Chapter 2
Maritime Terminal Operational Problems

Chapter 2 provides an introduction to maritime terminal configuration and operational problems. The structure of typical container terminals and associated decision problems are discussed. A case study of one of the largest container terminals, PSA Singapore Terminals, is given. A literature review on selected key problems in container terminals such as berth allocation, quay crane scheduling, transportation, yard crane scheduling, and the integration of decision problems are conducted.

2.1 Container, Container Vessels, and Configuration of a Container Terminal

Malcolm McLean was the first freight forwarder who used specially designed containers on its vessel in 1956 (Frank 2009). Standard sized containers can be efficiently stacked and be moved seamlessly between ships, trucks and trains. Vessels, trains, trucks and cranes can be designed to a single size specification. The use of standard sized containers for cargo shipment has been proved to be able to substantially reduce the handling time of vessels and reduce the amount of labour needed. Containers, container vessels have gained popularity rapidly in the past several decades.

The basic container unit is of size 20’×8’×8’6” (length × width × height), also referred to as a TEU (Twenty-foot equivalent unit). Cargo volume and vessel capacity are usually measured in TEUs. The 20-foot and 40-foot lengths are the two most important and most commonly used sizes. The 20-foot container, referred to as a TEU, has become the industry standard reference. The 40-foot length container, also known as the Forty-foot Equivalent Unit (FEU), has become the most frequently used container size today.

To enjoy economics of scale, the size of a container vessel has been increasing in the past few decades. In 2006, the world’s largest containership, EMMA MAERSK, has a capacity of 11,000 TEUs (PSA International, 2006). Today, the largest capacity of a container vessel has reached 18,000 TEUs (Maersk Group, 2014).

To meet the demand for an increasing number of container shipments, existing seaport container terminals were expanded, new container terminals, especially large ones, were constructed in recent years.
Nowadays, the top container terminals in the world include Shanghai, Singapore, Hong Kong, and Shenzhen (Salisbury 2013).

A container terminal (CT) is a very important node in container transport, which is used mainly to serve container vessels. Vessels are loaded and unloaded in the CT and containers are temporarily stored before moving to their next mode of transport—rail, road, or sea. A CT is depicted in Figure 2.1.

![Fig. 2.1. Corner of a CT (source: PSA Singapore Terminals)](image)

The layout of a typical container terminal is shown in Figure 2.2.

![Fig. 2.2 Typical container terminal](image)

Vessels mooring at the berths wait for quay cranes (QCs) to upload/discharge their containers. The PMs (prime mover) shuttle between the QCs and yard cranes (YCs) to move these boxes from the berth to the container yard area, and vice versa. Upon arrival at the blocks, the PMs queue in front of the YCs until they are served. Container yard areas serve as buffers where
containers are temporarily stored. PMs also shuttle between YCs and gates to move boxes to/from YC to/from hinterland through gates. Figure 2.3 shows the container vessels mooring at a CT.

Fig. 2.3 Vessels mooring at a CT (Source: PSA Singapore Terminals).

Three types of containers, export, import, and transhipment, may coexist in a container terminal. The flows of the three types of containers are shown in Figure 2.4.

Fig. 2.4 Container flows in a CT
Export containers are delivered from the hinterland by truck or train to the terminal gates. The containers are then checked and the transportation documents are processed at the gates before moving to the stack yards by the PMs. The YCs unload the containers from the PMs and put them into temporary storage in the yard stacks. When the vessel into which they are to be loaded arrives at the terminal, these export containers are retrieved from the stack yards by the YCs and carried by the PMs to the berth. The PMs send the containers to the appropriate QCs which lift the containers and load them onto the mother vessel. These export containers are then delivered to their destination ports by the vessels.

Import containers arrive by vessel and leave the terminal via truck or train to the hinterland. An import container is first unloaded from a vessel by a QC which puts it onto a PM. The PM then moves the container to a storage yard where a YC unloads it from the PM and places it in the stack yard for temporary storage. Later, these import containers are retrieved from the stack yard by the YCs and carried by the PMs through the gate to its hinterland destination.

A transhipment container transships from one (origin) vessel to another (destination) vessel with the help of a CT. Upon arrival of the origin vessel, a transhipment container is unloaded by a QC, moved from the quay to the yard by PMs, and then temporarily stored in a stack yard by a YC. Later when the destination vessel arrives at the terminal, the container is retrieved by a YC, moved from the yard to the quay and then loaded by a QC into the destination vessel. The destination vessel then unmoors from the berth and heads for its next port. A CT thus serves as a buffer (in space and time) between two container vessels.

Fig. 2.5 Schematic diagram of a CT
A CT typically consists of several areas: quay area, transport area, yard area, and hinterland, as shown in Figure 2.2. The hinterland refers to the origin of export containers and destination of import containers. External trucks (hauliers) or trains deliver containers from/to inland origins/destinations. Figure 2.5 is a schematic diagram of a CT. Vessels, QCs, and YCs are labelled with numbers for easy identification. The small black rectangles represent the PMs.

In Figure 2.6, we can find the quay area, transport area, and yard area of a CT.

Fig. 2.6 Quay area, transport area, and yard area of a CT (Source: PSA Singapore Terminals).

We describe the quay area, transport area, yard area in more detail as follows.

### 2.1.1. Quay Area

The quay area provides the berths for vessels to moor. The quay cranes perform the loading and unloading operations of containers at a quay area. After the arrival of a vessel, a berth is assigned to the vessel. The QCs load/unload the containers which are delivered by the PMs from/to a yard area. A QC handles the containers in a different section along the length of the vessel known as a hatch. Typically, three to four QCs work on a vessel at a time. Figure 2.7a shows some QCs operating in a CT. In Figure 2.7b, a QC is unloading a vessel and putting a container onto a PM.
There are two types of QCs: single-trolley and dual-trolley. A key indicator for the productivity of a quay area is the number of QC loading/unloading operations (called moves) per hour. QCs can operate 50-60 moves/hour technically (Dirk et al. 2005). The actual performance of the QCs is however much lower: around 30 moves/hour. The average number of lifts achieved at a terminal per QC working hour is known as the GCR (gross crane rate). The GCR is a very important performance measure of a CT. Careful synchronization of the operations of PMs and YCs is the key to improving QC performance. The flow of containers should be as smooth as possible and the idle time of QCs waiting for yard trucks (YTs) should be minimized. This bears some similarity to the concept of JIT (Just-In-Time) where the flow of jobs should be smooth and levelled. We note in Chapter 3 that a CT is often bottlenecked by slow YC movements.

### 2.1.2. Transport Area

In transport area, containers are moved back and forth by the transport equipment. There are two transport areas: quayside transport area which is between the quay area and yard area, and the landside transport area which is between the yard area and the gates. At a CT, the transportation provided by PMs between QCs and stack yards or between stack yards and gates are usually referred to as horizontal transportation. The loading/unloading movements of the QCs and the YCs are referred to as stacking transport.
The most common PMs are the YTs (Figure 2.8) and straddle carriers. The YTs are unable to lift the containers. Thus, the QCs and YCs are required to load/unload containers from/to the trucks. On the contrary, the straddle carriers can lift the containers themselves, which can increase the productivity of the QCs. However, the straddle carriers require higher purchasing, maintenance, and operational costs (Frank 2009). The other types of equipment, such as the Automated Guided Vehicles (AGVs), are also often used. AGVs can reduce the labour cost of a terminal because they do not need drivers but their transport productivity and average speed are relatively low (Frank 2009). Nowadays, the AGVs are usually not used between gates and stack yards. Terminals may use any of the PMs mentioned above. In this book, we focus only on the YTs. Thus, the PMs in this book refer to the YTs.

2.1.3. Yard Area

A CT acts as a buffer between the various trade routes. Vessels from different trade routes rely on the buffer to maintain their schedules which have substantial economic impact for the global economy. The physical entity lying at the heart of this buffer is the yard area, which is used for intermediate storage of the containers. The export, import, or transhipment containers are temporarily stored in yard area before they are sent to their final destinations. The yard area is the buffer to link the different types of transportation vehicles or the same type of vehicles but arrive at different time periods. In some CTs such as PSA Singapore, the majority of yard activity comes from the transhipment operations. Figure 2.9 shows a yard area at PSA Singapore.
A yard area is typically partitioned as contiguous rectangular blocks (see Figures 2.5 and 2.10) where a grid has been painted on the pavement indicating the positions where containers should be placed. In Figure 2.5, we illustrate 9 yard blocks. Bi-directional traffic lanes for the PMs occupy the space between the blocks. A typical block is six stacks (6*8.5 feet) wide and thirty stacks (30*20 feet) long. Blocks are divided along their length into 20 ft sections called slots. Each slot has several rows. The containers are stored alongside in each row and are stacked on top of each other. A typical block is six rows (6 x 8.5 feet) deep and forty slots (40 x 20 feet) long.

The YCs perform the stacking and retrieval operations in this area. The YCs transfer the cargo into and out of the storage blocks. The YCs straddle above the containers in each block and traverses parallel to the length of a block. A yard zone is defined as a group of adjacent blocks that are aligned lengthwise and that together they form a single lane for yard crane movement. In Figure 2.5, we illustrate three yard zones (e.g. block 1-3 are in zone 1; block 4-6 are in zone 2, etc). A YC can access all storage locations in a yard zone without changing lanes. The YCs move easily within yard zones but take a long time to move between different zones. The movement of
the YCs within (between) the zones is called a linear gantry (cross-gantry) movement. A yard area typically has space for many stacks of containers, with up to 6 containers per stack.

The YCs can be rail mounted gantry cranes (RMGCs), rubber tired gantry cranes (RTGCs), or overhead bridge cranes (OBCs). Figure 2.11 shows an OBC moving a container inside a yard area. The RTGCs are more flexible in operation. To avoid operations interruptions, the neighbouring YCs are kept apart from each other a safety distance, which is a multiple of a slot. The YCs operate approximately 20 moves/hour technically (Dirk et al. 2005). The tightest constraint on the movement of the YCs and PMs is the slow gantry movement of the YCs and the extremely slow cross-gantry movement of the YCs. In particular, if a YC wishes to move from its current zone to another yard zone, the minimum amount of time required is around 15 minutes. In addition, transhipments are the most crucial operations as they involve the vessels directly. A delay in transhipment operations (especially during a vessel loading) carries a very high immediate cost.

![Fig. 2.11 OBC moving a container in a yard area (Source: PSA Singapore Terminals).](image)

**2.1.4 Practical Example of CT: PSA Singapore Terminals**

As the world’s largest transhipment hub and the second largest container terminal (in terms of volume of TEUs handled) in the world, PSA Singapore Terminals handled 32.24 million TEUs in 2013 (Factsheet, 2014). It connects the cargo markets of the West (Europe, Africa, Middle East, and South Asia) with that of East Asia. PSA Singapore Terminals moves a seventh of the world’s shipping containers (Credit Opinion, 2014).
In Singapore, PSA Singapore Terminals operates seven container terminals (Tanjong Pagar, Keppel, Brani, Pasir Panjang Terminals 1, 2, 3 and 5) offering connections to 600 ports globally. There are a total of 57 container berths over a quay length of 17,350 meters (Factsheet, 2014). The layout of Keppel Terminal is depicted in Figure 2.12. There are 14 container berths in its 3200 meter quay length. A total of 40 QCs are working in this 105 hectare terminal (Factsheet, 2014).

In Figure 2.12, we can see a large yard area (with many long brown bars representing yard blocks). The quay area is set up along the sea side with white boxes representing the vessels and the short bars representing the QCs.

![Fig. 2.12 Layout of Keppel Terminal (source: Factsheet, 2014)](image)

### 2.2 Decision Problems

With the ever increasing international container traffic, CT operations are becoming more and more important, busy, and complex. Competition among the CTs is also increasing: The CTs compete with each other in providing better customer service to attract more container shipments. The CT operators such as PSA Singapore face constant pressure to reduce the turnaround time and increase flow efficiency. Efficient computational tools which can help increase container throughput are key. Large CTs have reached a degree of complexity that further improvements require scientific methods (Steenken et al., 2005). Many publications on CT operations have appeared in the past several decades. These publications provide the different tools for different types of decision problems in the CTs.
The decision problems in a CT are categorized into several different categories by different researchers. Böse (2011) introduced a three level model for CT planning:

Level 1 (Planning of terminal infrastructure): deals with decisions regarding preparation of the terminal (area, quay wall length, etc.) and terminal connection to the external networks.

Level 2 (Planning of terminal suprastructure): deals with decisions regarding what kinds of resources and in what quantity should be used in a CT. For example, CT layout design.

Level 3 (planning of terminal operation): deals with short-, medium- and long-term planning of the CT’s operations. Level 3 can be further broken down into short-term, medium-term, and long-term planning.

There are interdependencies among the three levels. A lower level provides support and restriction for the higher levels. A higher level generates the requirements to the lower levels. Basically, level 1 is the duty of the local communities and authorities, which is basically out of the scope of academic research (Böse, 2001). Levels 2 and 3 are areas of abundant research. Böse (2001) has compiled the research work on level 2: planning of terminal suprastructure.

Günther et al. (2005) categorized the decision problems in a CT into three categories: design, planning, and real time control problems. The design problems involve the types and the quantity of resources in a CT such as the type of handling equipment in the yard, the number of berths, QCs, YCs, YTs, storage slots, the yard layout, and the human operators, etc. Planning problems deal with how to schedule operations in a CT in advance to better utilize resources to maximize the efficiency of the operations. Real time control deals with the assignment of resources to tasks in real time. The detailed operations schedules can only be developed in a very short time (several minutes) for some resources, such as the PMs, and the YCs. Real time control necessitates very fast algorithms without sacrificing too much on the optimality.

The design problems found in Günther et al. (2005) corresponds to level 2 (planning of terminal suprastructure) in Böse (2011). The planning and real time control problems roughly belongs to level 3 (planning of terminal operations) in Böse (2011).

In this book, we apply the categories proposed by Günther et al. (2005) because level 1 in Böse (2011) is basically out of the scope of our research. We focus on the planning and real time control problems in this book.
2.3. Literature Review

In this section, we provide a literature review on selected key decision problems at different areas. Meersmans (2002), Vis and de Koster (2003), Steenken et al. (2004), Vacca et al. (2007), Stahlbock and Voß (2008), and Gharehgozli et al. (2014) provide excellent reviews of the extant literature on container terminal operations.

2.3.1 Quay Area

The quay area decision problems are critical and have a major impact on a CT’s operational performance. In the quay area, the corresponding decision problems include berth allocation, quay crane assignment, quay crane scheduling, and stowage planning. Stowage planning deals with how to assign the containers to empty slots in a vessel. We will not review the work on stowage planning.

2.3.1.1 Berth Allocation

A specific anchoring berth has to be allocated to a vessel by the CT planners before the vessel’s arrival. The Berth Allocation Problem (BAP) deals with how to optimally allocate such vessels to the berths or quay locations. The berthing (handling) time and the berthing position along the quay length of a vessel have to be determined. The input data to be considered include the technical specifications of the vessels (e.g. vessel length, arrival time), QCs (e.g. length of crane jib), data on the berth type (e.g. berth layout, number, length etc.), projected vessel handling time, mooring time windows, priorities of the vessels, dedicated berth areas, etc. The objectives of BAP include maximizing the productivity of the vessel handling, maximizing customer service levels, minimizing the total service time of vessels, and minimizing the costs.

Lim (1998) formulated the berth planning problem as a mathematical model and proved the problem’s NP-completeness. The problem was then transformed into a two-dimensional packing problem and a heuristic was proposed to effectively solve it. Nishimura et al. (2001) treated the berth allocation problem as a public berth system and applied genetic algorithms to obtain computationally efficient solutions. Guan and Cheung (2004) consider BAP with an objective of minimizing the total weighted flow time. Composite heuristics are proposed for large size problems.

Most BAP treat each arriving vessel equally. However, in actual operations, some vessels desire a high priority service and need to be allocated to the
berths earlier (Imai et al., 2003). The vessels with a large container handling volume generally are also given a higher priority over the smaller and less frequent vessels. In general, ship priority depends on the total throughput per shipping line (Imai et al., 2003). Thus, some researchers have incorporated service priority into the BAP. Imai et al. (2003) modified the BAP formulation to deal with the calling vessels with inhomogeneous service priorities. Instead of assigning weights to represent the vessel priorities, Golias et al. (2009) developed a multi-objective BAP. Different groups of vessels use different objective functions in their models. Vessel service is differentiated upon based on certain priority agreements. For other recent research on BAP, the interested reader is referred to Hendriks et al. (2010), Buhrkal et al. (2011), Du et al. (2011), Hendriks et al. (2012), and Xu et al. (2012).

2.3.1.2 Quay Crane Assignment

After assigning the berths to the vessels, we now need to allocate the QCs to the vessels for loading/unloading operations which are done through the Quay Crane Assignment (QCA) problem. The QCA is sometimes referred to as a crane split (Stahlbock and Voß, 2008). It deals with the allocation of the QCs to the vessels and the vessels’ bays. Given the vessels to be served and available QCs to be used, the QCA has to ensure that the overall loading/unloading operations of vessels can be completed as quickly as possible. The QCA is closely related to the BAP: The solution to the BAP serves as the input to the QCA while the solution from the QCA will in turn affect the BAP.

Peterkofsky and Daganzo (1990) pointed out the importance of crane operational efficiency and developed a branch and bound solution to speed up a cargo ship’s loading and unloading operations. Park and Kim (2003) formulated an integer programming model for the quay crane allocation problem with various practical constraints and designed a two-phase solution procedure for the mathematical model. Chang et al. (2010) applied objective programming to a dynamic allocation model for berth allocation and quay crane assignments based on a rolling-horizon approach. Due to the computational complexity of the problem, a hybrid parallel genetic algorithm was designed to solve the problem and its effectiveness was evaluated via simulation. Han et al. (2010) simultaneously addressed the berth and quay crane scheduling problems and took vessel arrival times and container handling uncertainty into consideration. A mixed integer programming model was proposed and a simulation based genetic algorithm procedure was employed to provide the proactive berth and quay crane schedules. Giallombardo et al. (2010) combined the berth allocation and the quay crane assignment problems and developed a mixed integer
quadratic program which was subsequently reduced to a mixed integer linear program. A heuristic which combined tabu search and mathematical programming techniques was developed to solve the problem.

2.3.1.3 Quay Crane Scheduling

The QCA allocates the available QCs to the vessel bays. The next decision is to decide the sequence of the loading/unloading tasks performed by these QCs, which is done by Quay Crane Scheduling (QCS). The tasks to the QCs are the containers on the vessel to be unloaded and the containers on the PMs delivered from the yard area to be loaded. A QCS model typically consider the following constraints: unloading tasks must precede any loading task, interference between neighbouring QCs, a QC can perform at most one task at a time, and the QC travel speed. The objective is usually to minimize the makespan of all tasks.

Kim, Kang, and Ryu (2004) considered the load sequencing of outbound containers. They decompose the problem into two subproblems: pickup schedule and load sequence. A beam search algorithm was used to solve the model.

In the QCS model developed by Lim et al. (2004), spatial and separation constraints were considered. In their QCS model, QCs cannot crossover; there is a minimum distance between the cranes, and jobs cannot be done simultaneously. Dynamic programming algorithms, a probabilistic tabu search and a squeaky wheel optimization heuristic were proposed to solve the model. Lim et al. (2007) further developed a scheduling procedure which assigns the containers to non-crossing quay cranes to minimize the maximum completion time of all operations. The problem was decomposed into start-time allocation and job-to-crane allocation. A dynamic programming, back-tracking exact algorithms, and a simulated annealing algorithm were employed to solve the proposed problem.

The QCS problem has also been studied by Kim and Park (2004) for quay cranes operating with non-crossing constraints, safety distance (at least one ship-bay), precedence constraints among the tasks in the same bay, and a consideration of container groups. A mixed integer programming model, a branch and bound algorithm, and a greedy randomized adaptive search procedure (GRASP) were proposed in order to minimize the total completion time. Moccia et al. (2006) strengthened the formulation of Kim and Park (2004) and developed a branch and cut algorithm which outperformed Kim and Park’s branch and bound approach. Sammarra et al. (2007) used tabu search to study the same problem as defined by Kim and Park (2004), and evaluated the performance of the tabu search approach against the branch and cut algorithm presented by Moccia et al. (2006), and
the GRASP approach proposed by Kim and Park (2004). Bierwirth and Meisel (2009) resolved the problem of Kim and Park (2004) in a limited solution space by using unidirectional schedules (i.e. the QCs are only allowed to move in one direction) and compared their proposed branch and bound approach with the results of Kim and Park (2004), Moccia et al. (2006), and Sammarra et al. (2007). Kaveshgar et al. (2012) used a genetic algorithm to cope with the drawback of Bierwirth and Meisel (2009)’s model where the QCs can only transfer in one direction.

The problem of identical quay crane scheduling has been discussed by Ng and Mak (2006) who introduced a decomposing scheduling heuristic which divides the vessel into non-overlapping zones, and a dynamic programming procedure was used to obtain the optimal partitioning. Zhu and Lim (2006) formulated the QCS problem as an integer programming model to minimize the overall makespan where the non-crossing constraint was considered, and all jobs in a particular bay need to be finished before the QC could move to another bay. The on-line version of the work of Zhu and Lim (2006) was followed by Zhang et al. (2008) and two heuristics namely (1) one-by-one or on-line list (i.e. without any available information about the remaining operation for consideration) and (2) release dates or on-line time (with regard to available information for release date) were proposed.

Lee et al. (2008a) focused on the QCS problem for a single vessel using handling priority for each ship bay and non-crossing operational restrictions. A mixed integer programming (MIP) model and a genetic algorithm (GA) was presented to minimize the total weighted completion time of every ship bay. In a similar study by Lee et al. (2008b), the QCS problem with non-crossing constraints was addressed for a container ship.

Tavakkoli-Moghaddam et al. (2009) tried to solve the quay crane assignment (QCA) and the quay crane scheduling (QCS) problems simultaneously. An MIP model and GA were used to analyze real-sized problems and to optimize the total completion time of the vessels and the QCs. In comparison with the results for the MIP model provided by the LINGO optimization software package, it was found that the proposed GA was better in terms of the solution times.

A new model for the QCS problem has been developed by Meisel (2011) with respect to the time window constraint for the quay cranes. A mathematical formulation, a lower bound and the unidirectional search heuristic (which is an extension from Bierwirth and Meisel, 2009) were used to deal with the minimization of the total vessel handling time.

Legato et al. (2012) presented a QCS model which includes different processing times for quay cranes for the container group as well as non-identical QCs, non-crossing constraints, time windows, and individual unidirectional schedules. They formulated a MILP, a Lagrangian relaxation
approach for finding a lower bound, and an enhanced branch and bound algorithm based on the work of Bierwirth and Meisel (2009) to solve the scheduling problem.

Lu et al. (2012) studied the QCS problem by considering multiple QCs which are able to process different operations (but not simultaneously) at a single bay (shared bay). A polynomial time complexity heuristic approach was developed to find the optimal assignment and sequencing of containers to contiguous QCs.

### 2.3.2 Transport Area

The transport area serves as the link between the quay area and yard area, and between the yard area and the hinterland. The decisions on the transport area aim to reduce the transport times and ensure QC and YC productivity. High speed QCs or YCs does not necessarily mean high productivity in a CT. Increasing the number or the speed of PMs may not always improve the performance of a CT. Careful synchronization among the transport, quay, and the yard areas is crucial to improving a CT’s performance. The decision problems in the transport area include the allocation of the PMs to the QCs, the allocation and sequencing of jobs to the PMs, and the routing of the PMs. The objectives in optimizing transport operations can be to minimize the transport times, minimize the QC idle times, or minimize the vehicle fleet size.

Nishimura et al. (2005) focused on PM control and considered PM routing. However, the study assumed a relatively static environment where each PM travels along a cyclical path with limited opportunities for re-routing.

There are many papers on AGV dispatching that are in the same vein as PM routing and scheduling. van der Meer (2000) provided a comprehensive introduction to the AGV dispatching methods. The research on multiple-load AGV systems commenced with Bilge and Tanchoco (1997), with contributions from a handful of other authors, including Sinriech and Kotlarski (2002). Levitin and Abezgaouz (2003) considered multiple-load AGVs with LIFO loading constraints. However, their study only considered a single vehicle operating in isolation. van der Meer (2000), Vis and Harika (2004), and Yang et al. (2004) considered a system of automated lifting vehicles (ALVs) for the internal transportation of cargo in container terminals. Evers and Koppers (1996), Cheng (2001), Ioannou et al. (2001), Kim and Bae (2004), among others, studied the AGV system in a CT where a single-load AGV capacity was assumed.

For multi-load AGV systems in a CT, Chan (2001) used simulation to compare two alternative dispatching strategies for the case where there are
n single-level, multi-load AGVs and m containers waiting to be moved. However, the assumption that all m jobs are ready at time 0 does not reflect the true nature of CT operations. Grunow et al. (2004) studied the dispatching of single-level, dual-load AGVs in an automated CT similar to ECT Rotterdam and CTA Hamburg. The AGVs can accommodate one 40’/45’ container or up to two 20’ containers. In these studies, the details of AGV routing and traffic control were ignored and only dispatching was studied. After carefully establishing the ideas of full and partial availability of the AGVs, the authors developed a novel heuristic dispatching algorithm. This algorithm was initiated whenever a new transport order appears within the look-ahead time window. For each partially available AGV, the algorithm generates up to three possible tours for the AGV, each one corresponding to a different ordering of the pick-up/drop-off operations of the new container with respect to the drop-off of the container already on the AGV. One noteworthy finding in Grunow et al. (2004) is that the benefit of dual loading on the AGVs diminishes as the CT becomes larger.

Nguyen and Kim (2009) discussed the dispatching of ALVs. An MIP model was developed to assign delivery tasks to the ALVs optimally. Petering (2010) developed simulated a real-time dual-load yard truck control in a transhipment terminal.

2.3.3 Yard Area

Yard area operations are often a potential bottleneck in a CT (Li et al., 2009). The performance of a CT depends heavily on the decisions made for yard area operations. In the yard area, the corresponding decision problems include assigning the yard blocks for the calling vessels, determining the storage locations for individual containers, yard remarshalling, and Yard Crane Scheduling (YCS). The YCS has two subproblems: YCS-I which deals with assigning the YCs to several yard blocks and YCS-II which deals with scheduling individual YC storage/retrieval operations within a block. In this book, we focus on YCS-II which has some unique features that make it unlike the other problems often studied with respect to material handling systems in the literature. For example, two working YCs in the same lane must remain separated by at least 160 feet; the YCs interact with the PMs only and there is no buffer at the YC/PM interface. A PM must be present at a location in order for a YC to complete handling at that location. YCS-II is also affected by the QC operations: The QC work list determines the target times and locations of storage and retrieval container moves. While the literature contains many articles on CT operations, only a few studies focus on YCS-II. We will provide more literature review on YCS-II in Chapters 3-7.
For YCS-I, Linn et al. (2003) introduced an MIP model for optimally deploy the YCs within a storage yard including cross gantry moves. The aim of this YC deployment procedure was to move the idle YCs to the other blocks at the beginning of the next planning horizon to even the YC workloads. Linn and Zhang (2003) studied the problem of finding an optimal number of YCs in each yard block for every planning period with cross gantry moves and developed an MIP model and a least cost heuristic to minimize the YC workload overflow.

Legato et al. (2009) investigated the deployment of multiple YCs among the yard blocks with particular respect to the linear- and cross-gantry movements. An integer programming model and a discrete event simulation were proposed in order to find the block pair for each YC, and to rank and select five proposed cross gantry deploying policies. The numerical results showed that the “crane transfer to the closet yard block” policy outperformed the other proposed configurations and rules.

Petering et al. (2009) studied the impact of a real time YC dispatching system in respect of the long term performance of a CT with the aim of obtaining a high quay crane usage rate in a transhipment terminal. Twelve rule-based yard crane dispatching algorithms and look-ahead yard crane dispatching systems were used in a discrete event simulation model to evaluate their performance.

The problems of linear- and cross-gantry interference have also been considered by He et al. (2010) in order to optimize the productivity of yard crane operations. Their proposed model minimized the total number of delayed jobs and the total time of YC inter-block moves by proposing a hybrid algorithm based on seven heuristics and a parallel genetic algorithm. Yan et al. (2011) used a knowledge-based approach for the YCS problem and improved the results compared with the approach previously adopted by He et al. (2010). Chang et al. (2011) presented a model for the YCS problem which assumed up to two YCs per block, cross gantry moves, and container groups to minimize the total job delay at the blocks. A dynamic rolling-horizon decision procedure and an integer programming model were developed to solve this problem. An integrated heuristic algorithm and simulation model were proposed, and a GA, which used the initial results from the heuristic algorithm, was introduced to cope with the scheduling problem. The computational results and simulation indicated that proposed procedure outperformed the static rolling-horizon technique presented by He et al. (2010).

Guo et al. (2011) noted that the problem of yard crane scheduling is NP-hard and hence adopted a three level hierarchical procedure for yard crane operation management, namely (1) the distribution of the YCs to different rows, (2) dispatching the YCs in a row, and (3) YC dispatching in individual
zones. Within this approach, their study focused on the single yard crane dispatching at Level 3. An evaluation of the two new heuristic algorithms that were used to minimize the average vehicle delay at the yard showed to be promising. In a follow-up study, Guo and Huang (2012) focused on Level 2 of the hierarchical procedure in Guo et al. (2011), and proposed a time partition algorithm and a space partition algorithm to arrange the YC workload in a single row of yard blocks. It was found that the use of a time partitioning algorithm outperforms that of a space partitioning algorithm.

The real time multiple YCS problem, and the associated consideration of loading operations and linear gantry movements has also been studied by Chen and Langevin (2011) in order to minimize the completion time of the handling of all the outbound containers. The model considered a minimum safety distance of five yard-bays and inter-crane interference between two adjacent yard cranes, also a maximum of two yard cranes in each block at any time. An MIP model was constructed, and GA and tabu search were utilized to obtain near optimal results.

Sharif and Huynh (2012) investigated both the centralized and decentralized methods for scheduling multiple YCs. A comparison of the efficiency of these two methods in respect of five solution qualities (optimality, scalability, computational efficiency, adaptability, and fault tolerance) in truck operations showed that the centralized method outperformed the decentralized method by an average of 16.5% but the decentralized approach demonstrated strong local adaptability to real-time truck arrivals.

For the YCS-II problem, Ng and Mak (2005a, 2005b) addressed the problem of scheduling a single YC carrying out a given set of jobs with different ready times in order to minimize the total job waiting times. To achieve this, Ng and Mak (2005a) first developed a heuristic and a lower bounding algorithm to benchmark the performance of the heuristic. A branch and bound algorithm was then proposed in Ng and Mak (2005b). In addition, Ng (2005) presented an integer program, a dynamic programming-based heuristic, and a lower bound to solve the problem of scheduling multiple YCs. While the inter-crane interference was considered; there was no consideration of the safety distance requirement and no differentiation of the storage and retrieval jobs.

The single block twin-crane scheduling problem with both storage and retrieval jobs was also investigated by Vis and Carlo (2010). They constructed a mathematical model to minimize the makespan for both cranes. An algorithm to derive a lower bound for the makespan was then introduced and a simulated-annealing based heuristic was proposed to solve the problem. Stahlbock and Voß (2010) investigated the effectiveness of double-rail-mounted gantry cranes via simulation. They evaluated different online algorithms for the sequencing and scheduling of jobs for the
automated double-rail-mounted gantry cranes at a storage block for instances derived from actual terminal operations.

Park et al. (2010) utilized the heuristic-based and local-search-based real-time scheduling methods for twin rail-mounted gantry cranes working in a block at an automated container terminal. They rolled planning horizon ahead whenever a crane finishes one job and therefore claimed the approaches are suitable for real-time scheduling. Specifically, they singled out the containers that require re-handling and treated them as independent jobs, which greatly facilitated the cooperation of the quay cranes and addressed the workload balancing and crane interference problems.

Recent container handing technologies, such as the double-rail-mounted gantry (DRMG) crane system, have been used in practice to improve the handling efficiency. The DRMG system consists of two rail-mounted gantry cranes of different height and width which allow them to cross each other. Cao et al. (2008) focused on providing an operational strategy for the system to load outbound containers. They proposed an integer programming model and then applied a greedy heuristic, a simulated annealing algorithm and a combined scheduling heuristic to tackle the problem. Dorndorf and Schneider (2010) scheduled triple cross-over stacking cranes in an automated container storage block with asynchronous handover at the transfer areas at both block front ends. They treated the problem as an online optimization problem and constructed a new crane schedule for a certain planning horizon whenever a new job arrives or a job is completed.

Gharehgozli et al. (2014) modelled the one-YC scheduling problem with both the storage and retrieval jobs in a single container block into a continuous time integer programming problem. A two-phase solution method, which utilized the intrinsic properties of the model, was adopted to optimally solve the problem. The first phase included a merging algorithm which tries to patch sub-tours into a complete crane tour taking consideration of extra travel time, while the second phase used a branch-and-bound algorithm to reach optimality.

We developed discrete and continuous time models for YCS-II in Chapter 3 and Chapter 4. Realistic operational constraints such as non-passing YCs, the safety distance between two YCs, and simultaneous storage/retrieval moves have been considered. High quality solutions are obtained for actual problems with very short solution times.
2.3.4 Integration of decision problems

The three level model introduced by Böse (2011) for CT planning are hierarchically interconnected: the lower level provides support and restrictions for the higher levels. The higher level generates requirements to the lower levels. This is also the case for the three category decision problems introduced by Günther et al. (2005). Even at the same level, the decision problems are also closely related. The solution to the BAP can serve as the input to the QCA, the solution to the QCA can in turn serve as an input to the QCS. However, the solution from the QCA and QCS will affect the BAP: the allocation of the QCs to the vessels and the detailed schedule of the QCs will affect the vessel handling time of the vessels, which is often assumed to be constant in BAP. The productivity of the QCs cannot be high if the efficiency of the PMs or YCs are low. A QC or YC will have to wait if the PMs arrive late. Similarly, the PMs have to wait if the QCs or YCs are busy. Thus, the operations of the quay, yard, and the transportation areas have to be carefully coordinated. The flow of containers should be as smooth as possible. Only through a systematic integration of all areas, a CT is able to achieve a high performance and fulfil its overall function as a key node in international transportation networks.

This interdependency nature calls for the integration of decision problems: at the hierarchical level (i.e., the decision problems of design, planning, and real time control be integrated), and at the area level (i.e., the decision problems of quay, yard, and transportation areas be integrated). However, due to the extreme complexity of the decision problems at the different levels and areas, the model size often becomes intractable. There are not many research work considering the integration of such decision problems. Most research only consider the optimization of an individual problem in one area as listed in Sections 2.3.1-2.3.3.

Recently, more researchers realized the importance of integrating the decision problems. Meersmans (2002) considered the integrated scheduling of automated stacker cranes (ASCs), AGVs and QCs at a container terminal. Some algorithms including branch and bound and beam search heuristics, are tested. Bish (2003) considered the combined problems of QCs loading and unloading sequence, YT dispatching, and container storage locations at yards. The objective is to minimize the maximum time taken to serve a given set of vessels. A heuristic was proposed to solve the model. Chen (2007) presented integrated model to schedule various types of equipment at the same time. Both the loading and unloading operations are considered simultaneously. The problem is formulated as a job shop scheduling problem and solved with a tabu search algorithm. Lee et al. (2008) considered the integrated scheduling of cranes and YTs. The problem is formulated as an MIP model and solved with GA to minimize the makespan
of the cranes’ tasks. Lau and Zhao (2008) addressed the integrated scheduling of QCs, AGVs and automated yard cranes (AYCs). They proposed a heuristic called GA plus maximum matching (GAPM) to obtain near-optimal solution. Cao et al. (2010) developed an integrated model for the yard truck and yard crane scheduling problems for loading operations. Solution methods based on Benders decomposition are proposed to solve the model. Yuan et al. (2011) proposed the integrated straddle carrier path planning and task allocation in yard operations including the QC and YT related jobs. A job grouping strategy was proposed to solve the problem.

Chen, Lee, and Cao (2012) studied BAP, QCA, and QCS in an integrated model and applied combinatorial Benders cuts algorithm to solve the model. Homayouni et al. (2013) formulated a model to optimize the coordinated scheduling of cranes and vehicles in CTs. The objectives are to minimize the total travel time of the trucks and delays in the tasks of the cranes.

Meisel and Bierwirth (2013) proposed a framework to integrate the BAP, QCA, and QCS models. They used three phases to solve the three subproblems. Problem scenarios of realistic size can be solved in their integrated framework.

Chen et al. (2013) considered the simultaneous scheduling of QCs, YT, and YC so that the makespan for serving a set of vessels is minimized. The problem is formulated as a constraint programming model. A three-stage algorithm is developed to combine the different subproblems. Tierney et al. (2014) developed an integer programming model for analyzing inter-terminal transportation (ITT) in sea ports. The movement of the containers between terminals (sea, rail or otherwise) within a port was analyzed and modeled. A two-step solution procedure was proposed to solve the overall ITT problem. Homayouni et al. (2014) used GA to solve the integrated scheduling of the QCs, AGVs, and YCs. They showed that GA outperforms the simulated annealing algorithm for the problem.

2.4 Acknowledgments

The authors would like to acknowledge the valuable suggestions and data provided by PSA Corporation Ltd., which have helped to improve this chapter.

References


