., Nagaiwa, K., Tat, C-W., 1997. Efficient planning of berth allocation for container terminals in Asia. Journal of Advanced Transportation 31, 75–94.

[2] Imai, A., Nishimura, E., Papadimitriou S., 2001. The dynamic berth allocation problem for a container port. Transportation Research Part B, 35,401–417.

[3] Meisel, F., 2009. Seaside operations planning in container terminals. Physica-Verlag Berlin Heidelberg.

[4] Bierwirth, C., Meisel, F., 2009. A survey of berth allocation and quay crane scheduling problems in container terminals. European Journal of Operational Research, 202(3), 615-627.

[5] Ballis, A., Dimitriou, L., Paravantis, J., 2010. Quay to Storage Area Container Transfer: Critical Review of Modeling Techniques and Practical Outcomes. In: Proceedings of the 89th Annual Meeting of the Transportation Research Board, 10-3020.

[6] Moorthy, R., Teo C-P., 2006. Berth management in container terminal: The template design problem. OR Spectrum 28(4), 495-518.

[7] Golias, M.M., Boilé, M., Theofanis, S., 2007. The stochastic berth scheduling problem. In: Proceedings of the Second TRANSTEC Conference, Prague, Czech Republic.

[8] Zhou P-F, Kang, H-G., 2008. Study on berth and quay crane allocation under stochastic environments in container terminal. Systems Engineering Theory and Practice 28(1),161-169.

[9] Lemesre, J., Dhaenens, C., Talbi, E.G., 2007. Parallel partitioning method (PPM): A new exact method to solve bi-objective problems. Computers and Operations Research 34,2450–2462.

[10] Golias, M.M., Boilé, M., Theofanis, S., 2009. Service time based customer differentiation berth scheduling. Transportation Research Part E 45(6),878-892.

[11] Eiben, A.E., Smith, J.E., 2003. Introduction to Evolutionary Computing. Springer.

[12] Golias, M.M., Boilé, M., Theofanis, S., Taboada, A.H., 2010. A multi-objective decision and analysis approach for the berth scheduling problem. International Journal of Information Technology Project Management 1(1), 54-73.

[13] Hansen, P., Oguz, C., Mladenovic, N., 2008. Variable neighborhood search for minimum cost berth allocation. European Journal of Operational Research 131(3), 636-649.

7.0 ACKNOWLEDGMENT(S)