OR Spectrum (2008) 30:53–75
DOI 10.1007/s00291-007-0082-7

REGULAR ARTICLE

Optimizing the landside operation of a container terminal

Gary Froyland · Thorsten Koch · Nicole Megow · Emily Duane · Howard Wren

Published online: 4 August 2007

Abstract This paper concerns the problem of operating a landside container exchange area that is serviced by multiple semi-automated rail mounted gantry cranes (RMGs) that are moving on a single bi-directional traveling lane. Such a facility is being built by Patrick Corporation at the Port Botany terminal in Sydney. The gantry

Research supported by Patrick Technology and Systems, the Australian Research Council Centre of Excellence for Mathematics and Statistics of Complex Systems (MASCOS), and the German Research Foundation Research Center Mathematics for key technologies: Modelling, simulation, and optimization of real-world processes (MATHEON).

G. Froyland
School of Mathematics and Statistics, University of New South Wales,
Sydney NSW 2052, Australia
e-mail: g.froyland@unsw.edu.au

T. Koch (✉)
Zuse Institute Berlin, Takustr. 7, 14195 Berlin, Germany
e-mail: koch@zib.de

N. Megow
Institute of Mathematics, Technische Universität Berlin, Straße des 17. Juni 136,
10623 Berlin, Germany
e-mail: nmegow@math.tu-berlin.de

E. Duane
Department of Mathematics and Statistics, University of Melbourne,
Melbourne VIC 3010, Australia
e-mail: e.duane@ms.unimelb.edu.au

H. Wren
c/o Andrew Rembel, Patrick Technology and Systems, Level 2, 4b Lord St,
Botany NSW 2019, Australia
e-mail: a.rembel@patrick.com.au

cranes are a scarce resource and handle the bulk of container movements. Thus, they require a sophisticated analysis to achieve near optimal utilization. We present a three-stage algorithm to manage the container exchange facility, including the scheduling of cranes, the control of associated short-term container stacking, and the allocation of delivery locations for trucks and other container transporters. The key components of our approach are a time scale decomposition, whereby an integer program controls decisions across a long time horizon to produce a balanced plan that is fed to a series of short time scale online subproblems, and a highly efficient space-time divisioning of short-term storage areas. A computational evaluation shows that our heuristic can find effective solutions for the planning problem; on real-world data it yields a solution at most 8% above a lower bound on optimal RMG utilization.

Keywords Container terminal · Yard crane scheduling · Storage space allocation · Integer programming

Mathematics Subject Classification (2000) 90C90 · 90B06 · 90C06 · 90C10

1 Introduction

The past decade has witnessed a tremendous growth in marine transportation. Rising competition among ports and technical progress in ship design, resulting in higher capacity vessels, has put enormous pressure on port operators to develop efficient container handling systems.

One notable ongoing trend is the use of automated container handling and transportation technology Günther and Kim (2006). The full potential of such high-technology facilities can be utilized only with sophisticated optimization methods. Furthermore, research in this direction usually reveals the need for redesigning the layout of the terminal. Intelligent terminal layouts can increase the terminal capacity, reduce the time for container transport, and thus, reduce the turnaround time of ships enormously.

Patrick Corporation's container terminal at Port Botany in Sydney, Australia, was designed to handle 700,000 TEU (twenty-foot equivalent unit) per year, and by 2005 it was already processing 800,000 TEU. To further increase the capacity of the terminal and speed up transshipment processes, a change in operational design was necessary and led to significant changes in terminal layout as well as in systems design. The new design of the rail/road exchange area is intended to reduce the area requirements while at the same time to increase the exchange capacity. The principal change is the switch from a manned straddle carrier exchange for containers arriving by road to a semi-automated *Rail Mounted Gantry* (RMG) operation. The semi-automated RMGs will also handle containers arriving or departing by rail.

An overview of the overall port operation can be seen in Fig. 1. Export containers arrive by road or rail in the exchange area to be unloaded by RMG and to be transferred by straddle carriers to the main long-term storage yard within a few days. Later, straddle carriers will take the containers to the quayside where they will be loaded onto a ship by quay cranes. Import containers arrive by ship at the quayside and are transferred across the terminal to the rail/road exchange area in the opposite direction.

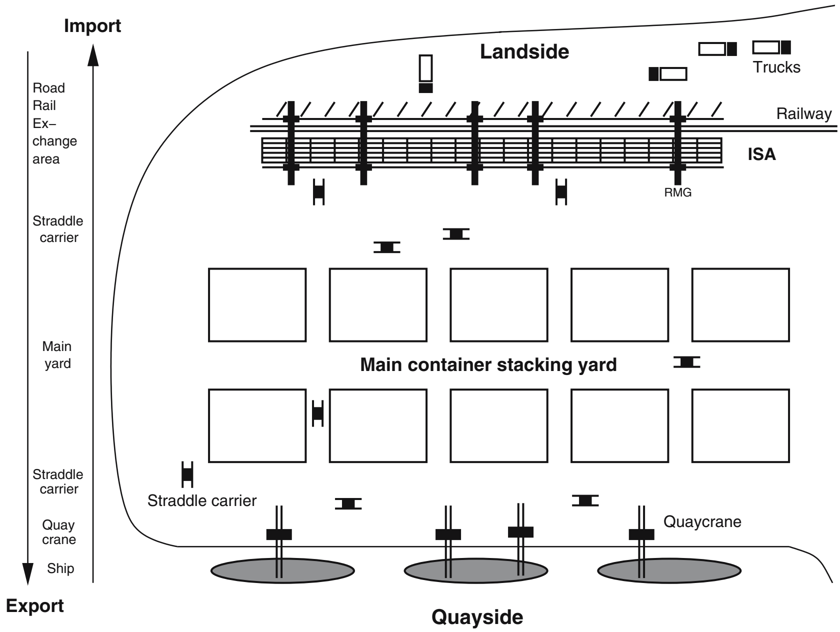


Fig. 1 Simplified Port Botany layout

The gantry operation is a new concept designed to achieve significant cost and operational efficiencies. The RMGs however are a limited resource and will handle almost every non-transshipment container passing through the terminal; thus sophisticated management systems must be established to ensure resource utilization is optimized.

The RMG operation includes an *Intermediate Stacking Area* (ISA) where containers will be placed once booked by road or rail carriers ready for prompt transfer by RMGs onto trucks or trains. Containers arriving for export by train or truck will also be stacked in this area until a suitable time for transfer to the main container stacking yard, initially operated with manned straddle carriers and eventually with automated straddle carriers (AutoStrads). The ISA allows the transfer of containers to be carefully planned to optimize both RMG and AutoStrad resources. Without the ISA handling, the highly unbalanced container flow leads to a similarly unbalanced use of resources in the terminal.

The management system for the RMGs and the ISA requires sophisticated analysis and allocation of container transfers to specific equipment. Our approach to building a near optimal resource utilization proceeds in three phases. The main idea is to decompose the complex problem into subproblems with small time planning intervals of 1 h length. The time interval of 1 h is motivated by the currently used hourly truck booking time slots.

In the first phase, a master problem is solved by an integer program which selects a 1 h time interval for the movement of each container from the quayside into the Intermediate Stacking Area, or vice-versa. The solution of the master problem determines all container transport into or out of the intermediate storage area for each time interval because transfers on the landside are fixed by arrival times of trucks or trains.

In the second phase, based on this allocation of container movements to single hours, the stacking positions for import containers on the quayside (served by straddle carriers) are determined by a set of integer programs.

Finally, in the third phase, an online heuristic efficiently solves the single-hour subproblems, which involves scheduling the RMGs, assigning the short-term positions of containers, and determining truck bays to be used. The online algorithm has to deal with the unknown information on exact truck arrival times. A static partitioning of the container stacking area further simplifies this third phase by decomposing the multiple-crane problem into single-crane problems.

This approach has proved very successful. For a particular simulation sourced from real world data we obtain a solution within 8% of the RMG operating time optimum. In addition, only 0.4% of the RMG operating time is spent on reshuffling containers. The solution is also optimal with respect to the maximum number of straddle carriers needed.

2 Related work

Due to the growing importance of marine transportation, operations on seaport container terminals have received increasing attention by researchers. Recent overviews that include detailed descriptions and classifications of major logistic operations on seaport container terminals are provided by Vis and de Koster (2003), Steenken et al. (2004), Kim (2005), and Günther and Kim (2005). The problems arising in the design and operation of inter-modal terminals are investigated among others in Kozan (2000), Alicke (2002), Ballas and Golias (2002), and Corry and Kozan (2006).

Most investigations in the literature are concerned with effectively allocating and scheduling key resources, such as berths, yards, quay cranes, yard cranes and container transporters. In fact, the focus is currently not on optimizing the transport chain as a whole but on optimizing several separate parts of the chain (Steenken et al. 2004). Two subproblems are related to our integrated approach: the storage area management and the yard crane (in our case RMGs) scheduling.

Storage area management addresses the assignment of storage locations to containers, which includes the allocation of space for containers moving into and out of the storage yard as well as reshuffles (rehandling). A limited number of scientific publications deal with the problem of stacking containers; see (Vis and de Koster 2003; Steenken et al. 2004; Kim 2005; Günther and Kim 2005). Most of the previous work does not determine the exact storage location within the yard; more common is to restrict the storage location to sub-blocks (sections of many locations), see, e.g., the recent work by Lee et al. (2006). In many other investigations it is assumed that the container positions are fully or partly predetermined. Recently, Dekker et al. (2006) considered various strategies for stacking containers in a simulation environment. Here, the focus is indeed on deciding where to place a new or reshuffled container. However, for most of their stacking strategies it is crucial that container categories are given, and containers of the same category can be exchanged.

Several approaches on scheduling different types of cranes have been published. The task is to find a schedule and route for one or more cranes. While most of the studies are restricted to single-crane scheduling, multiple cranes are rarely addressed. Zhang et al. (2002) and Cheung et al. (2002) consider multiple *rubber-tired gantry cranes* (RTGs). Given the forecasted workload distribution in yard blocks for different time periods, they find a schedule and routes for crane movements among blocks such that the total unfinished workload in the yard in each period is minimized. The problem is formulated as a mixed integer program and solved by different modifications of the Lagrangian relaxation method. While RTGs are very flexible in operation, *rail mounted gantry cranes* (RMGs) are more stable. Ng (2005) investigates the problem of scheduling multiple RMGs, going further than previous Double-RMG considerations. Based on a dynamic-programming approach, Ng decomposes the multi-crane problem into single-crane problems and solves these subproblems with a greedy heuristic.

There are also a few publications that consider several operational decisions in an integrated solution approach; see the reviews (Steenken et al. 2004; Vis and de Koster 2003). Typically, decision support systems are designed for particular terminal layouts utilizing meta-heuristics (see e.g., Kozan and Preston 2006 for recent work) and/or simulation techniques.

Even though automation of port operations is the current trend (Günther and Kim 2006), a semi-automated container exchange facility for loading and unloading trucks and trains using an intermediate buffer as designed by Patrick Corporation appears to be unique at this time.

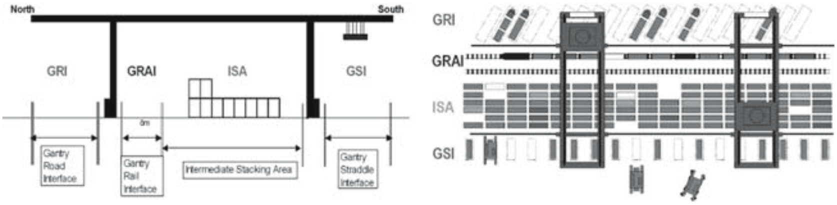
An additional complicating issue is the uncertainty of truck arrivals. Most of the published studies assume that truck arrival times or distributions are known a priori or use decision trees (Kim et al. 2000) as an attempt to support real world decisions. However, these approaches do not reflect the online character of the actual problem. General problems such as scheduling and routing arise frequently in optimization of logistics systems and have been considered in various settings (Bramel et al. 2005), including online environments (Jaillet and Wagner 2006; Murthy and Manimaran 2001; Pruhs et al. 2004). However, most of these approaches do not apply to seaport container terminals which have their own special characteristics.

3 Problem description

The container exchange between the terminal and road or rail is a very complex and time critical part of a port. Both container flows, inbound and outbound, are handled simultaneously. Thus, its effective operation determines terminal efficiency to a great extent.

Patrick Corporation designed a semi-automated container exchange facility that is currently under construction in Port Botany, Sydney. This facility comprises the Gantry-Road Interface (GRI), the Gantry-Rail Interface (GRAI), the Intermediate Stacking Area (ISA), and the Gantry-Straddle Interface (GSI) as shown in Fig. 2.

Up to five RMGs will handle all import and export containers utilizing the intermediate storage area for prompt transfer onto trucks or trains. The task is to find a schedule and routes for the RMG moves. Additionally, the stacking has to be managed



(a) Elevation **(b)** Ground plan

Fig. 2 Layout of the container exchange area comprising the Gantry-Road Interface (GRI), the Gantry-Rail Interface (GRAI), the Intermediate Stacking Area (ISA), and the Gantry-Straddle Interface (GSI)

including the allocation of all container positions in ISA, GSI and GRI. The RMGs are clearly a bottleneck resource and its operation needs careful planning to guarantee optimal resource utilization.

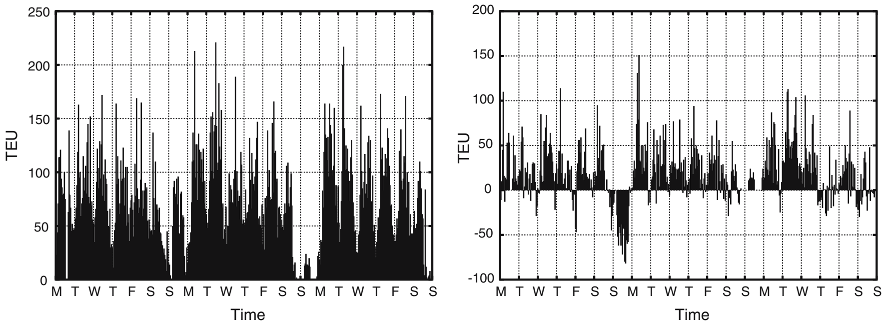
In the following sections we describe separately the three major components of the container exchange area complemented with some first analyses and statistics of relevant numbers in the data set. The data on container movements used in this article is based on historical work loads in Port Botany and was provided by Patrick Corporation. The data extends for 30 days, with total traffic slightly less than 50,000 TEU, just under 60% being imports. 20′ (twenty-foot) containers represent 57% of all containers by number, both for import and export. The remaining containers are 40′ containers. Container movements that were only partly in the observed time frame were removed. For the statistics shown, the first week and the last 4 days are removed to minimize temporal boundary effects.

3.1 The Gantry-Road Interface (GRI)

In the north is the *Gantry-Road Interface* (GRI) consisting of 60 truck slots. Located south of the truck slots is the Gantry Rail Interface (GRAI) consisting of two railway tracks. The GRAI is not considered in this investigation. All containers to be moved by train in the given data set were changed to truck arrivals/departures in the same hour. Further we assume that each truck either delivers or picks up only a single container.

The landside turnover per hour is highly varying. While the average in a week might be about 80 TEU per hour, the minimum is zero and the maximum well over 200 TEU in a single hour as depicted in Fig. 3a.

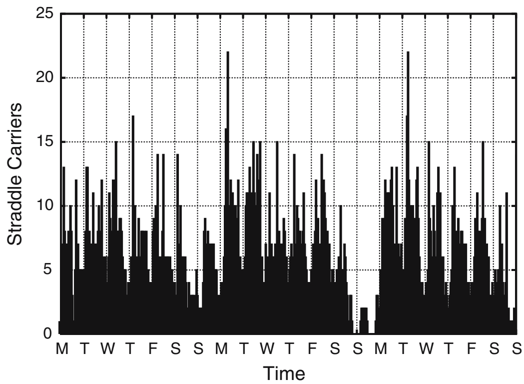
Trucks have to book their arrival hour (both for import and export) 24 h in advance. The exact time and order of truck arrivals within an hour are unknown beforehand. Patrick Corporation requires containers to be ready for truck pick up at least 4 h in advance. Thus for import containers there is an effective time window of 20 h in the ISA if we assume the container is immediately transferred from the main yard to the ISA once a truck arrival is booked. Direct imports from the GSI to GRI bypassing the ISA are prohibited for the following reason: One of the main goals of the new design is to minimize truck waiting times. Since neither the precise truck nor straddle arrival times are known, planning for a just in time delivery by the straddle carrier is impossible. Therefore, a direct transfer would require the container to wait for some



(a) Turnover **(b)** Import minus Export

Fig. 3 Landside container turnover per hour and dominating transfer direction

Fig. 4 Straddle requirements computed as a moving average of demand assuming both import and export containers reside in the GSI for only 1 h



time in the GSI to be ready when the truck shows up. This is not desirable, as it might block GSI positions for several hours, which would interfere with the GSI operation. In contrast, direct exports from GRI to GSI are allowed, because the above problems do not exist in this direction.

Nearly all container movements within the terminal are done by *straddle carriers*. Since straddles are an expensive resource both in terms of purchase and operation, it is important to minimize the maximum number of straddles needed at any particular hour; this also frees them up for other tasks. Straddles can perform approximately six trips between the main yard and the GSI per hour, moving at most 12 TEU in either direction. To reach this maximum, the number of containers arriving and departing has to be equal, which is usually not the case, as depicted in Fig. 3b. We call the consecutive movement of two containers in opposite directions by a straddle carrier a *pairing*.

Figure 4 demonstrates the imbalance in straddle carrier utilization that would occur without an ISA facility. It shows the number of straddles required per hour if all arrivals and departures have to be serviced "just in time" within the same hour when all possible pairings are performed. In Section 4.1.2, we show that the introduction of an intermediate buffer, such as the ISA, and strategic planning of this infrastructure can balance straddle requirements significantly.

3.2 The Gantry-Straddle Interface (GSI)

The *Gantry-Straddle Interface* (GSI) is located in the south. It consists of two rows of 132 slots for 20′ containers. Stacking up to three containers high is possible. Altogether, a maximum of 792 TEU can be placed in the GSI.

Export containers are required to be in the yard at least 12 h prior to ship arrival time. Since trucks might deliver the containers a week early, the usable time window for export container movements is typically much larger than the one for import containers. We will exploit this to reduce the number of straddle carriers needed by increasing the number of possible pairings.

3.3 The Intermediate Stacking Area (ISA)

Between the GRI and GSI lies the *Intermediate Stacking Area* (ISA) where import containers will be placed once booked by road or rail carriers such that they can be transferred by RMGs onto trucks or trains; see Fig. 2. Containers arriving for export by train or truck will also be stacked in this area until a suitable time for transfer to the main container stacking yard.

The ISA layout is seven rows by 100 columns of 20′ container spaces. At each space, it is possible to stack up to three containers high, resulting in a maximum capacity of 2,100 TEU. Since the flexibility of the ISA will be reduced substantially as the maximum ISA utilization limit is approached, it is unlikely that all 2,100 spaces can be utilized in practice.

The container movement through the ISA is done by five semi-automated rail mounted gantry cranes. In this study we consider RMGs that share the same tracks and cannot cross each other. Each RMG is capable of moving between 20 and 40 containers an hour. An RMG can *cross travel*, i.e., along ISA rows, and *long travel*, i.e., between GRI and GSI, at the same time with roughly the same speed. All RMGs can handle both 20′ and 40′ containers. While in principle possible, twin-lift operations of 20′ container were not considered, since we assume that each truck carries only a single container. Rotating a container can be done while moving it. The truck loading and unloading operations are semi-automated using tele-operators. Loading and unloading operations in both the ISA and GSI are fully automated.

Some of the ISA slots have facilities to host refrigerated (reefer) containers. For this study we did not make any special consideration for reefer or other containers with special requirements, e.g., dangerous goods, other than guaranteeing that the number of reefer containers in the ISA never exceeds the number of slots with power connectors.

3.4 Goals for managing the container exchange area

Patrick Corporation has identified multiple objectives for optimizing the operation of the container exchange area; these are in decreasing importance:

- move all containers while respecting the corresponding time windows, the ISA capacity limit, and causing no truck to wait more than 15 min;
- minimize the maximum number of straddle carriers needed. This includes a maximization of pairing opportunities for straddle operations;
- balance the work load of RMGs within each hour;
- balance the total RMG load over all hourly time slots;
- minimize the total usage of RMGs.

Notice that there is a conflict between the goals of minimizing the number of needed straddle carriers, balancing the RMG utilization, and having a low ISA utilization. Improving one of these goals has immediate repercussions for the other two.

3.5 Issues not considered and the consequences

While this study has incorporated a great deal of detail, there are still several aspects encountered in practice that are missing in our model. None of them is important enough to change the results substantially. However, some issues, such as incorporating trains, for example, may lead to different solutions.

- In some hours the landside turnover is simply too high for the RMGs to serve. In the current model, if required, we extend an “hour” until all trucks are served. In practice there is the opportunity of handling overload by directly transferring containers between trucks and straddle carriers.
- Trucks delivering and/or picking up more than one container at the same time or RMGs using twin-lifts are not considered. Including such considerations is likely to improve the performance of the facility. Regarding the integer program in Sect. 4.1 no changes are necessary. Since it is known beforehand where a twin-lift might be possible, it can be modeled the same way as a 40′ container.
- Precise truck routing is currently not considered. In fact, the time a truck needs to get to its assigned GRI slot is neglected. In practice, slots shall be assigned such that the truck throughput is maximized.
- A special treatment of certain container types is not considered. Special containers are those that require refrigeration, must be stored for a certain period of time, or will be x-rayed. Their special handling might lead to a decomposition of the ISA. The integer program in Sect. 4.1 is easily capable of ensuring capacity limits, but due to limited choice of RMGs the load balancing at the operational stage might suffer. These special containers might be handled best by direct exports bypassing the ISA.
- The simulation does not handle 40′ containers regarding stacking. Incorporating 40′ container stacking is likely to affect the performance due to limited choices for stacking.
- Our results are achieved assuming that the data is known for the entire month. Incorporating booking times for exports requires solving the integer program in Sect. 4.1 with a moving time horizon as new booking information becomes available. The strategic IP will get much smaller and thus easier to solve, but the solution quality might suffer.

- Time windows for the GSI delivery and non-uniform straddle availability could be incorporated to improve the flexibility of the solution method in adapting to other yard operations that require straddle carriers.
- An important issue that is simplified in the current simulation is the handling of trains. Due to the completely different arrival process, the introduction of trains will affect the optimal solution.

4 Solution method

We divide the decision making into three sequential subproblems.

1. **Strategic planning.** All container moves are temporally assigned at an hourly level. This is done by an integer program.
2. **Tactical setup.** Import containers are assigned to positions in the GSI. A set of integer programs solves this assignment problem.
3. **Operational decision-making.** Containers are assigned to short-term stacking positions (export containers to GSI positions, trucks to GRI positions) and precise RMG operations are planned within each hour. An online algorithm takes decisions concerning the RMG moves and placements depending on the online information about truck arrivals.

The decomposition of the complex planning problem into 1-h subproblems motivates a partitioning of the ISA into "corridors". These corridors define areas in which RMGs move. Moreover, we assign to each corridor a fixed time period in which a crane "mainly" moves in this corridor.

Firstly, we restrict the movements of each of the five RMGs to one-fifth of the ISA. We call these five sets of 20 columns *moving areas*. One of these moving areas is shown in Fig. 5. Secondly, for each hour we try to restrict each RMG to a *moving corridor* of five columns within its moving area. The four corridors cycle, so every 4 h the same corridor is used; see Fig. 5. The advantage of the corridor construction is the following: since the RMGs can long and cross travel at the same time it is possible to move about five columns across at no time cost while moving from GRI to GSI. The corridor construction aims to absorb most cross travel time into the unavoidable long travel time. This concept clearly reduces the solution space, but also allows us

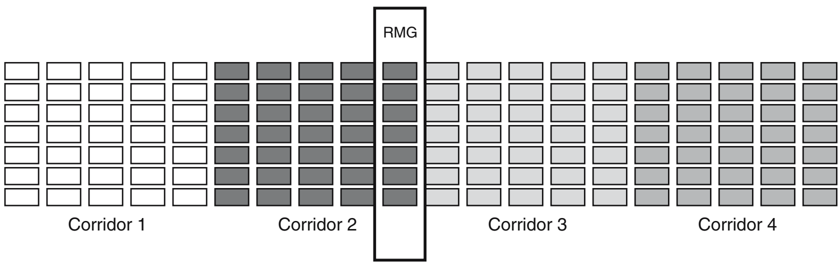


Fig. 5 RMG moving area with four moving corridors

to solve this complex planning problem in reasonable time and still yield a solution which we prove to be within 8% of the optimum performance.

4.1 Stage 1—A strategic IP to determine ISA-GSI transit times

In the first step, a *strategic integer program* (strategic IP) fixes the movement times of the containers between the ISA and GSI. The overall goal is to find a rough schedule with hourly precision. That reduces the complexity of the online problem to be solved in each single hour and enables a simple algorithm to succeed. The objectives of the strategic IP are:

- assign each container a GSI movement time within its allowed time window. The GSI movement and the GRI movement times should correspond to the same moving corridor;
- the maximum number of straddle carriers needed for an hour should be minimized assuming optimal pairing;
- the load of the RMGs should be balanced over all hours and stay within reasonable limits;
- the utilization of the ISA has to stay within prescribed limits.

4.1.1 Model description—Stage 1

Data

H	set of hours $\{1, \ldots, \mathcal{H}\}$
C	set of containers transiting GRI
C_h	$\subset C$, subset of containers transiting GRI in hour $h \in H$
I	$\subseteq C$, import containers
E	$\subseteq C$, export containers, $I \cup E = C$, $I \cap E = \emptyset$
I_h	$:= C_h \cap I$, import containers transiting GRI in hour $h \in H$
E_h	$:= C_h \cap E$, export containers transiting GRI in hour $h \in H$
W_c	$\subseteq H$, set of available hours for GSI movement by RMG for container $c \in C$ (time window in GSI)
L	set of RMG utilization levels $\{1, \ldots, \mathcal{L}\}$
M	set of ISA utilization levels $\{1, \ldots, \mathcal{M}\}$
λ_l	$\in \mathbb{N}_+$ Number of RMG operations below utilization level $l \in L$, $\lambda_l < \lambda_{l+1}$
μ_m	$\in \mathbb{N}_+$ Number of containers below ISA utilization level $m \in M$, $\mu_m < \mu_{m+1}$
τ_c	TEU count of container $c \in C$ (either 1 or 2)
$\tau(X)$	$\sum_{c \in X} \tau_c$, $X \subset C$.

Variables

We define variables $x_{ch} \in \{0, 1\}$ with $c \in C$ and $h \in W_c$.

$$x_{ch} = \begin{cases} 1, & \text{if container } c \text{ is moved between GSI and ISA in hour } h; \\ 0, & \text{otherwise.} \end{cases}$$

Furthermore, we define the following non-negative integer variables:

t_h^i	number of RMG movements from GSI in hour $h \in H$
t_h^e	number of RMG movements to GSI in hour $h \in H$
t	$\geq \max_{h \in H}\{t_h^i,\ t_h^e\}$, i.e., the maximum number of straddle operations in any hour.
v_h	total number of RMG operations in hour h; direct exports count as two operations.
w_{hl}	number of RMG operations in hour $h \in H$ in excess of λ_l, $l \in L$.
p_h	amount of TEU resident in ISA at the end of hour $h \in H$.
q_{hm}	amount of TEU resident in ISA at the end of hour $h \in H$ in excess of μ_m, $m \in M$.

Constraints

Each container has to be moved exactly once:

$$\sum_{h \in W_c} x_{ch} = 1, \quad \text{for all } c \in C.$$

Count the GSI import and export operations per hour:

$$\sum_{c \in I} x_{ch} = t_h^i, \quad \text{for all } h \in H,$$
$$\sum_{c \in E} x_{ch} = t_h^e, \quad \text{for all } h \in H.$$

Compute the maximum number of straddle (GSI) operations needed in any hour. Assuming optimal pairing, the number of straddle operations needed for a particular hour is the maximum of the number of import operations of any hour and the number of export operations of any hour:

$$t \geq t_h^i \quad \text{and} \quad t \geq t_h^e, \quad \text{for all } h \in H. \tag{1}$$

Determine the number of RMG operations (GRI+GSI) per hour:

$$v_h = |C_h| + t_h^i + t_h^e, \quad \text{for all } h \in H.$$

Compute the number of RMG operations exceeding level l:

$$w_{hl} \geq v_h - \lambda_l \,, \quad \text{for all } h \in H, l \in L.$$

Determine the amount p_h of TEU resident in ISA at the end of the hour $h \in H$. On the GRI side, import containers in I_h leave the ISA, whereas export containers in E_h enter the ISA. On the GSI side, the solution variables x_{ch} define the containers that are moved into or out of the ISA in hour h. Thus,

$$p_h = p_{h-1} + \tau(E_h) - \tau(I_h) - \sum_{c \in E} \tau_c x_{ch} + \sum_{c \in I} \tau_c x_{ch} \,, \qquad \text{for all } h \in H. \tag{2}$$

with $p_0 = 0$, being the amount of TEU initially in the ISA.
Compute the number of TEU resident in ISA at the end of hour h in excess of level m:

$$q_{hm} \geq p_h - \mu_m \,, \quad \text{for all } h \in H, m \in M. \tag{3}$$

Objective

Define penalty coefficients d^t, d_l^w, d_m^q, d^x with $d^t \gg d_l^w \approx d_m^q \gg d^x$, $d_l^w < d_{l+1}^w$ and $d_m^q < d_{m+1}^q$. Let $\delta(c, h)$ be the dwell time of container $c \in C$ in the ISA, if it is moving to or from the GSI in hour h. Our objective is

$$\begin{aligned} \min \quad & d^t t && \text{minimize maximum number of straddle carriers} \\ & + \sum_{h \in H} \sum_{l \in L} d_l^w w_{hl} && \text{penalize high RMG utilization} \\ & + \sum_{h \in H} \sum_{m \in M} d_m^q q_{hm} && \text{penalize high ISA utilization} \\ & + \sum_{c \in E} \sum_{h \in W_c} d^x \delta(c, h) x_{ch} && \text{penalize long dwell times} \end{aligned}$$

While this model looks rather simple, it is the result of extensive examinations of several approaches. Note that after a lower bound for t is found by solving the model, t can be dropped out of the objective and given suitable bounds that leave some room for improvement of the other objectives.

To encourage moving containers in time slots corresponding to their corridors, there are two possibilities that both work satisfactorily. The first is to give some benefit d_{ch}^x for movements within the corridor. The other method is to force adherence to the corridor by restricting W_c accordingly. For the computations shown later, the corridor was strictly enforced. In this setting we apply constraints (2) and (3) separately to each corridor. Reefer containers can be easily incorporated by duplicating Eqs. (2) and (3) restricted to reefer containers.

The model tries to minimize the maximum number of straddle carriers needed in any hour. In practice, one would expect to get a list with the maximum number of available straddle carriers for each particular hour, subject to the number of carriers needed for other work in the terminal, e.g., unloading a ship. This can be easily incorporated by introducing non-negative integer variables t_h for each hour $h \in H$ and a constraints similar to (1):

$$\omega_h \geq t_h \geq t_h^i \quad \text{and} \quad \omega_h \geq t_h \geq t_h^e, \ \text{ for all } h \in H,$$

where ω_h is the maximum number of straddle operations allowed in hour $h \in H$.

4.1.2 Results—Stage 1

The IPs resulting from the above model typically have about 350,000 variables, 50,000 constraints and one million non-zeros when using the data for a whole month. The modeling was done using ZIMPL (Koch 2004). The solving time with CPLEX 10.0 on Sun V40z is less than 5 min.

An important factor for the size of the model is the time windows for the containers. We assumed that truck arrivals for import containers are booked 24 h in advance. The time window for import containers is therefore 20 h, as the container has to be in the ISA 4 h before the truck arrives. If the container is discharged from the ship less than 24 h before the truck arrives, the time window is shortened accordingly. For export containers, the time window starts with the arrival of the truck, and ends 12 h before the ship arrives with a maximum of 192 h. If the truck arrives less than 12 h before the ship the arrival time of the ship is used as the end of the time window.

When solving the problem for a whole month of data, the truck booking times for export containers are neglected. In reality it would be necessary to repeatedly solve the model with a moving time horizon. This will lead to significantly smaller IPs. On the other hand, due to an increased knowledge horizon, the solution we present here constitutes an upper bound on the performance in practice.

Figure 6 shows the number of RMG operations per hour. An operation includes moving the RMG to the container, hoisting it, moving it to its destination and lowering it. Whenever a container enters or leaves the ISA this is counted as one operation. This means a move going directly from the GRI to the GSI, i.e., a direct export, counts as two moves. Since this is the longest possible move, and because we do not take precise moving times into account, this is giving an acceptable estimate.

In real operations the ISA design has an overflow mechanism, i.e., whenever more trucks arrive than can possibly be processed by the RMGs, there are the opportunity to serve them directly by straddle carriers. This allows one to dimension the ISA facility to run most of the time at a high utilization level. In this investigation we did not take this overflow mechanism into account. As a result, at certain hours more than 150 RMG operations have to be scheduled, since more than 150 trucks arrive. As mentioned before, it is possible to further balance the number of RMG operations from hour to hour at the expense of using either more straddle carriers or increasing the utilization of the ISA.

Figure 7 depicts the ISA utilization. Note that the number of 40′ containers is much less varying than the number of 20′ containers. This has the effect that whenever a high number of containers is in the ISA, the percentage of 40′ containers is below average. The solution stays completely below 1,400 TEU, i.e., stacked two containers high,

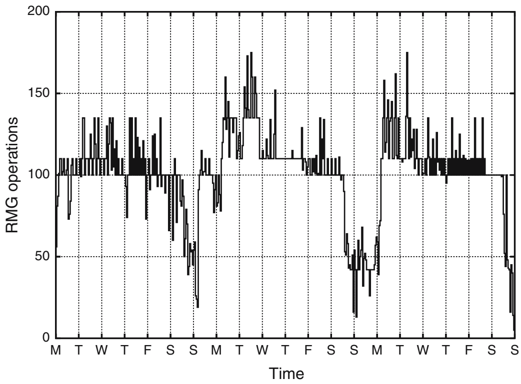


Fig. 6 Number of RMG operations per hour

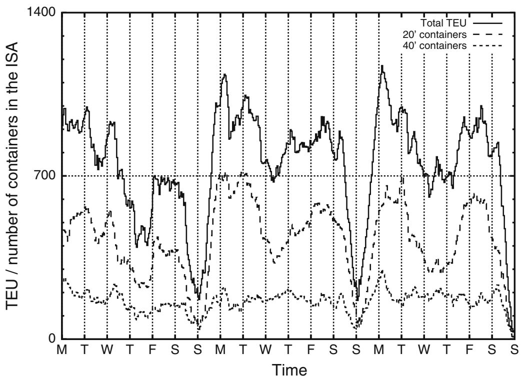


Fig. 7 ISA utilization per hour in TEU

even though we could fill up the ISA completely. The two sharp increases in the ISA utilization between Sunday and Monday result from the short time windows for the import containers. We simulated truck booking times of 48 h in advance instead of 24 h in advance and these spikes were mostly gone. Since the real booking system requires trucks arriving on Monday to be booked on Friday evening, further improvements are possible.

Figure 8 depicts how long containers stay in the ISA. Since we forced container movement to the time corridors, all dwelling times are multiples of four. Containers staying zero hours in the ISA are direct export containers. Apparently about one-third of the export containers do not touch the ISA. All containers staying more than 24 h in the ISA are also export containers, due to the time window for import containers. The peak at 4 h consists mainly of import containers, meaning that most of the import containers are delivered from the yard to the ISA at the latest possible time, which is 4 h before truck pick-up. The other peak from import containers at hour 24 is due to their time windows expiring. Some export containers stay more than 5 days in the ISA, even though this is penalized by the objective function of the IP.

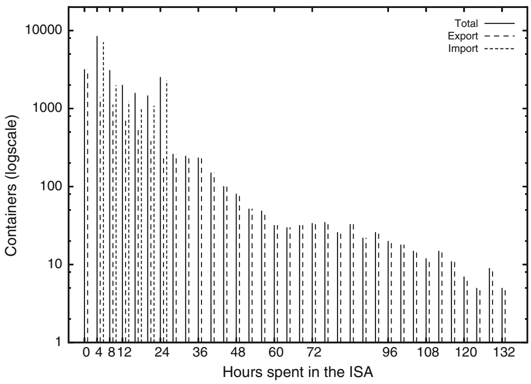


Fig. 8 Distribution of ISA dwelling time

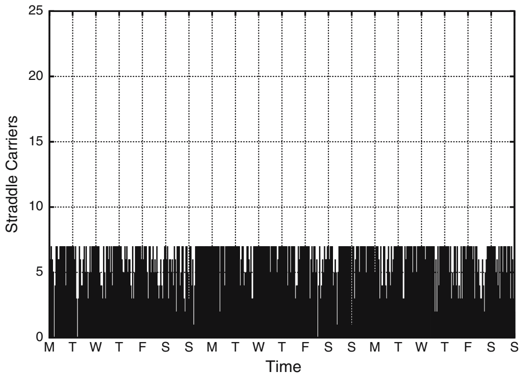


Fig. 9 Straddle requirements after Stage 1 optimization with a $24 - 4 = 20$ h effective time window for import containers, and if both import and export containers reside in the GSI for only 1 h

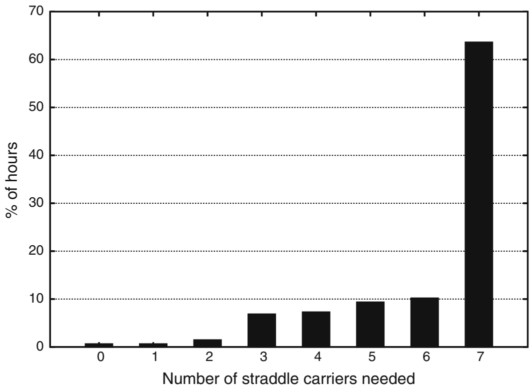


Fig. 10 Distribution of straddle fleet utilization

Some additional experiments have confirmed that these long dwelling times are indeed necessary.

In Fig. 9 we see the number of straddles needed if we assume that import containers have to be delivered in the hour before their movement into the ISA and export containers have to be removed from the GSI to the main yard in the hour after they were put into the GSI. Only seven straddle carriers are needed. Thus, by strategically using the ISA, we can significantly reduce the "no ISA" or "just in time" straddle requirement of twenty two straddle carriers shown in Fig. 4. The seven carriers is in fact optimal; we verified this by solving the Strategic IP (see Sect. 4.1) with the objective to minimize the maximum number of straddle carriers needed.

Our GSI utilization is 215 TEU on average and 414 TEU at maximum, which is well below the maximum capacity of 792 TEU. In practice, it is likely that the time a container stays in the GSI is more than 1 h. Since there is enough room for lengthier container dwell times in the GSI, there may be additional opportunities for pairing straddle moves or otherwise improving straddle usage for a particular hour.

Figure 10 shows the distribution of the number of straddles needed per hour for the usage pattern shown in Fig. 9 assuming a 1 h GSI time window.

4.2 Stage 2—A set of tactical IPs to determine GSI locations

In each hour a number of import containers has to be placed into the GSI by straddle carriers. The position in the GSI determines which RMG will handle the container. To allow a balanced operation, the distribution of containers between the RMGs has to be balanced with regard to the number of containers, the actual container volume (TEU), and the ISA dwelling time. Furthermore, the containers shall be placed in the correct moving corridor for the particular hour if possible. The problem described below can be shown to be $\mathcal{NP}$-hard as it contains the $\mathcal{NP}$-hard scheduling problem of minimizing the makespan on parallel machines as a special case. Nevertheless, the even distribution of containers is crucial for the good performance of the online algorithm in the third stage. Therefore, the GSI positions are determined by solving an integer program for each single hour.

4.2.1 Model description—Stage 2

Let C be the set of containers, P be the set of possible positions (2 rows of 132 slots with up to three containers as described in Sect. 3.2), and R the set of RMGs. Each position in P is assigned to a specific RMG in R and $P(r) \subset P$ is the set of feasible positions for RMG $r \in R$.
Define variables $x_{cp} \in \{0, 1\}$ with $c \in C$ and $p \in P$,

$$x_{cp} = \begin{cases} 1, & \text{if container } c \text{ is placed at position } p; \\ 0, & \text{otherwise.} \end{cases}$$

Further define non-negative integer variables z_r which equal the number of containers assigned to each RMG $r \in R$. Moreover, we have two non-negative integer variables $\bar{z}$ and $\underline{z}$, which equal the maximum and minimum number of containers assigned to any RMG, respectively. Now, we have the following constraints.

Each container has to be placed somewhere, that is,

$$\sum_{p \in P} x_{cp} = 1\,, \quad \text{for all } c \in C.$$

For each RMG, we count the number of containers assigned to it

$$\sum_{c \in C} \sum_{p \in P(r)} x_{cp} = z_r\,, \qquad \text{for all } r \in R.$$

Compute the minimum and maximum number of containers

$$\underline{z} \le z_r \le \bar{z}\,, \qquad \text{for all } r \in R.$$

The constraints for TEU and dwelling time are similar. Finally, we use the following objective function:

$$\min \;\; \bar{z} - \underline{z} + \sum_{c \in C} \sum_{p \in P} d_{cp} x_{cp}$$

with $d_{cp} \ll 1$ and $d_{cp_1} \ll d_{cp_2}$ if p_1 is a position on the ground and p_2 is a stacked position, indicating that a position on the ground has to be used before a position on top of the stack. The objective function additionally contains terms to penalize differences in TEU and dwelling time between the RMGs simlar to $\bar{z} - \underline{z}$.

4.2.2 Results—Stage 2

This approach leads to about 750 IPs with up to 50,000 variables, 40,000 constraints and 270,000 non-zeros. We did not solve all IPs to optimality, but stopped after a time limit of 120 s. While some of the IPs are very hard to solve to optimality, we believe that usually the optimal solution is found. This approach worked quite well. The overall difference in the workload of the RMGs is less than half a percent regarding the assigned number of containers, TEU, and dwelling time.

4.3 Stage 3—An online algorithm to serve the GRI and GSI

The strategic planning in Stages 1 and 2 has decomposed the problem into several subproblems, each with a 1 h planning horizon. Due to the unknown arrival times of trucks within each hour, an online algorithm must be applied. The decisions to be made by the online algorithm are: deciding the order in which trucks are served by RMGs, deciding when GSI containers are moved by RMGs, assigning containers to temporary positions in the ISA and tracking which containers lie above or beneath other containers, and precise tracking of the movements of RMGs at the level of seconds.

4.3.1 The algorithm—Stage 3

The online algorithm maintains two queues for each RMG. Firstly, the "admitted" queue contains jobs to be processed by the respective RMG. The order and timing for jobs in this queue are fixed. Whenever a new truck arrives, it is immediately assigned to an RMG. Then an RMG job is constructed depending on the truck: a truck delivering a container creates a "GRI to ISA" move, whereas a truck that picks up a container creates an "ISA to GRI" move. This job is added to the corresponding "admitted" queue. For import containers the RMG is fixed by the location of the container in ISA. In contrast, an export container is assigned to the RMG that will become idle first after processing all jobs in its "admitted" queue.

All movements between GSI and ISA are known at the beginning of an hour. These jobs are placed into the "open" queue. If an RMG is idle then it takes the *nearest* job from the "open" queue and adds it to the "admitted" queue. The *nearest* job is a move that starts at a container position that has minimum distance to the current RMG position.

For a job to be admitted, it has to be *feasible*, i.e., a target position has to be available. Containers are only stacked onto each other if the lower container will be moved at least 4 h later. Whenever a move is not possible because no feasible stacking position can be found in the corridor, it is checked if a *packing operation* is possible. We define a *packing operation* as taking a container that occupies the bottom position of a slot and moving it to some feasible location, freeing the slot; see Fig. 11.

There are two types of packing operations: in the so-called "nice" packing operation, containers are only moved in the right direction, i.e., towards the GRI for import containers and towards the GSI for export containers. If no nice packing can be found, any packing is considered. If no position is found within the moving corridor, the full moving area of the RMG is searched. Whenever an RMG is idle, both queues are empty, and there is a nice packing operation, it is executed.

The algorithm as described has no way to handle the situation whereby a truck with an export container arrives, is assigned to an RMG, and cannot find a feasible position within the RMG's moving area in the ISA to place the container. However, we now argue that this situation cannot occur. If there is no feasible position in an RMG moving area and no packing operation is possible, the RMG moving area must be near its capacity of 420 containers. However, Stage 2 ensured that the distribution of containers over each RMG moving area is approximately the same. This means the *total* utilization of the ISA has to be very high, which cannot happen, as Stage 1 ensured that the overall ISA utilization stays within reasonable limits.

Details on stacking and packing: Suppose that the truck arrival for an import container X is booked 12 h in advance. Moving X from the yard into the ISA might take up to 4 h. Once X is in the ISA, which other containers can be possibly stacked upon it? Note, that any container to be stacked on top of X has to be removed at least 2 h ahead of the time X is scheduled for moving (this is a constraint enforced by Patrick Corporation). This leaves only 6 h usable for stacking. If the moving corridor scheme is employed this is further reduced to 4 h, because containers to be stacked at this position will only arrive in 4 h intervals. This leaves only containers to be stacked on top of X that are delivered in the same hour to the GSI, but move to the GRI 4 h earlier.

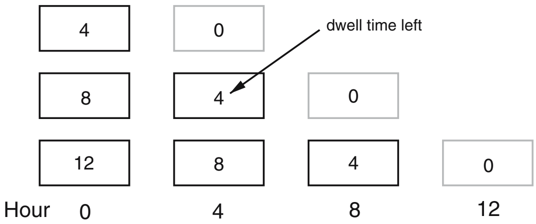


Fig. 11 Packing operation

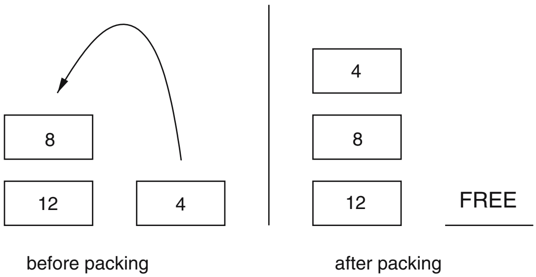


Fig. 12 Slot occupation time

```
1 Put all GSI ↔ ISA jobs into the "open" queue
2 if a truck arrives then
    construct a job and add it to the "admitted" queue
3 if the "admitted" queue is not empty then
    process the oldest job from "admitted" queue
    Goto 2
4 if the "open" queue is not empty then
    if there is a feasible job then
        select the job and move it to the "admitted" queue
    else
        if a packing operation is possible then
            put the packing job into the "admitted" queue
    Goto 2
5 if a nice packing job is possible/available then
    put the packing job into the "admitted" queue
6 Goto 2
```

Algorithm 1: Online sequencing of container moves

Figure 12 depicts what usually happens regarding import containers. Even if we succeed in stacking three containers high, it takes 16 h until the position is available again for new containers.

4.3.2 Results—Stage 3

The above online algorithm has been used as part of a simulation environment. The simulator can only handle 20′ containers. To get usable results two of the seven rows

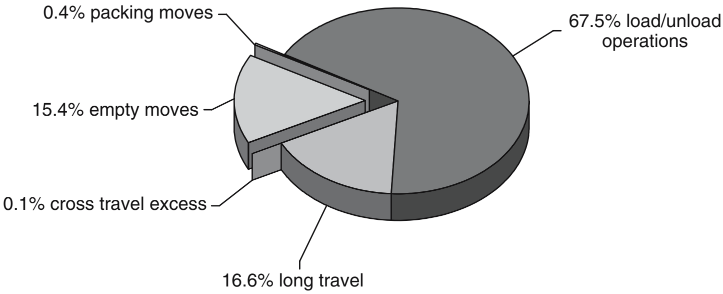
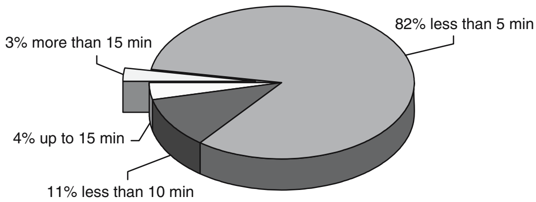


Fig. 13 RMG moving time distribution

Fig. 14 Truck waiting time distribution



of the ISA were blocked in advance. Since 43 % of the containers are 40′ this leads to approximately the correct adjusted ISA size TEU-wise.

Note that for the following results each hour was computed separately. The simulation ran until all the work for the hour was completed, even if this meant running for more than 1 h. In a real continuous operation this is of course impossible. The reasons for doing this were twofold: first, there are hours which clearly exceed the possible capacity of the RMGs. Lacking the overflow procedure possible in practice, these hours would spill the results into several succeeding hours. Second, computing hour by hour allows a much clearer insight into what is happening and where the performance thresholds are.

Figure 13 shows how RMGs spend their time. The most time-consuming tasks are the load/unload operations which are unavoidable apart from increasing the number of direct export containers. The time needed for long travel is also unavoidable. RMG movements with a container are nearly optimal: only 0.1 % of time exceeds the time needed for the mandatory long travel. This shows that dividing the RMG areas into hourly corridors works remarkably well. About 15.5% of the time the RMG moves without a container. Since we assume that each truck only delivers or picks up a single container, there has to be an empty move between two loaded moves. If we assume the empty moves to be the shortest possible moves (for example, to an adjacent GRI/ISA/GSI position) nearly half of the empty move time is unavoidable. The time needed for packing containers can essentially be neglected as it only comprises 0.4 % of the total time. Altogether we can say that under the assumptions used, the RMG operation is within 8 % of the theoretical optimum. While there is no guarantee that the Stage 3 algorithm will perform as well on every possible data set, we remark that the data used in this study describes a high-load situation.

One of the operational goals was to have a waiting time for trucks of less than 15 min. As can be seen in Fig. 14, this goal is met for 97 % of all arriving trucks.

Why does excessive waiting happen for 3% of arriving trucks? In some instances, this is due to the conversion of trains to trucks: a large number of trucks artificially appear simultaneously at the beginning of an hour and the RMGs cannot service them quickly enough, even without packing and with no cross travel. For other hours, the number of containers within the hour is just too high to be serviced in the available time.

5 Conclusion

In this paper, we studied the problem of managing a container exchange facility with multiple RMGs. We proposed an integer programming-based heuristic consisting of three stages. Computational experiments show that the strategic IP (in Stage 1) works well: we compute feasible schedules that are globally optimal. This global perspective minimizes and balances the resource usage to a large extent and enables an online algorithm (in Stage 3) to work effectively. The space–time divisioning of the corridor construction leads to very efficient RMG operations. It is evident from Fig. 13 that the performance of the facility heavily depends on the time needed for loading/unloading operations. We have demonstrated that if the ISA is intelligently controlled it can cope with the anticipated levels of container throughput with the planned RMG exchange infrastructure. Moreover, our framework provides a valuable guide on the minimum size of the required straddle carrier fleet.

International container exchange facilities are becoming increasingly sophisticated with higher degrees of automation of their components. We believe that our approach of time scale decomposition, whereby a strategic master IP controls decisions across a long time horizon to produce a balanced plan that is fed to a series of short time scale online subproblems, can be highly effective for decision making at such facilities. The space–time divisioning employed on the ISA may also be adapted to manage generic storage components of container exchange facilities in a very effective way.

Acknowledgments We thank Chris Knott, Adrian Sandrin, Andrew Rembel, Benjamin Rogers, and David Whitburn from Patrick Technology and Systems for providing the historical data used to test our solution method, and for explaining the technical details of the container exchange facility. We also thank Andre Costa from the University of Melbourne's Department of Mathematics and Statistics for his contributions to the project.

References

Alicke K (2002) Modeling and optimization of the inter-model terminal Mega Hub. OR Spectrum 24:1–17

Ballas A, Golias J (2002) Comparative evaluation of existing and innovative rail-road freight transport terminals. Transp Res A 36:593–611

Bramel J, Chen X, Simchi-Levi D (2005) The logic of logistics: theory, algorithms, and applications for logistics and supply chain management, 2nd Edn. Springer, Heidelberg

Cheung RK, Li C-L, Lin W (2002) Interblock crane deployment in container terminals. Transp Sci 36(1): 79–93

Corry P, Kozan E (2006) An assignment model of dynamic load planning of intermodel trains. Comput Oper Res 33:1–17
Dekker R, Voogd P, van Asperen E (2006) Advanced methods for container stacking. OR Spectrum 28:563–586
Günther H-O, Kim KH (2005) Container terminals and automated transport systems. Springer, Heidelberg
Günther H-O, Kim KH (2006) Container terminals and terminal operations. OR Spectrum 28:437–445
Jaillet P, Wagner MR(2006) Online routing problems: value of advanced information as improved competitive ratios. Transp Sci 40:200–210
Kim KH (2005) Models and methods for operations in port container terminals. In: Langevin A, Riopel D (eds) Logistics systems: design and optimization, Gerad 25th Anniversary Series, chap. 7. Springer, Heidelberg, pp 213–243
Kim KH, Park YM, Ryu KR (2000) Deriving decision rules to locate export containers in container yards. Eur J Oper Res 124:89–101
Koch T (2004) Rapid mathematical programming. PhD Thesis, Technische Universität Berlin
Kozan E (2000) Optimising container transfers at multi-modal terminals. Math Comput Modell 31:235–243
Kozan E, Preston P (2006) Mathematical modelling of container transfers and storage locations at seaport terminals. OR Spectrum 28:519–537
Lee LH, Chew EP, Tan KC, Han Y (2006) An optimization model for storage yard management in transshipment hubs. OR Spectrum 28:539–561
Murthy CSR, Manimaran G (2001) Resource management in real-time systems and networks. MIT Press, Cambridge
Ng WC (2005) Crane scheduling in container yards with inter-crane interference. Eur J Oper Res 164:64–78
Pruhs KR, Sgall J, Torng E (2004) Online scheduling. In: JY-T, Leung (ed) Handbook of scheduling: algorithms, models, and performance analysis, chap 15. Chapman & Hall/CRC
Steenken D, Voß S, Stahlbock R (2004) Container terminal operation and operations research—a classification and literature review. OR Spectrum 26(1):3–49
Vis IFA, de Koster R (2003) Transshipment of containers at a container terminal: an overview. Eur J Oper Res 147(1):1–16
Zhang C, Wan YW, Liu J, Linn R (2002) Dynamic crane deployment in the container storage yards. Transp Res B 36(6):537–555