MODELING AND OPTIMIZATON

FOR AN AIR CARGO TERMINAL

HUANG PENG

(B. Eng., Tsinghua University)

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF ENGINEERING

DEPARTMENT OF INDUSTRIAL AND SYSTEMS ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2005

Acknowledgements

I would like to express my deep and sincere gratitude to my supervisor, Prof. Huang Huei Chuen. Her wide knowledge and her logical way of thinking have been of great value for me. Her invaluable advice, guidance and patience throughout my study and research have made the completion of this work possible.

I am deeply grateful to my supervisor, Dr. Lee Loo Hay, for his detailed and constructive comments, as well as his understanding, encouraging and personal guidance. I wish to thank him for his stimulating suggestions and kind encouragements which have helped me during all the time of research for and writing of this thesis.

I owe my most sincere gratitude to Prof. Ellis Johnson, Prof. Chew Ek Peng, Dr. Lee Chul Ung, Dr. Wikrom Jaruphongsa and other professors in my departments, who gave me the opportunity to work with them in the air cargo research group. Their valuable advice and friendly support have been very helpful for my research.

My warmest thanks also go to the colleagues and fellows in my department. They are Dr. Bao Jie, Dr. Cheong Wee Tat, Chen Gang, Dr. Dai Yuanshun, Gao Wei, Dr. Paul Goldsman, Jiang Feng, Lai Xin, Leong Chun How, Liang Zhe, Liu Bin, Dr. Ivy Mok, Dr. Alec Morton, Dr. Sun Gang, Dr. Tang Yong, Wang Xiaoyang, Wang Wei, Dr. Yang Guiyu, Xu Zhiyong, Zhang Caiwen, Zeng Yifeng, to name a few (alphabetically).

Finally, my gratitude is due to my family for their encouragement, help, support, and understanding throughout my study and research.

Table of Contents

Summary

This research work studies the modeling and optimization for an air cargo inbound terminal. Operations in the terminal include cargo receiving, checking-packing, order-picking, and shipping. There are many factors that affect the operation performances in the terminal. The factors investigated in this thesis are the cargo flow time, workload balancing, and congestion effects. To address these factors, a cargo assignment plan is studied in detail.

Because of the various factors of consideration for this problem, it is neither possible to be formulated as a single objective problem, nor practicable to be modeled as a linear programming or integer programming problem, given the existing modeling techniques. Therefore a multi-objective mixed-integer programming model is formulated to improve the assignment plan. It aims to provide a series of non-dominated solutions.

These solutions are then input to a simulation framework which will identify the best solution(s) to the preference of the decision maker. This simulation is able to model the cargo handling operations. It not only evaluates the effects of cargo assignment on the overall performance, but also examines the congestion effects due to imbalanced assignment and system randomness. The performances of these solutions in simulation are collected and compared for decision making.

Such a research approach including MIP formulation and simulation modeling is applied to an inbound air cargo terminal. Extensive computational experiments are

conducted with actual data as the input sources. This approach is demonstrated to be capable to support the decision makings for the terminal.

List of Tables

List of Figures

1 Introduction

Since the momentous globalization of world trading and economy, the airline industry has been playing a pivotal role in the integration of world markets. Along with the growing demands of international trading and exchange, global air transportation is experiencing an excellent opportunity to boom again after the 911 incident and the global economic recession in 2001. With the paces of globalization and regionalization, the world is marching towards a new phase of peaceful development. The recent trend in financial integration and energy market liberalization further stimulate the up-stream supply for the airline industry. Evidence suggested that the airline industry is soaring again despite the recent events like epidemics and turbulence in the Gulf. A robust global supply chain network is shaping itself to accommodate the start of another economic growth cycle. Airline industry is therefore becoming more and more crucial in the global supply chain.

Air cargo terminal connects different modes of shipment together, and therefore serves as a significant and indispensable link in the global commerce chain. Recent advances in information technology and computer hardware pave the way for possible improvement on the air cargo terminal’s strategic and tactical performance.

This research is motivated by a study at an air cargo terminal which handles the inbound and transshipment cargos for a top-tier international airline at its hub airport. We observe that cargos shipped by the airline arrive at the terminal in the form of a pallet or a Unit Load Device (ULD) which often consists of a few consignments belonging to different cargo agents. (Generally, cargo agent is used by the cargo terminal to address all the shippers, freight forwarders, and consignees that have consignments handled by the cargo terminal.) In addition, the concerned airport may not be the intended destination for some of the consignments.

This chapter aims to explain the related backgrounds about the research project. It thus starts with a brief introduction about the background of the research, followed by the detailed description on the function and layout of an inbound cargo terminal. Subsequently, the cargo inbound handling process and its related assignment planning approach are introduced to give some lights on the origination of the research problems. After the descriptions about the research motives, the contribution of this research work is briefed. Finally, the structure of this entire thesis is outlined in details.

1.1 Background

This section provides an overview of the research problem. It gives an overall understanding about where the problem comes from, how the problem is related to our research, and how we elaborate it in the future. The general description in the thesis is based on our observations at a leading international airport.

An air cargo terminal is essentially a fast-moving warehouse. The inbound terminal needs to do breakbulking in order to facilitate cargo agents’ collections and to transfer the cargos to the outbound terminal for further processes to be ready for the connecting flights. In an inbound terminal, the cargos are moved through various facilities, and finally reach the outbound terminal or shipment dock. The cargo travels within the terminal via different types of facilities and transferring equipments. The details about the cargo terminal will be introduced in Section 1.2.

Due to the varied cargo characteristics, the cargo movement in the terminal exhibits different patterns. The cargo airplanes touch grounds at the airfield within the airport. As we can see in Figure 1.1, after the cargos are unloaded from the airplane and towed to the ramp side of terminal, the cargos start their movement within the terminal.

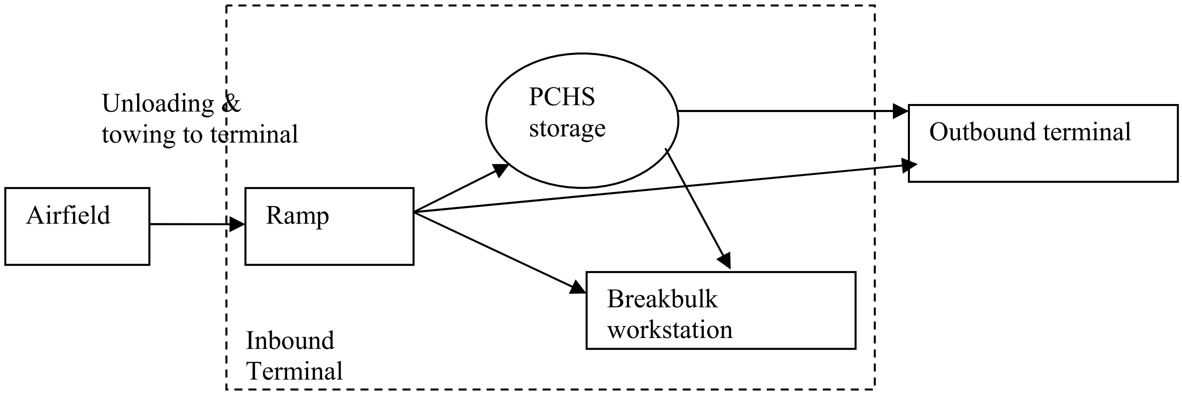

Figure 1.1 Cargo flow process in this inbound terminal

It is obvious to see from Figure 1.1 that there are two directions for the cargos to travel within the terminal. One of the directions is to transfer to the outbound terminal immediately after they arrive at the inbound terminal or through the intermediate storage (PCHS storage, more details in Section 1.2) to the outbound terminal. This direction is for transshipment cargos which need to be sent to the connecting flights. The other cargo movement is to transfer them to the breakbulk workstation where they are broken loose at the breakbulk workstation (more details in Section 1.2) in anticipation of the collection from cargo agents. The process of the inbound cargo handling will be further introduced in Section 1.3.

In order to handle huge volume of cargos, the terminal is equipped with multiple facilities and equipments. These facilities include the ramps, storage places, and breakbulk workstations. The ramps are divided into several ramp zones for the ease of management. They are the places to receive the inbound cargo. The storage places in the PCHS (Pallet Container Holding System) and the breakbulk workstations are also grouped into clusters. The breakbulk workstations are where the breakbulk job is taking place. The equipments within the PCHS are the hoists and the transferring vehicles which assist the cargo movements. It is observed that the transferring time between different facilities vary and some equipments are shared between groups.

Since there are multiple ramp zones and breakbulk workstation areas in the terminal, the present work practice for this international airline is to designate the suitable ramp zone and the workstation area for each flight according to their flight number. Therefore a fixed assignment plan which dictates the ramp zones, workstation areas, and the storage places belonging to a particular flight is adopted. A more comprehensive introduction about the cargo assignment planning is given in the upcoming Section 1.4.

Such a fixed assignment plan would make it easy for the management of cargo dispatching. In addition, since the transferring time between different facilities varies, it helps to choose the shorter traveling path to take advantage of this difference. Furthermore, a fixed assignment makes it possible to estimate the workload condition for each facility, since the flights allocated to each facility are already known beforehand. It is therefore obvious that the efficiency of the terminal operations depends much on the quality of this assignment.

Our research is to measure and identify a good assignment for the terminal operation so as to improve its efficiency. In the following sections, more detailed introductions about the terminal operations, function, layout, cargo handling process, and cargo dispatching planning are described to elicit our research motivations and its performance measures. After the necessary background information, the contributions of this work and the structure of the thesis are discussed.

1.2 Introduction on air cargo terminal

The purpose of this section is to give some basic description about the function, components, and layout of an air cargo inbound terminal.

The basic layout of the inbound cargo terminal can be illustrated by the graph below in Figure 1.2. It primarily consists of ramp zone facilities, PCHS system, and breakbulk workstation areas.

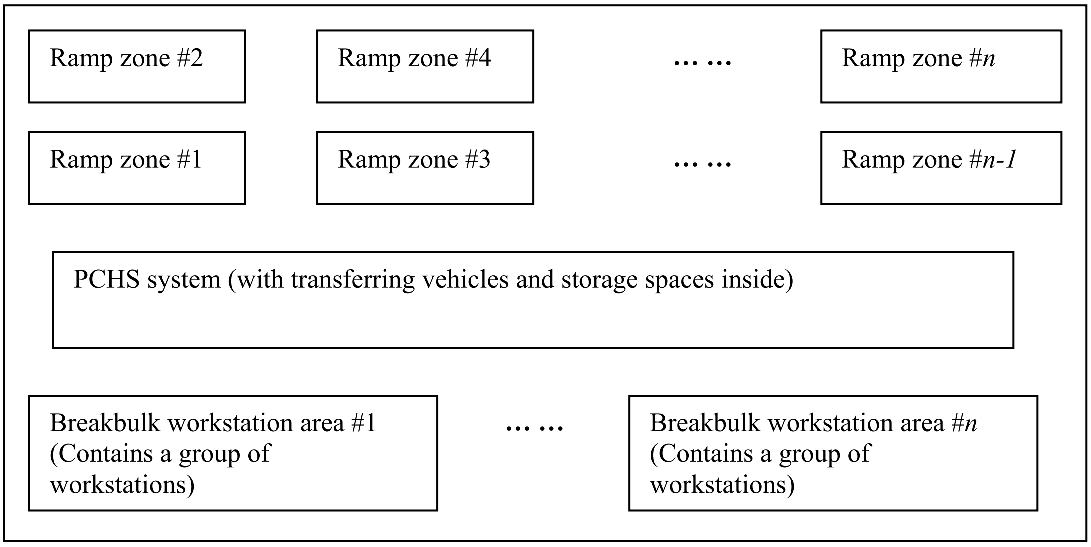

Figure 1.2 A simple illustration of the basic layout for a terminal

The air cargo in movement is packaged in a unit load device (ULD). It is important to firstly explain the basic layout of the terminal and the structure of the cargo handling system in a detail manner to build an understanding of the air cargo inbound terminal.

The *cargo terminal* is essentially a multi-level warehouse building. The inbound and outbound functions of the terminal are differentiated and there are dedicated sub-terminals to serve either the inbound or outbound function.

A *PCHS* is the same concept as a MHS (Material Handling System), which can hold the cargo for short time on-purpose storage.

A *ramp zone* is the receiving dock of the inbound terminal for the cargo unloaded from the airplane. Once a ULD is towed from the airside of the airport to the terminal, the ULD will be introduced into the ramp zone. There are multiple ramp zones located at different areas of the terminal. Ramp zones are located at the ground level of the terminal building. The cargo is placed onto the conveyor queue lane of the ramp zone after they arrive at the terminal. From the queue lane, the cargo is thereby transferred into the PCHS by the transferring vehicles.

The *transferring vehicle* is also referred as the *ETV*. ETVs are electrically driven equipments within the PCHS which are controlled by the computerized control system. These vehicles move along the vehicle channels within the PCHS. It is comparable to the AGV (Automated Guided Vehicle) of an ASRS (Automated Storage / Retrieval System). The transferring vehicles serve various purposes such as moving the cargo between the queue lane and PCHS, transporting cargo between different positions within the PCHS, and transferring cargo between the PCHS storage positions and exit positions of the PCHS.

The ultimate purpose of the inbound cargo terminal is to move the cargo to the outbound terminal or to the breakbulk workstation. The cargo goes to the outbound terminal may be checked and palletized again for another flight in the outbound terminal. The breakbulk workstation performs the breakbulk job for the palletized cargo.

PCHS highway serves as the direct linkage between the inbound terminal and the outbound terminal. The inbound cargo with transshipment purpose and without breakbulk requirement will be moved directly via this direct link to the outbound terminal. This PCHS highway locates horizontally in the space above the ramp zones.

Hoist serves as the linkage between different levels of the terminal. A hoist is an electricity-driven lift for the purpose of moving cargoes vertically between different levels. It has fixed capacity so that it could carry fixed amount of ULDs each time.

Breakbulk workstations locate outside the PCHS and near the exit dock of the terminal warehouse building. These workstations are grouped into several areas to ease the management and resource dispatching. These areas are called the *breakbulk workstation areas*. At each workstation, the checking team performs the breakbulk job according to an eight-hours-per-shift schedule. The palletized cargos are broken loose and rearranged, and then moved by forklift to the outbound terminal or the receiving dock for the cargo agents' collection.

1.3 Introduction on the cargo inbound process of an air cargo terminal

In this section, we address the cargo inbound handling process in a thorough way. The cargo inbound process is the subject of our study, and the purpose of this study is to improve the process via our modeling and simulation approach.

An illustration of the process is given in Figure 1.3 for an incoming flight from the time it arrives at the airport to the time it leaves the breakbulk workstation in the inbound terminal. Obviously, this chart doesn't consider the case of direct transshipment of which the cargo moves from PCHS to the outbound terminal directly. Since the ULDs of a flight arrive on a unit-by-unit basis, it is possible that when the first ULD is being processed at the next process, the last ULD could be still at the initial process. Therefore, there is some overlapping between the time frames of two adjacent processes in the chart.

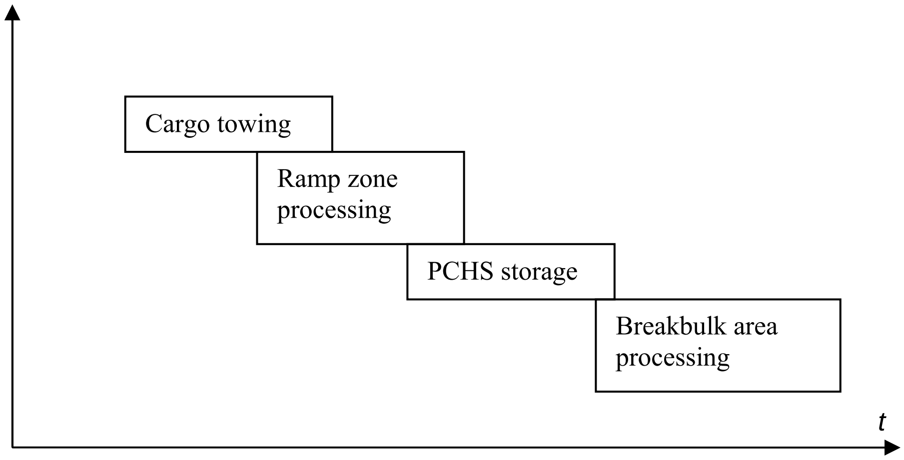

Figure 1.3 Cargo movement process for inbound operations

Due to the fact that there are multiple ramp zones and breakbulk workstations, it is necessary to decide the allocation of these facility resources to the cargo beforehand. The cargo dispatching procedure follows the planned assignment to assign the cargos from different flights to different facilities. Thus, this cargo dispatch plan is a tactical planning problem of assigning flights to ramp zones and to breakbulk workstation areas.

As mentioned before, there are various cargo flow patterns within the cargo terminal. Hence, the different sequences of cargo flow need to be introduced in further detail. The general cargo flow process can be broken down according to its associated origin-destination. The following flow chart mainly describes the cargo flow process of inbound cargo operations.

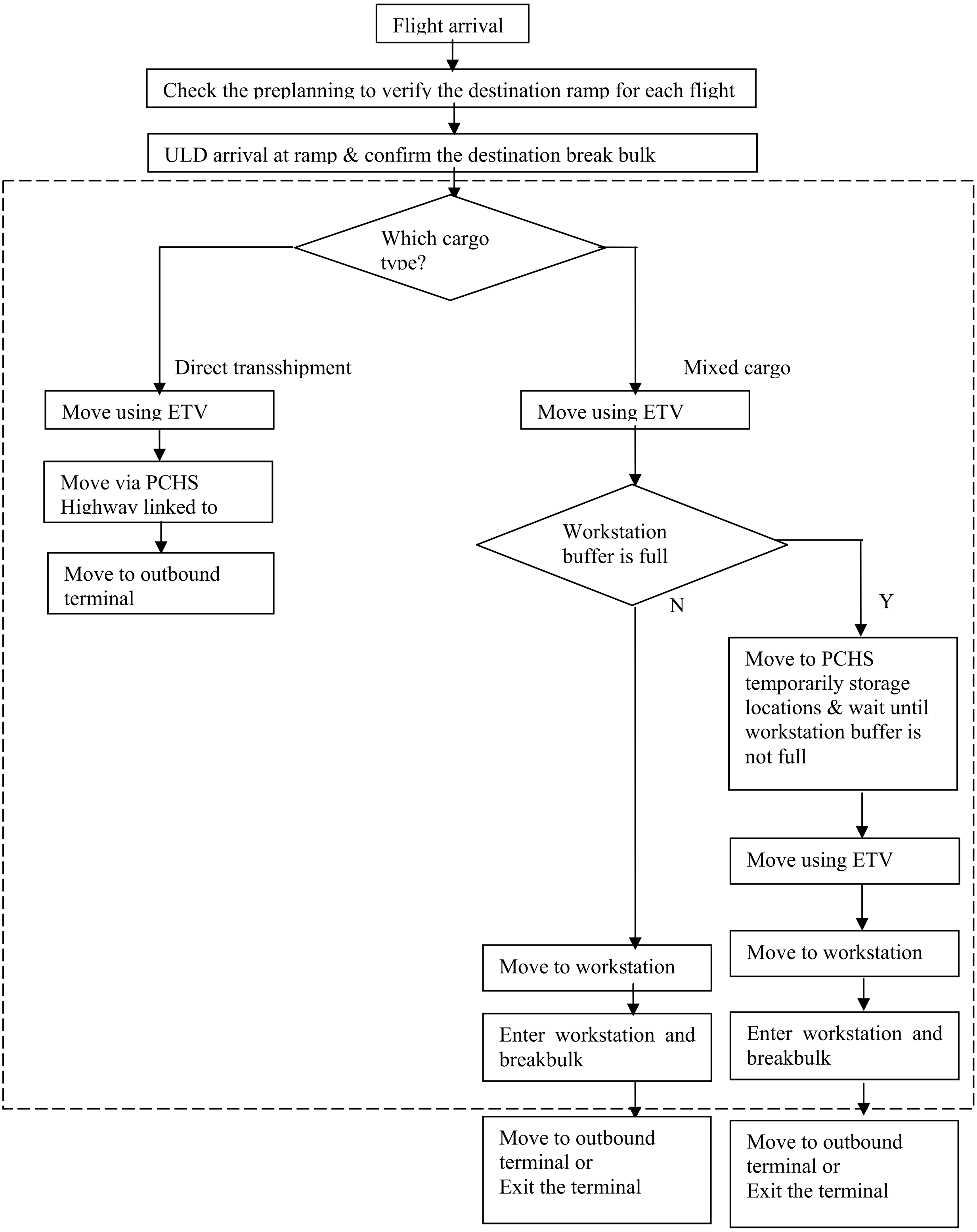

Figure 1.4 General flow of a ULD

The cargo flow process shown in the dashed box is solely decided by its own origin and destination positions. All the movements included in this dashed box are the sub flow process of the ULD movement within the PCHS.

The cargo flow process in the dashed box involves the choice of different paths within the PCHS based on the cargo characteristics. As mentioned in Section 1.1, different cargo characteristics, such as mixed shipment, or direct transshipment can determine how the cargo moves within the PCHS system. This can be seen from the decision making on whether to use the PCHS highway to move the direct transshipment cargo. If the cargo is for direct transshipment purpose, it will be lifted onto PCHS highway through which the cargo will reach the outbound terminal. Otherwise, the cargo will be moved into PCHS for breakbulk purpose.

The cargo flow process of the mixed cargo is more complicated. The mixed cargo includes both the imported cargo which needs breakbulk before exports and the cargo which requires import transactions only. The choice of whether to use the temporary storage space is also of our interest here. As a generally accepted practice, it is more preferable for the cargo to travel through the shorter and less congested path if this proposed path is free to use. Otherwise, if the shorter path were not available due to congestion or malfunction, the cargo would stay in the storage temporarily. When the path becomes available once more, the cargo movement will start again. The mixed cargo will be eventually moved to the outbound terminal or breakbulk workstations.

1.4 Introduction on the tactical planning of an air cargo inbound terminal

It is mentioned in Section 1.3 that the cargo dispatch planning is a problem of allocating cargos to different facilities during different time frames. In this section, we try to explore the outcomes of such a planning, as well as the relationships between the tactical planning of the cargo dispatching and the operation efficiency and service quality of the terminal.

It is discovered that the preplanning of the assignment of flights to different ramp zones and then to different workstation areas could affect the ULD's flow pattern and in turn, influences the facility utilization and overall throughput time. The capacity utilization becomes the main concern for the improvement of handling efficiency. Plus, as a service provider, the terminal serves as the linking node between the carriers and the cargo agents. It is therefore crucial to improve the quality of service by reducing the overall throughput time within the terminal. Thus, the operation efficiency and service quality of the cargo terminal are heavily dependent on the preplanning assignment.

We observe currently, in this inbound terminal, some of the ULDs need to travel relatively longer distance to reach their assigned areas. Many transshipment cargos are assigned to travel through an unreasonable longer way to the outbound terminal instead of some shorter path. In addition, the handling volumes of the workload at different workstation areas are imbalanced. One of the possible reasons for these observations is the imbalanced assignment of flights to the ramp zones and the workstation areas.

It appears that the current system does not operate at the ideal level due to its assignment planning. Such similar problems also exist elsewhere in air freight or sea freight terminals. The current inefficiencies of the terminal operations are mainly as follows:

1. Imbalanced workload and congestion

The data we collected show that, different segments of the ramp zones handle different workloads resulting some of the equipments highly utilized while others under utilized. In other words, during certain hours, the highly utilized ramp zone would suffer from ULD congestion on the ramp queue lane, waiting for the ETV. The imbalanced workload creates the problem of congestion, and the congestion at the ramp queue lane causes the longer time to handle these congested cargos.

In the current practice, the cargos are staged at the PCHS locations which are close to the workstation for a short duration of no more than 60 minutes before transferring to breakbulk workstation. The reason is that the cargos have to wait for the workstation to be free, i.e., when the workstation reaches its capacity limit, the cargos need to be stored at the PCHS temporarily.

2. Inefficient ULD flow

Since the mixed cargos need to be directed to the breakbulk workstations for the next stage operation, they will make use of the transferring equipments within the PCHS to reach the designated workstations.

It is observed that some of the not-so-good ULD flows within the PCHS use more equipment and take longer time unnecessarily because of their long traveling paths, even though there are multiple available paths with no congestions to go. Such detours cause the cargo flow in the terminal suffering from long movement time.

1.5 Problem description

After some piloting collection and analysis of the data, the findings based on the analysis suggest that the current fixed assignment of flights to different ramp zones and to different workstation areas is not efficient. Therefore, it would be necessary to revise the preplanning assignment in order to improve the cargo flow pattern. It appears that the current system does not operate at the maximum efficiency level due to the current assignment planning.

1.5.1 Motive of the research project

For a busy facility with certain peak periods such as an air cargo terminal under the dynamic condition, congestion reduction and diffusion are equally important. The ULD might encounter longer flow time by traveling through a congested shorter path than through the longer but less congested paths. Hence, this research project is expected to bring shorter flow time and less congestion to this cargo terminal.

The aim of the project is to optimally allocate the incoming cargos from different flights to the ramp zones and to the workstation areas so as to improve the material flow in the cargo terminal. It plans to reduce the movement time for cargo to move from ramp to PCHS locations, export terminal, and breakbulk workstations, along with the consideration that the congestion effects can be lessen. In other words, the model aims to find a flight-to-ramp-to-workstation assignment, to reduce the flow time and to streamline the cargo flow.

1.5.2 Performance measures

The objective of this problem is to improve the quality of cargo assignment to reduce the congestion occurred due to imbalanced workload while not compromising on the flow time. The flow time reduction problem is the most straightforward issue for the cargo terminal, while the congestion problem and the imbalanced workload problem arose from our observations and insights derived from the initial stage data analysis.

There are three performance measurements proposed for this study:

1. The average flow time of the ULD is one of the most straightforward measures for evaluating the system efficiency. The flow time in terms of an ULD normally is made up of movement time and processing / queuing time. i.e., flow time = movement time + processing / queuing time. The long storage time could be the result of other reasons

such as the delay in paperwork or late notification for the customers, rather than congestion. Here the term flow time will not consider the long intermediate storage time resulted from the above reasons.

2. To avoid the imbalanced assignment of the workload at the ramp zones and the workstation areas, we use the maximal pair-wise difference of the *capacity ratio* as the second measure. We approximate the degree of workload congestion by the *capacity ratio*. The capacity ratio is the ratio of the current assigned workload at the equipment to the nominal processing capacity of the equipment during a fixed time unit. The capacity ratio (*CR* for short) should be less than 1 in reality, but may be greater than 1 if it is calculated in a relaxed manner by including works waiting in queue. Therefore the capacity ratio is a measure of the utilization of each facility.

3. In order to measure the seriousness of the congestion, we suggest the exceeding value of the capacity ratio over 1 as the third measure. If taking into account the possible over-utilization at some facilities, the capacity ratios at these facilities would be greater than 1 during some intervals. To avoid such risk of over-utilization, it is therefore valuable to reduce the overall exceeding values of capacity ratios.

1.6 Research contributions

This thesis tackles the operations enhancement plan in an air cargo inbound terminal. The research work suggests a novel and comprehensive approach to address the flight-to-ramp-to-workstation assignment problem. The main contributions are:

1. It provides an analytical modeling approach to formulate a mixed-integer programming for an air cargo terminal. Such an MIP model is not only able to evenly allocate the cargo workload to the equipments, but also it could improve the overall movement efficiency.

2. It proposes an applicable hybrid framework which entails both optimization and simulation techniques for air cargo terminals. The optimization process provides the non-dominated solutions, while the simulation process further tests these solutions and helps to make the decision. It has an edge over other conventional singular approach to pinpoint the best decision.

3. Such an approach extends the planning problem from daily operations to the weekly tactical plan for an air cargo terminal. Therefore it provides assistance in the mid-term / long term decision making for the business process reengineering of inbound air cargo terminals.

1.7 Organization of the thesis

The rest of the thesis is structured into 5 chapters to present the study in a more specific and detailed scale. The upcoming Chapter 2 provides a review of the available literature. The current literature provides an overview of the related research works about the operations management aspects of a cargo terminal. The reviewed articles range from managing the operations in a container terminal, strategic design issues such as the layout, shape, number of facilities of a warehouse, to the workload balancing issues.

Chapter 3 approximates the situations described in Chapter 1 by a mixed-integer programming which addresses the flight to ramp zone and to workstation area assignment problem. This mathematical formulation aims at reducing the overall flow time related cost, as well as balancing the workload among the facilities. Such an MIP formulation with multiple objectives is expected to give a series of efficient solutions with proper “quality” of the assignment to streamline the cargo handling process. The coefficients estimate and necessary assumptions are also stated in the chapter.

In Chapter 4, we suggest a simulation model for the cargo movement process, in order to capture some important random features which are not considered in the MIP model. Such a simulation model is described in Section 4.1, including the objectives of the model, the required inputs, the model layout design and the logic design, etc. The simulation model is further verified, pilot run, and validated in Section 4.2, 4.3, 4.4. Such a simulation model serves as the test ground for the efficient solutions to identify the most favorable assignment from them. The performance measure for the simulation is the overall flow time for all cargos since it is considered as the most important requirement for a cargo terminal.

The computational results for the MIP multiple objective optimization and the simulation outputs are presented in Chapter 5. An ϵ-constraint approach to find out suitable efficient solutions is proposed in Section 5.1.1, with the solution results given in Section 5.2 and 5.3. The simulation running results along with data statistics are also given in Section 5.2 and 5.3 followed by proper explanations. Accordingly, the suitable assignment is identified and suggestions are made.

Finally, Chapter 6 concludes this research and suggests some future research directions.

2 Literature Review

The problems associated with locating items or facilities, assigning works or products, and scheduling production or fleet arise frequently in modern logistics systems. These problems have been extensively studied in management sciences / operations research / operations management, in various contexts of production planning and scheduling, container terminal management, fleet management, or warehouse management.

However, most of the articles are not directly related to air cargo terminals which show a resemble manner to cross-docking. The research in air cargo terminal, thus, requires similar techniques to those of other contexts but in an entirely different setting, to take into account the characteristics in the air cargo terminal.

In this regard, the literature review is organized into three topics, namely container terminal operations, air cargo planning and operations, and load balancing.

2.1 Container terminal operations

Our problem involves the improvement of the operations of an air cargo terminal by providing a new tactical design for the system. Such a problem is considered similar to the tactical design for a container terminal. The improvements to make in our study are motivated from various perspectives, such as to increase the throughput, or to decrease the turnaround time or cycle time, like in Preston and Kozan (2001) and Taniguchi et al. (1999), or, to allocate the space to effectively allocate the operations to reduce the traveling cost or delivery cost, like in Kim and Park (2003). The objective functions for these problems became the overall turnaround time, or the traveling cost, and constraints came from different types of resources. Those problems were basically formulated as a mixed-integer linear programming with certain degree of simplifications. Then some problem specific solution methods were adopted to resolve these mathematical programming.

The problem under our study involves with a fixed flight schedule which is repeated weekly. Nozick and Morlok (1997) presented a model for the planning of operations of an inter-modal truck-rail service. This model strictly followed a fixed schedule, which is similar to our problem, since our problem also deals with the fixed weekly schedule of all flights. The service operation in Nozick and Morlok (1997) was comprised of moving trucks and containers on rail cars between terminals, with transportation by truck at each end. It aimed at redesign such systems to produce more reliable services, with multiple service classes, and better equipment and facility utilization. An integer linear programming model was developed with the objective to minimize the overall cost covering all elements of the operations, which is also the ultimate goal of our problem if more study is given in the future. This model was directed toward the intermediate horizon planning, that is, the planning for a period of one week or a month or so. Its constraints included the different service levels, flow conservation equations, fleet size constraints, and terminal physical capacity constraints. The inputs for this model were the forecast of cargo amount, equipment specifications, the vehicles information, and terminal capacities.

Our problem in cargo terminal requires the proper assignment of cargo contents to processing facilities. Such type of problems often occurred in container terminals, too.

Bish (2003) considered a container terminal where the regular operations are the loading and unloading of containers to and from a set of ships, and storing the containers in the terminal yard. Each ship was served by multiple quay cranes, which were used to load and unload containers to and from ships. The transportation equipments for the containers were a fleet of vehicles, each with unit capacity. The problem was to assign a storage location for each unit container, as well to dispatch the vehicles to the containers, and to schedule the loading and unloading operations, in order to minimize the maximum time to serve a given set of ships. This was an NP-hard problem, and therefore a heuristic algorithm was developed. The above study can be considered as a comprehensive example of the research in container terminal operations. Its modeling approach also gave implications for our problem modeling on one of the performance measures, namely the overall flow time of the cargoes.

Another paper has also provided sufficient insights for our research problem. Vis and Koster (2003) finished a complete overview about the container transshipment problem. In the article, the "docking time" of the transshipment of containers at a container terminal was presented as the major factor for evaluation, which is the same as our proposed objective – the overall flow time. It provided a classification of several decision stage problems at container terminals. It examined individual types of material handling processes as well as the overall planning problem for a container terminal.

Simulation technique also plays a vital role in the operational planning for terminal operations. With the help of simulation, a more clear and straightforward image of the system under study could be suitably presented to the management as in Gambardella et al. (1998), Marco and Samli (2002), and Yun and Choi (1999). In Gambardella et al. (1998) a decision support systems for the operations management of an inter-modal container terminal was presented. It addressed the allocation of containers on the yard, the allocation of resources, and the scheduling of operations, in order to maximize the performance of the system. This problem was further solved with other techniques like genetic algorithm and mixed-integer linear programming. Furthermore, the simulation model of the terminal was developed with the purpose to present the results to management. This simulation focused on the efficient allocation of resources. Similar application of simulation tools can also be seen in Marco and Samli (2002) and Yun and Choi (1999). These research works contributed to prompt the thought of using simulation in this thesis for the pinpointing of the desirable assignment planning.

These above works contributed much to the origination of our problem modeling. They suggested the use of a mathematical model as well as a simulation model to address the performance enhancement of air cargo terminal operations. However, the lack of measurements for congestion effects was common in them, and which became another concern for our problem.

2.2 Freight terminal strategic planning

Some articles in this particular field looked at the strategic issues such as the layout, or shape design of a terminal, while others concentrated on the total number of vehicles, and equipments within the terminal, and the network design related to the terminal. Especially, the strategic issues about the freight terminals are becoming more and more significant particularly because of the expansive development of cross-dock like terminals.

Most of the strategic planning problem looked at the reduction of traveling cost and handling cost within the system. Layout design of a terminal, such as the terminal door placement, or shape of a dock terminal, was extensively studied in the literature. Tsui and Chang (1990) proposed a bilinear programming model and a straightforward solution method for a local optimal solution to a freight terminal. Based upon this research work, later Tsui and Chang (1992) used another heuristics approach to solve the same problem and improve the solution time up to 70%. Although these works provided significant improvements over previous planning, their models only considered one-stage assignment which assigned jobs to outgoing docks.

A problem-specific study about reducing the material flow cost for a long term planning problem was introduced in Gue (1999). It suggested a two-step approach for the incoming trailer scheduling based on the terminal layout. The first was to determine the optimal assignment of trailers to dock door based on the “look-ahead” schedule for a given layout; the second step was to search the solution space of all possible layouts with the lowest cost. This first problem was formulated as a linear programming with the decision variables representing the material flow from incoming doors to outgoing doors. The objective was to reduce the cost of assigning incoming trailers to doors. In the second step, the local search algorithm continued to swap searches for a better layout until there was no further improvement to make. The first step problem to assign the incoming trailers is somewhat similar to our assignment problem in which work contents are assigned to facilities.

Some other work investigated the components of the total cost in a cargo terminal. Bartholdi et al. (2000) described a set of models that guided a local search routine to generate a layout, in which the total cost would be minimized. The balancing of the traveling distance and congestion was also addressed in this model. The total labor cost was broken down into two parts, one is the worker traveling time, and the other one is the worker waiting time due to congestion. Their work further investigated the possible causes of congestion, with the help of queuing theory. Simulated annealing procedure was adopted to refine the best plan of total cost based on an initial layout. Therefore, in our study, the same attentions similarly are paid to both the traveling time and the waiting time.

2.3 Load balancing

With the advent of modern manufacturing technologies, the load balancing issues are frequently discussed in literature, while there is still little seen in the area of freight handling terminals. The concepts and implications from Toyota Production Systems, Kanban systems, and Just-In-Time could also serve as the basic methodologies for a modern freight handling terminal, in particular in an air cargo terminal in which the swiftness and efficiency are mostly concerned. The JIT philosophy also contributed much to the conceiving process of this specific study on air cargo terminal operations. Based on their similarity and resemblance, this load balancing approach could also be applied in the context of performance improvement in a logistics terminal.

Although the load balancing issues were rarely addressed, there exists a selection of articles on the workload balancing for a complex system especially in a manufacturing system. Such articles provoked the thoughts of balancing the workload among facilities during each time window. They shed lights on the development of the load balancing aspect of the mathematical model.

Load balancing issue often works with the planning together, as in our problem. Houghton and Portougal (1997) presented their study on the balancing of workload variations and WIP inventories. In their study, a production planning model was initiated, with the multiple objectives of minimizing the capacity requirement planning cost and inventory holding cost, along with capacity constraints, inventory supply and stock constraints. Under the planning settings, several steps of planning procedures as well as the trade-off analysis were carried out to search the optimal solution for the planning model. Dynamic programming approaches were employed here to search for the solutions to each trade-off in each step. Although we do not use dynamic programming approach for our problem, this study is worthy mentioning since it suggested the complexity of creating such a dynamic programming model.

A majority of articles in this field were focused on the integer programming model of machine loading. Both Berrada and Stecke (1986) and Wilson (1992) modeled their problems with an integer programming approach. In their models, the tools and operations were allocated to machines with limited capacities. An approximate integer model was developed for this problem. In Berrada and Stecke (1986) a branch and bound approach was used to solve this problem with extensive demonstrations on how to find the lower bound and the selection of the branching variables. As in Wilson

(1992), this model was further modified with revamped objectives and a heuristic algorithm was carried out to handle this problem. Also in Khouja and Conrad (1995), the authors tried to assign the customer groups to employees, with the consideration to minimize the deviation of the processing time of different groups, as well as to minimize the deviation from an employee point of view. The problem is formulated and further solved with both a heuristic approach and a zero-one goal programming approach. The final suggestion for this study was to use heuristic approach to provide a good initial point for the zero-one goal programming solution method. These research works motivated the thinking of using integer programming or mixed-integer programming model to represent the load balancing problem with the cargo terminal background.

Some even more complicated model was devised to address the precedence of job sequences in balancing issue. Sawik (2002) proposed an integrated formulation for both the scheduling and balancing of an assembly line system. This integrated formulation took into account the task precedence information, time limitations, as well as other essential information for this problem. It aimed at minimizing the overall completion time for the operations. In addition, in order to find the optimal decision for this problem, an integrated method and a decomposed method are both applied on it. This integrated method could resolve both scheduling and balancing simultaneously, while the hierarchical approach handled the problems sequentially. Thus they were compared in terms of computation time and efficiency. It was recommended that, for large size problems, the hierarchical approach was more suitable to produce reasonable results within certain time. As in our problem, since the precedence of jobs is solely determined by their arrival times, the task priority is not of our concern. Hence, our problem only addresses the assignment and balancing issues, while it could determine the scheduling of job sequences automatically after the proper assignment plan is achieved.

Apart from the applications in business and manufacturing systems, Amiouny et al. (1992) suggested a unique approach to balance the load stowed into an airplane. A heuristics method motivated from the “center-of-gravity theory” was used. This heuristic is shown to be able work well on this one-dimensional balanced loading problem given its structure. The knowledge from mechanic design was shown complementary to the traditional question of this type. And this “combinatorial mechanics” approach, according to the authors, was able to sufficiently tackle this class of problems. Although the background of this problem is similar to ours, such an unusual and unique approach is beyond our knowledge. Furthermore, this model only balances the load without considering the time issue. The only purpose of reviewing this article is to present an unconventional way to deal with a conventional problem.

3 Mathematical Formulation

In this chapter, a mixed-integer programming model is presented to resolve the flight to ramp zone and flight to breakbulk area assignment problem. This mixed-integer programming model is formulated with multiple objectives based on the performance measures suggested in Chapter 1. A deterministic estimation about the coefficients in the model is also discussed in this chapter. The computation experiments are later implemented using solution package ILOG® CPLEX 8.0, on a PC Pentium IV 2.60 GHz platform, with 512 MB build-in memory.

This chapter is organized into two major sections. It starts with the mixed-integer programming model, along with a detailed introduction on the underlying ideas. And then, it shows the coefficients estimate process for the model.

3.1 The mixed-integer programming model

The objective of this assignment problem (flight to ramp zone and flight to breakbulk area assignment problem) is to determine the specific ramp zone, and the specific breakbulk workstation area for any given flight. The purposes of this problem are multiple. The most crucial objective is to minimize the overall flow time for all cargos in movement. The other objectives include balancing the workload on each facility, and reducing cargo overloading at the facilities.

Since the flights handled by the terminal operate on a weekly repetitive basis, it is logical to set the time horizon to be one week. The facility has its natural processing capacity in terms of the number of cargos processed by it within a given time interval. The duration of this time interval is treated as an adjustable parameter. It could be adjusted smaller to capture the workload more precisely, or greater to make the model easier. Therefore, the duration of the time interval is a delicate choice for the mathematical formulation and its computational tractability. For our model, we set it depending on the length of its time horizon, for instances, we set it as 5 minutes for the one-day problem, and 1 hour for the one-week problem.

This section is organized as follows: first, the necessary assumptions and approximations are made to facilitate the problem formulation; then, the mathematical model is presented with an overall look; finally, the elaborations on each constraint and objective of the model are given to explain the underlying principles behind this problem formulation.

3.1.1 Assumptions

To approach the problem, some assumptions about the mathematical model need to be made. The purpose of making these assumptions is to facilitate the problem formulation with plausible relaxations on some of the stringent conditions.

The assumptions are:

1. In our formulation, the “towing and unloading time” between the arrival of flights at the airport and the arrival of cargos at the cargo terminal is ignored because we consider the problem from the perspective of tactical planning. In other words, the

starting arrival time of cargos at the cargo terminal is treated the same as the scheduled arrival time at the airport in the estimate for the workload coefficients.

2. We assume there is no interaction between the arrivals of cargos from different flights, i.e., the towing of cargos to the terminal ramp from one flight is independent of the towing of cargos from the other flights. Under this assumption, the cargo arrival behavior and the workload profile for each flight are easier to estimate.

3. Besides that, the arrival process of cargos at the ramp is assumed to be at a constant arrival rate.

4. Also the processing rate of cargos at the facilities is assumed to be constant.

5. We also assume that there are unlimited resources at the ramp zones and workstation areas, hence, given the assignment of the ramp zone and the breakbulk workstation area for each flight, the traveling path and the traveling time for each flight will be fixed. With this information, we can estimate the workload of a flight at a given facility during a given time period. The workload at this facility would be zero if no flight is currently being processed on this facility, or some number if some flights assigned and currently being processed during this time interval.

6. Another assumption is needed for the processing of cargos from freighter flights. As the freight flights carry a great amount of cargos, it is a common practice to allocate more workforce to perform the checking and breakbulk job. Based on our observation, two times of the regular size of workforce for a passenger flight are allocated for each

freighter flight, at the breakbulk workstation. Therefore, it is reasonable to assume that the processing rate for freighter flights is twice as that of passenger flights at breakbulk workstation area.

Note that the double workforce requirement for freighter flights occurs only at the breakbulk workstation, thus the processing of freighter flights at the ramp zone is still the same as that of passenger flights at the ramp zone. Hence, the processing rate for freighters flights at ramp zones are the same as that of passenger flights.

3.1.2 Model formulation

The purpose of this section is to provide a mixed-integer programming model for our problem. In this section, the objectives, constraints, and variables of the mathematical model are stated.

Set notations and indices

I^p	the set of all incoming passenger flights;
I^f	the set of all incoming freighter flights;
J	the set of ramp zones;
K	the set of breakbulk workstation areas;
T	the set of time intervals;
i^p	an incoming passenger flight, $i^p \in I^p$;
i^f	an incoming freighter flight, $i^f \in I^f$;
j	a ramp zone, $j \in J$;

k a break bulk workstation area, $k \in K$;

t a time interval (time unit) in one week, $t \in T$;

Variables

$x_{i^p jk}$ = 1, if the passenger flight i^p is processed at ramp zone j , and then goes to workstation area k for break bulk; 0 otherwise;

$y_{i^f jk}$ = 1, if the freighter flight i^f is processed at ramp zone j , and then goes to workstation area k for break bulk under parallel processing by double workforce; 0 otherwise;

CR_j^t the capacity ratio for ramp zone j during time interval t, which denotes the ratio of actual workload to the nominal processing capacity of a ramp zone;

CR_k^t the capacity ratio for workstation area k during time interval t, which denotes the ratio of actual workload to the nominal processing capacity of a workstation area;

a_j^t the exceeding value of CR_j^t over 1, if CR_j^t is greater than 1; 0 otherwise. It is an auxiliary variable which denotes the exceeding value of the capacity ratio of real workload over the processing capacity of a ramp zone;

b_k^t the exceeding value of CR_k^t over 1, if CR_k^t is greater than 1; 0 otherwise. It is an auxiliary variable which denotes the exceeding value of the capacity ratio of real workload over the processing capacity of a workstation area;

Input parameters

U^{i^p} the number of ULDs on passenger flight i^p ;

U^{i^f} the number of ULDs on freighter flight i^f ;

T_{jk} equipment transferring time from ramp zone j to workstation area k ;

C_j the processing capacity of ramp j ;

C_k the processing capacity of workstation area k ;

$m^t_{i^p jk}$ the workload at ramp zone j in terms of the number of ULDs during interval t for a passenger flight i^p , which is assigned to ramp zone j , and workstation area k ;

$m^t_{i^f jk}$ the workload at ramp zone j in terms of the number of ULDs during interval t for a freighter flight i^f , which is assigned to ramp zone j , and workstation area k ;

$n^t_{i^p jk}$ the workload at workstation area k in terms of the number of ULDs during interval t for a passenger flight i^p , which is assigned to ramp zone j , and workstation area k ; these ULD are under processing by one checking team;

$n^t_{i^f jk}$ the workload at workstation area k in terms of the number of ULDs during interval t for a freighter flight i^f , which is assigned to ramp zone j , and workstation area k ; these ULD are under processing by two checking teams;

The mathematical formulation is:

Objectives Type I (minimize the overall flow time):

Minimize $$\sum_{i^p \in I^p} \sum_{j \in J} \sum_{k \in K} U^{i^p} T_{jk} x_{i^p jk} + \sum_{i^f \in I^f} \sum_{j \in J} \sum_{k \in K} U^{i^f} T_{jk} y_{i^f jk} \ , \tag{3.1}$$

Objectives Type II (minimize the maximal pair-wise difference of workload):

Minimize $$CR_{\max\_r} \ , \tag{3.2}$$

Minimize $$CR_{\max\_b} \ , \tag{3.3}$$

Objectives Type III (minimize the overall exceeding value of workload):

Minimize $\sum_{j\in J}\sum_{t\in T} a_j^t$, (3.4)

Minimize $\sum_{k\in K}\sum_{t\in T} b_k^t$, (3.5)

Subject to:

Assignment constraint:

$$\sum_{j\in J}\sum_{k\in K} x_{i^p jk} = 1, \qquad \text{for } \forall i^p \in I^p, \qquad (3.6)$$

$$\sum_{j\in J}\sum_{k\in K} y_{i^f jk} = 1, \qquad \text{for } \forall i^f \in I^f, \qquad (3.7)$$

Capacity Ratio constraint for each ramp zones / workstation area:

$$\sum_{i^p\in I^p}\sum_{k\in K} x_{i^p jk} m_{i^p jk}^t + \sum_{i^f\in I^f}\sum_{k\in K} y_{i^f jk} m_{i^f jk}^t = CR_j^t C_j, \qquad \text{for } \forall t \in T, \forall j \in J, \qquad (3.8)$$

$$\sum_{i^p\in I^p}\sum_{j\in J} x_{i^p jk} n_{i^p jk}^t + \sum_{i^f\in I^f}\sum_{j\in J} y_{i^f jk} n_{i^f jk}^t = CR_k^t C_k, \qquad \text{for } \forall t \in T, \forall k \in K, \qquad (3.9)$$

The additional constraints for $CR_{\max\_r}$ and $CR_{\max\_b}$:

$$CR_{j_1}^t - CR_{j_2}^t \le CR_{\max\_r}, \qquad \text{for } \forall t \in T, \forall j_1 \in J, \forall j_2 \in J, \quad (3.10)$$

$$CR_{k_1}^t - CR_{k_2}^t \le CR_{\max\_b}, \qquad \text{for } \forall t \in T, \forall k_1 \in K, \forall k_2 \in K, \quad (3.11)$$

Additional constraint for Objective Type III:

$$CR_j^t - 1 \le a_j^t, \qquad \text{for } \forall t \in T, \forall j \in J, \qquad (3.12)$$

$$CR_k^t - 1 \le b_k^t, \qquad \text{for } \forall t \in T, \forall k \in K, \qquad (3.13)$$

Integrality and non-negativity constraint:

$$x_{ijk} \in \{0,1\},\ y_{i^f jk} \in \{0,1\},\ CR_j^t \ge 0, CR_k^t \ge 0, a_j^t \ge 0,\ b_k^t \ge 0 \qquad (3.14)$$

The time horizon of the problem is chosen as one whole week which wraps around from the end to the beginning, since the airline's flights are scheduled on a weekly repetitive basis. In order to express the performance throughout a week, the entire time horizon is divided into equally small time intervals. The collection of these intervals is denoted as the set *T*. To formulate our problem, we need to estimate the utilization of each facility during each discrete time interval *t* first.

Before the use of the term "*workload*", it is essential to articulate the meaning of it. The definition of *workload* is defined hereinafter: the number of ULDs that is currently assigned to be processed during the specific time interval.

An assignment of a flight *i* to ramp zone *j* and breakbulk workstation area *k* can be identified by the unique combination of *i*, *j*, and *k*. Thus the workload of a flight *i* at ramp zone *j* during a given time interval *t* is denoted as m_{ijk}^{t}. The workload of a flight *i* at workstation area *k* during a given time interval *t* is denoted as n_{ijk}^{t}. Hence, a unique flight assignment defines the workload profiles for this flight spanned over the entire time horizon at both ramp and workstation. The coefficients m_{ijk}^{t} and n_{ijk}^{t} are the consequences of each assignment.

The ratio of the cumulated workload to the nominal capacity of a facility during a given time interval *t* is mentioned as capacity ratio, which is denoted as CR^{t}. The capacity ratio is also the result of the assignment.

3.1.3 Model description

This section explains the rationale and thinking process in conceiving these constraints and objectives in the MIP model.

There are several considerations for the construction of the constraints. One of the most apparent constraints is the assignment constraint. Besides that, in order to describe the exceeding values of the workload beyond nominal capacities, the capacity ratio constraints are introduced. In addition, the non-negativity and integrality constraints are essential, too. These are elaborated in the following:

Assignment constraint:

$$\sum_{j\in J}\sum_{k\in K} x_{i^p jk} = 1, \qquad \text{for } \forall i^p \in I^p,$$

$$\sum_{j\in J}\sum_{k\in K} y_{i^f jk} = 1, \qquad \text{for } \forall i^f \in I^f,$$

Constraint (3.6) ensures that only one ramp zone and one breakbulk area are assigned for one passenger flight. Constraint (3.7) ensures that one freighter flight is processed by exactly one ramp zone and one breakbulk area.

Capacity Ratio constraint for each ramp zones:

$$\sum_{i\in I^p}\sum_{k\in K} x_{i^p jk} m^t_{i^p jk} + \sum_{i\in I^f}\sum_{k\in K} y_{i^f jk} m^t_{i^f jk} = CR^t_j C_j, \qquad \text{for } \forall t \in T, \forall j \in J,$$

Capacity Ratio constraint for each workstation:

$$\sum_{i\in I^p}\sum_{j\in J} x_{i^p jk} n^t_{i^p jk} + \sum_{i\in I^f}\sum_{j\in J} y_{i^f jk} n^t_{i^f jk} = CR^t_k C_k, \qquad \text{for } \forall t \in T, \forall k \in K,$$

Constraints (3.8) and (3.9) capture the workload at the ramp zone and breakbulk area during each time window. $m^t_{i^p jk}$ and $m^t_{i^f jk}$ are the workload coefficients for each flight at the ramp zone during each time window. $n^t_{i^p jk}$ and $n^t_{i^f jk}$ are the workload coefficients for each flight at the breakbulk area during each time window. These coefficients represent the workloads in terms of the number of ULDs during each time window.

The coefficients $m^t_{i^p jk}$ and $m^t_{i^f jk}$ are estimated under the assumption that only one unit of workforce handles the cargos at the ramp zone, regardless whether the cargos are from passenger flights or freighter flights. Thus the estimation methods for m of both passenger and freighter flights are the same. However, the coefficients $n^t_{i^p jk}$ and $n^t_{i^f jk}$ are estimated under the assumption that, one checking team handles the cargos from passenger flights while two checking teams handle the cargos from freighter flights, at breakbulk workstation areas. Thus the estimation method for n of passenger flights and freighter flights are different only in dealing with the processing rate of checking team(s).

This estimation procedure is further elaborated in Section 3.2.

It is not difficult to see that, the left-hand-side of the above equation gives the workload assigned to a given facility at a given time interval. The right-hand-side is the product of the capacity ratio (for this facility during a given time interval) and the facility's capacity. The capacity ratio variables indicate the usage of this facility, it can be less, equal, or more than 1. If it is less than 1, it means this facility is not fully utilized; if it is greater than 1, it implies that this facility is over utilized.

The additional constraints for $CR_{\max\_r}$ and $CR_{\max\_b}$:

$$CR_{j_1}^t - CR_{j_2}^t \le CR_{\max\_r}, \qquad \text{for } \forall t \in T, \forall j_1 \in J, \forall j_2 \in J,$$
$$CR_{k_1}^t - CR_{k_2}^t \le CR_{\max\_b}, \qquad \text{for } \forall t \in T, \forall k_1 \in K, \forall k_2 \in K,$$

Constraints (3.10) to (3.11) are meant to capture the maximum pair-difference of each two capacity ratios. By reducing $CR_{\max\_r}$ and $CR_{\max\_b}$, the difference between any two facilities could be lessen, and as a result the workload could be distributed more balanced and evenly for each ramp zone and each breakbulk workstation area, during each time interval.

Additional constraint for Objective Type III:

$$CR_j^t - 1 \le a_j^t, \qquad \text{for } \forall t \in T, \forall j \in J,$$
$$CR_k^t - 1 \le b_k^t, \qquad \text{for } \forall t \in T, \forall k \in K,$$

Constraints (3.12) and (3.13) are the additional constraints to capture the exceeding amount of workload over the nominal processing capacity, at each facility. If the capacity ratio is greater than 1, the facility is over utilized. In this case, the exceeding value a_j^t or b_k^t will be greater than 0.

Integrality and non-negativity constraint:

$$x_{ijk} \in \{0,1\}, \ y_{i^f jk} \in \{0,1\}, \ CR_j^t \ge 0, CR_k^t \ge 0, a_j^t \ge 0, \ b_k^t \ge 0$$

Constraint (3.14) is the integrality and non-negativity constraint for all variables.

Objective type I: minimize the overall flow time

Minimize $$\sum_{i^p \in I^p} \sum_{j \in J} \sum_{k \in K} U^{i^p} T_{jk} x_{i^p jk} + \sum_{i^f \in I^f} \sum_{j \in J} \sum_{k \in K} U^{i^f} T_{jk} y_{i^f jk} ,$$

The overall flow time of all ULDs is the most important measurement for evaluating the system efficiency. The overall flow time includes the pure traveling time and the waiting plus storage time due to congestion. The expression (3.1) captures only the pure movement time for each flight. As the waiting and intermediate storage times are hard to estimate in the linear model. These are, however, indirectly captured through the capacity ratios. Later in the simulation part, these waiting time and intermediate storage time will be counted directly.

Objective type II: minimize the maximal pair-wise difference of workload

Minimize $CR_{\max\_r}$,

Minimize $CR_{\max\_b}$,

Objectives (3.2) and (3.3) aim at minimizing the maximal pair-wise differences of workload at ramp zones and breakbulk areas during each time interval. Through these two objectives, it is expected to give a more balanced assignment of workload to the different facilities.

Objective type III: minimize the overall exceeding value of workload

Minimize $$\sum_{j \in J} \sum_{t \in T} a_j^t ,$$

Minimize $$\sum_{k \in K} \sum_{t \in T} b_k^t ,$$

Objectives (3.4) and (3.5) are meant to minimize the amount of the exceeding values of workload, over the nominal processing capacity at the ramp zones and breakbulk areas. Through these two objectives, the possibilities of the workload greater than the nominal capacity would be reduced and, the utilization ratio of equipment will be maintained at an ideal rate.

3.2 The estimation of the coefficients

The purpose of this section is to explain how to derive the coefficients *m* and *n* in the MIP model.

These coefficients stand for the workload of each flight assignment during each time interval. In real situations, the amount of cargo processing at ramp zones and breakbulk workstations for each flight is difficult to estimate, because of the randomness of the loading profile of a flight, the randomness of the arrival time, and the unexpected congestion due to the randomness of cargo amount and arrival time. In our study, a rough estimation methodology is proposed based on the assumption of sequential, deterministic patterns of arrival. To reflect the actual process, the parameters used in deriving the workload coefficients are obtained from the actual data through statistical means.

The estimation consists of two stages:

Stage 1. Time estimate for a given assignment.

This part identifies the arrival and departure time of ULDs of a flight for a given assignment. That is, using the parameters of the arrival and departure rates of ULDs, we try to find out the time window for each flight cargo at a ramp zone and a breakbulk area. This effort provides the estimated entering time and leaving time of the ULDs at the ramp zone and the breakbulk workstation area.

Stage 2. Cargo volume estimate based on time information.

At this stage, we calculate the values of workload level for each ramp zone and each breakbulk area for all the time intervals within the reviewing time horizon. These values indicate how many ULDs of a particular flight will be at a ramp zone or a breakbulk workstation area during any given time interval.

In order to carry out the calculation for all the coefficients, we further make the following assumptions:

1) The ULD arrival rates for each flight at both the ramp zone and the breakbulk workstation are assumed to be constant.

2) The arrival rate at the breakbulk area is equal to the departure rate at the ramp zone.

The following two sub-sections elaborate the approach to estimate the various times for a flight assignment, as well as the workload coefficients. The last section gives a brief description of the relevant data.

3.2.1 Estimate of the times

The following notations will be used in the next two sections:

a	the ULD arrival rate at a facility;
p	the ULD processing rate at a facility;
Δt	the time required to finish one ULD at a facility;
u	total number of ULDs on a flight;

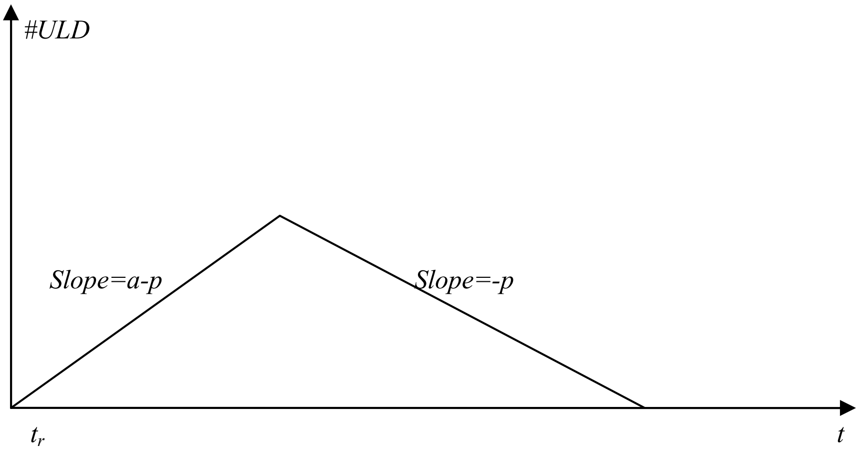

Figure 3.1 The workload profile of a flight

The first entering time of a flight at a given facility, is defined as the arrival time of the first ULD of this flight at this facility: t_r. The first entering time of a flight is approximated as the flight arrival time at the airport, as our assumption in Section 3.1.1.

The last entering time of a flight at a given facility, is defined as the arrival time of the last ULD of this flight at this facility: $t_r + (u-1)/a$.

The first leaving time of a flight at a given facility, is defined as the departure time of the first ULD of this flight at this facility: $t_r + \Delta t_u$.

The last leaving time of a flight at a given facility, is defined as the departure time of the last ULD of this flight at this facility: $t_r + \Delta t_u + (u-1)/p$.

In all cases, we have $\Delta t_u = 1/p$.

The above procedure illustrates the basic approach to estimate the workload in terms of the number of ULDs at both ramp zones and breakbulk workstation areas. As stated in the assumptions in Section 3.1.1, the processing rate p for freighter flights at workstation areas is two times of that for passenger flights at workstation areas. Hence, when calculating the coefficients using this method, the estimation for coefficients of freighter flights at workstation areas should be treated differently by doubling up the processing rate.

To implement such an estimation procedure, the arrival flight information is needed. Relevant information such as the weekly flight schedule, airplane loading capacity, typical utilization rate of the airplane space are therefore collected. As mentioned earlier, we use the flight arrival time at the airport as the first entering time of the flight at the cargo terminal.

3.2.2 Estimate of the workload coefficients

Based on the assumptions and information provided in previous sections, we can calculate the exact workload profile of a flight for a given assignment.

The procedures to calculate the workload at the ramp zone within the time interval *[t, t+Δt)* is:

Step1: find the maximum (*max*) of the time *t* and the first entering time at the ramp zone or breakbulk area for this flight assignment;

Step 2: find the minimum (*min*) of the time *t+Δt* and the last leaving time at the ramp zone or breakbulk area for this flight assignment;

Step 3: if $max \le min$, then the number of ULD being processed during time interval *[t, t+Δt]* is given by, p * (*min* - *max*); else, the number of ULD being processed is zero. Return the value of workload as the workload measurement for this time interval *[t, t+Δt)*.

An illustrative example:

Total 6 ULDs of passenger flight AA001 start arrival at the ramp at 11:58pm. This flight is assigned to ramp zone #2 and breakbulk workstation area SA.

The following essential information is available.

Ramp Zone #2:

Arrival rate: 0.5 ULD/min

Departure rate: 0.28 ULD/min

Single ULD processing time: 1 / 0.28 = 3.53 min / 1 ULD

Traveling time from ramp zone #2 to breakbulk area SA: 0.77 minutes

Breakbulk Area SA:

Arrival rate: 0.28 ULD/min

Departure rate: 0.2 ULD/min

Single ULD processing time: 1 / 0.2 = 5 min / 1 ULD

Time estimates at ramp zone:

First entering time at ramp zone, t_r = 11:58,

Last entering time at ramp zone, $t_r + (u-1)/a$ = 11:58 + 10 = 12:08,

First leaving time at ramp zone, $t_r + \Delta t_u$ = 11:58 + 3.53 = 12:01.53,

Last leaving time at ramp zone, $t_r + \Delta t_u + (u-1)/p$ = 11:58 + 3.53 + 17.65 = 12:19.18.

Time estimates at breakbulk area:

The processing time for one ULD at ramp zone is 3.53 minutes, and the traveling time from ramp zone 2 to breakbulk area SA is 0.77 minutes,

First entering time at breakbulk area, t_r = 12:01.53 + 0.77 = 12:02.30,

Last entering time at breakbulk area, $t_r + (u-1)/a$ = 12:02.30 + 17.65 = 12:19.95,

First leaving time at breakbulk area, $t_r + \Delta t_u$ = 12:02.30 + 5 = 12:07.30,

Last leaving time at breakbulk area, $t_r + \Delta t_u + (u-1)/p$ = 12:02.30 + 5 + 25 = 12:32.30.

Estimate of workload coefficients:

If given time interval [12:05, 12:10), then the workload coefficient at ramp zone for this flight is 1.40. The complete ramp zone workload coefficients for this flight assignment are {0.57, 1.40, 1.40, 1.40, 1.23} for continuous time intervals { [11:55, 12:00), [12:00, 12:05), [12:05, 12:10), [12:10, 12:15), [12:15, 12:20) }, and 0 for other time intervals.

Similarly, the workload coefficients at breakbulk workstation area for this flight given time interval [12:05, 12:10) is 1.00. The complete breakbulk area workload coefficients for this flight assignment is {0.54, 1.00, 1.00, 1.00, 1.00, 1.00, 0.46 } for time intervals { [12:00, 12:05), [12:05, 12:10), [12:10, 12:15), [12:15, 12:20), [12:20, 12:25), [12:25, 12:30), [12:30, 12:35) }, and 0 for other time intervals.

3.2.3 Data information for estimate

To estimate the various times and workload coefficients of each flight, we need the ULD traveling time from different ramp zone to different breakbulk area. The movement data for these different flow patterns were deduced from the equipment specifications and extracted from the data collected. The processing rate and the arrival rate of ULDs at ramp zones and breakbulk areas are also estimated statistically based on actual data.

Furthermore, the work content for handling each flight is measured in terms of the number of ULDs. The number of ULDs loaded on each flight is estimated from the actual data statistically. In general, they vary according to different aircraft types.

4 Simulation Modeling

This chapter describes the simulation model of the cargo terminal operations. In the simulation model, we simulate the ULD movement and processing within the cargo terminal. It can capture the randomness and queuing effects which have not been addressed in the deterministic optimization model. The purpose of the simulation is to evaluate and compare the performances of different assignment plans which are obtained by solving the multiple-objective optimization models.

In Section 4.1, the simulation model is described. It starts from model flow descriptions and assumptions. Then the rules and policies used in the simulation are presented, followed by the simulation inputs and performance measures. This simulation model is implemented in AutoMod®. After the proper verification of the simulation modeling approach in Section 4.2, the simulation setup is illustrated in Section 4.3. This simulation model is validated based on the statistics of the pilot-run results in Section 4.4.

4.1 Simulation model design

This section starts from introducing the major process flow of the simulation in Section 4.1.1. Then the system operation rules and policies are given in Section 4.1.2. The inputs and output performance measures for the simulation are discussed in Section 4.1.3 and Section 4.1.4 respectively. In Section 4.1.5, the implementation of this model in AutoMod® is briefed introduced.

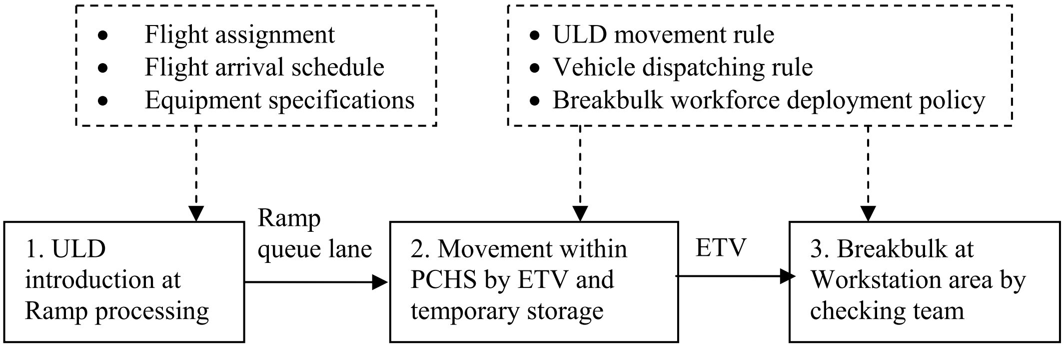

Figure 4.1 The simulation model framework

The major framework of this simulation is presented in the Figure 4.1. The three solid boxes stand for the major process flows in the simulation model, which will be described in more detail in Section 4.1.1.

A graphical illustration of the simulation layout plan is given in Figure 4.2.

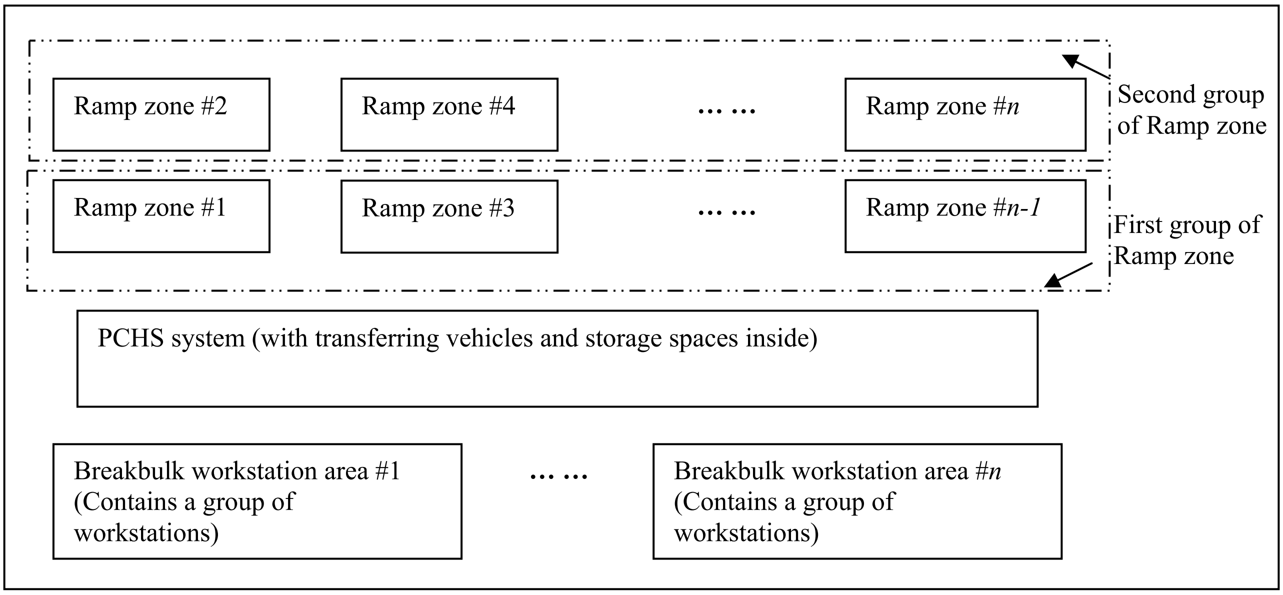

Figure 4.2 An illustration of simulation layout

As can be seen from Figure 4.1, the first process of cargo flow starts from the ULD introduction process at the ramp zone. Hence in order for future processing, the ULDs are moved onto ETV by the ramp queue lane, and then into PCHS by the ETV. This process in simulation is performed based on the direction of the flight assignment plan, the flight arrival schedule, and the equipment specifications. The flight assignment plan provides the movement pattern for each ULD from various flights. Along with the specifications of the ramp zone equipments, the flight arrival schedule gives the timing of these movements. After this process, the ULD is moved to next process which is the movement within the PCHS and temporary storage.

The second step of ULD movement is represented in the simulation by various movement procedures in the PCHS. The common properties of these movements are that they are all performed by the transferring vehicles, i.e. ETV. These ETVs are the electricity-driven vehicles with fixed capacities moving between the origin and destination positions for each ULD movement within the PCHS. In this process, the ULD movement rule and the vehicle dispatching rule together control the movement mechanism of ULDs within PCHS.

Finally, the ULD is transferred to breakbulk workstation areas via the ETV. The breakbulk process is performed by the checking teams at the workstations. Hence, at this stage, the ULD movement rule and the breakbulk workforce deployment rule dictate the processing at workstation areas.

The dashed box at the top-left of Figure 4.1 indicates the inputs information for the simulation, which will be further introduced in Section 4.1.3. It provides the flight assignments, arrival times, aircraft types, and equipment specifications to help decide the times, paths, and quantities for the ULD movements.

The dashed box at the top-right includes the various rules and policies for ULD movements, vehicle dispatching, and workforce deployment. They will be further discussed in Section 4.1.2.

As denoted in Figure 4.2, the ramp zones are grouped into two different sections in order to facilitate the analysis and modeling of the problem. The ramp zone group which is more close to the workstation areas has to be distinguished from the ramp zone group which is further away from the workstation areas, as the ULDs exhibit completely different movement patterns at these two ramp zone groups.

There are two floors in the cargo terminal, with a few breakbulk workstations on each floor. The ramp zone which locates at the ground floor serves as the cargo entrance to the cargo terminal. It is the starting position for the cargo movement within the terminal. As mentioned, the ramp zone group that is closer to the ground floor workstation areas serves the cargoes which are dispatched for the ground floors, the second floor, and the outbound terminal as well; however, the ramp zone group that is further away from the ground floor workstation areas serves only the cargoes which are dispatched for the second floor and the outbound terminal. PCHS stands at the opposite side of the breakbulk workstation areas with the coverage of each floor.

The PCHS and the Hoist within PCHS serve as the transportation equipments linking various cargo movements, such as moving from the ramp zone to the ramp zone, moving from the ramp zone to the workstation, and moving from the ramp zone to the outbound terminal, etc.

The workstations on each floor are the equipments which serves the breakbulk process of the cargoes.

4.1.1 Model description

The purpose of this section is to describe the cargo handling process in the simulation model in more detail.

The following assumptions for the model are made:

1) The arrival time of the flight at the airport is taken as the first entering time of ULD of this flight at the cargo terminal.

2) The arriving process of ULDs from the same flight follows an empirical distribution which is an approximation for the arrival process.

3) The number of ULDs on each flight also follows an empirical distribution, which varies by aircraft types.

4) The ULD movements are executed in the system according to the order of their arrival times.

5) The ULD always moves along the shortest path within the PCHS, if no special requirements.

6) A freighter flight needs two checking teams at workstation, while a passenger flight usually requires one only at breakbulk stage. If there are not enough checking teams to handle the flights, the ULDs of these flights have to wait in the PCHS until there are enough teams to perform the breakbulk job, according to the checking team deployment rule which will be elaborated in Section 4.1.2.

7) There are a fixed number of checking teams at the cargo inbound terminal.

8) The checking teams are assigned to the flights based on a workforce deployment rule which will be given in more detail in Section 4.1.2.

As described in the previous section, due to their different positions in the terminal, these two distinguished group of ramp zones are defined as follows: The ramp zone group which is closer to the workstation areas is defined as the first group of ramp zones, while the ramp zone group which are further away to the workstation areas is defined as the second group of ramp zones.

Hence, the descriptions about the system cargo flow logic are presented in two parts. The first part (Section 4.1.1.1) will mainly discuss the cargo movement at the first group of ramp zones, while the second part (Section 4.1.1.2) will concentrate on the cargo movement at the second group of ramp zones. A general logic flow of the overall cargo movements is presented in Figure 4.3.

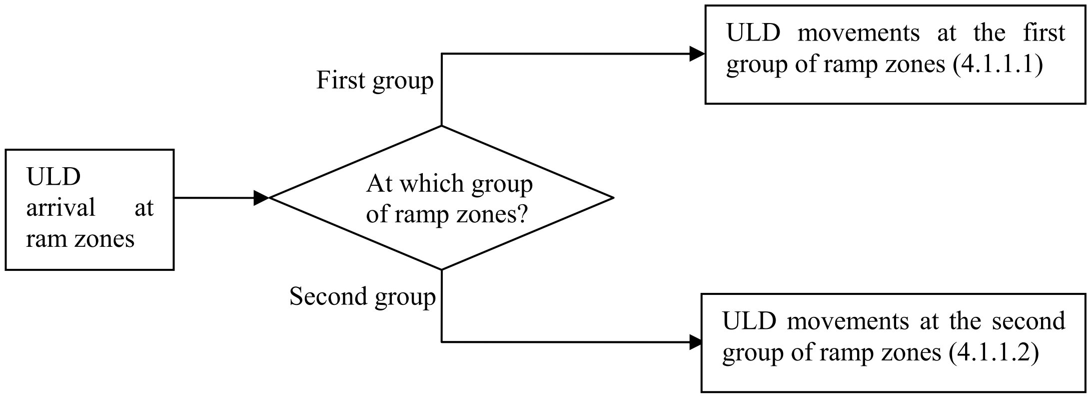

Figure 4.3 General flow of ULD movements

Such a general flow pattern is in line with the general ULD flow within the PCHS as illustrated in Figure 1.4. Figure 1.4 gives a general flow pattern of all ULDs regardless their arrival positions; however Figure 4.3 focuses on the more in-depth differentiation of the ULDs according to their different arrival positions, i.e., the ramp zone group where they start movements within the system.

In the upcoming sections, two different movement patterns will be described according to this differentiation as shown in Figure 4.2. They are also in line with the illustrations in Figure 1.3 and Figure 1.4 with a stronger focus on the detailed ULD movements within PCHS and the ULD breakbulk process at the workstation areas.

4.1.1.1 The cargo movement at the first group of ramp zones

The ramp zones in this group are located at the closer side to the breakbulk workstations. Figure 4.4 below illustrates the basic ULD flow process within the system.

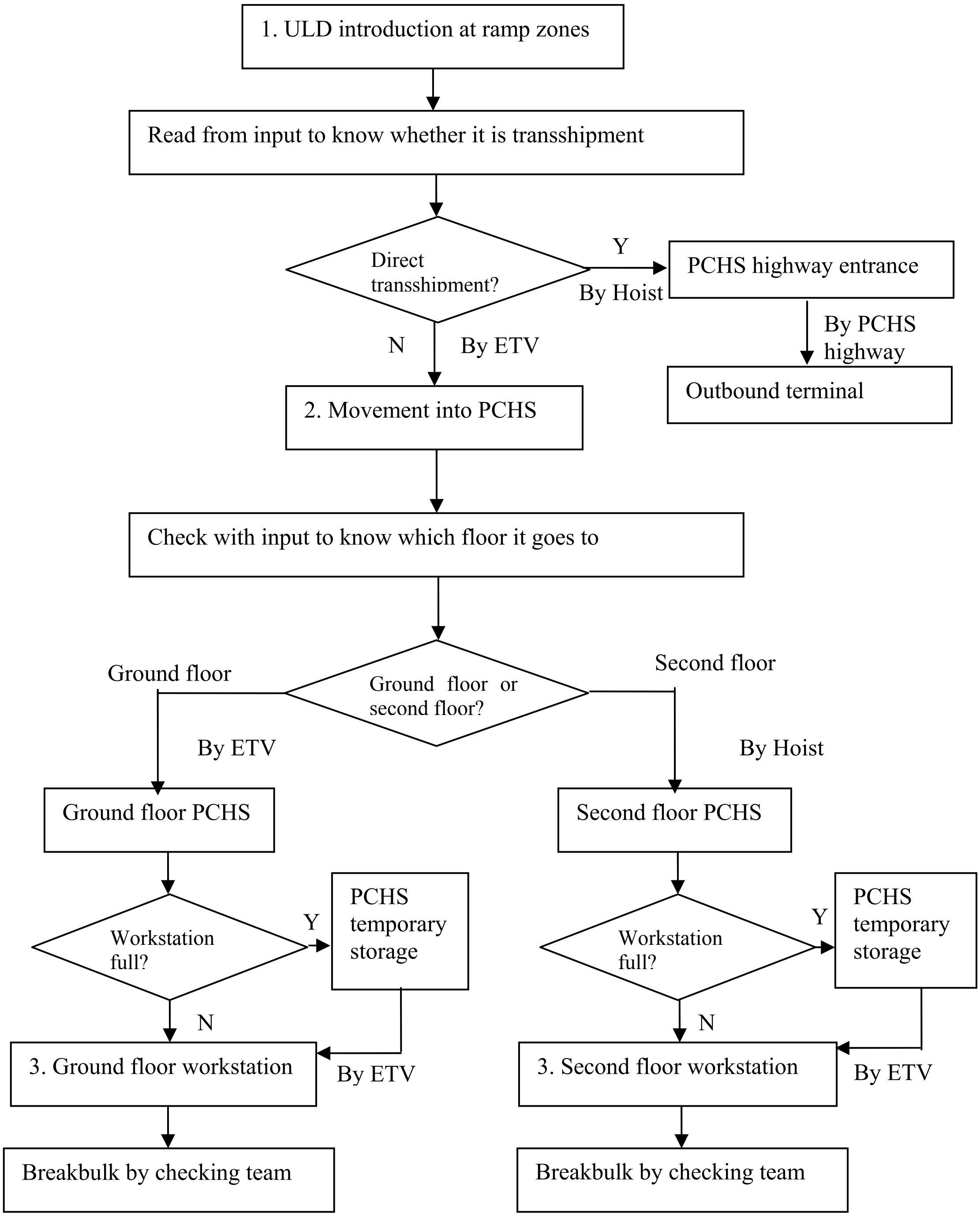

Figure 4.4 Flow chart of the ULD movement at the first group of ramp zones

As illustrated in Figure 4.4, there are three major steps during the ULD movement at the first group of ramp zones. They are described in more detail as follows.

1. ULD introduction at the ramp processing

The ULDs assigned to this group of ramp zones will move either to the ground floor or to the second floor of the terminal, except some direct transshipment ones which move directly to the outbound terminal through the PCHS Highway. The PCHS highway is another electricity-driven conveyor which locates right between the inbound terminal and the outbound terminal horizontally.

ULDs are loaded onto ramp queue lanes as soon as they arrived at the terminal.

2. Movement within the PCHS and the temporary storage

ULDs to breakbulk at the ground floor of the terminal, are transferred from the ramp queue lane to the workstation via the ETV. The ETV moves the ULD to the ground floor breakbulk workstations if these workstations are not occupied. If these workstations are occupied, the ULD is therefore moved to the buffer space in PCHS for temporary storage.

ULDs to breakbulk at the second floor of the terminal, are transferred from the ramp queue lane to the hoist via the ETV. The ETV moves the ULD to the hoist entrance. The PCHS buffer serves as the queuing space for the ULD to wait for the hoist to be free. If the hoist is occupied, the ULD is therefore moved to the buffer space in the PCHS for temporary storage.

The ULD designated for the second floor then arrives at the PCHS second floor after exiting from the hoist. It is transferred from the PCHS to the workstation at second floor via ETV. Again, the PCHS can also serve as the queuing space if the workstation is currently not free.

The ULD movement and ETV scheduling rules in Section 4.1.2 provide the necessary directions on these movements. Thus the ULD which arrives early will be processed early in the system. And also the ETV processes its service requests by the time orders it received from various ULD movements.

3. Breakbulk at the workstation area by checking teams

As stated, right after the ULD arriving at the workstation, the checking teams start the breakbulk process according to the checking team deployment policies in Section 4.1.2. The productivity and size of the workforce together determine the time duration for each ULD at breakbulk stage. After breakbulk, the ULDs are moved to their respective destinations according to their respective needs.

4.1.1.2 The cargo movement at the second group of ramp zones

The ramp zones of this group are stationed at the farther side to the breakbulk workstations. The graph below in Figure 4.5 gives a clear illustration of the ULD movements within this group.

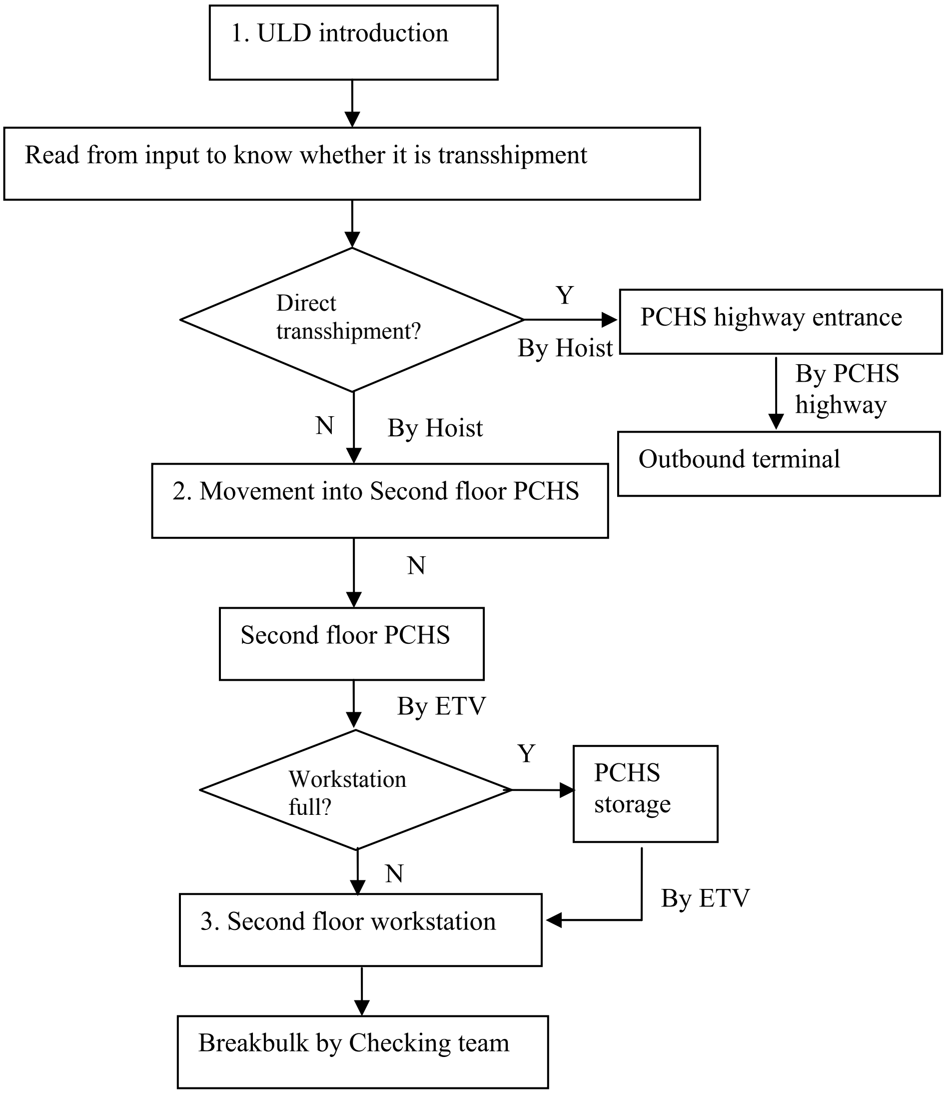

Figure 4.5 Flow chart of the ULD movement at the second group of ramp zones

As illustrated in Figure 4.5, there are three major steps during the ULD movement at the second group of ramp zones. They are described in more detail as follows.

1. ULD introduction at the ramp processing

The ULD assigned to this group of ramp zones will move to breakbulk workstations at second floor, or to the outbound terminal through the PCHS highway.

The ULD is initially introduced to an elevating hoist via a queue lane.

2. Movement within the PCHS and the temporary storage

After the ULD is transferred onto the second floor of the PCHS, it is then moved to the second floor breakbulk workstations by ETV. The ULD movement and ETV dispatching rules in Section 4.1.2 together determine the movement route of each ETV. Thus the ETV moves the ULD to its planned destination under these directions. The ULD which arrives early will be moved early by its related ETV. Furthermore, the ETV processes its service requests by the time orders it received from various ULD movements.

However, if congestion occurs, the ULD still needs the temporary storage within the PCHS. This is in line with the movement and dispatching rules which will be described in Section 4.1.2.

3. Breakbulk at workstation area by checking teams

As stated in the previous section, the checking teams start the breakbulk process right after the ULD arriving at the workstation according to the checking team deployment policies in Section 4.1.2. After breakbulk, the ULDs are moved to their respective destinations according to their respective needs. The productivity and size of the workforce together determine the amount of time for each ULD at the breakbulk stage.

4.1.2 Rule and policy description

This section describes the ULD movement and the vehicle dispatching rules in the system, as well as the checking team deployment policy.

These rules and policies are abstracted based upon actual operations. Hence, by incorporating these rules and policies into our simulation model, it could be more consistent with the actual practices in the air cargo inbound terminal.

4.1.2.1 ULD movement rule: shortest traveling time

The ULD movement rule is incorporated into the simulation model as shown in Figure 4.1. Together with the vehicle dispatching rule, it directs the ULD's traveling paths in the PCHS given their arrival positions and destinations at the inbound terminal.

The exact dispatching rules to allocate the ULD to different segments of the ramp zone and workstation area need to be further investigated. The two main reasons are:

1. Due to the complexity of the real system, it is impossible for the system to identify exactly which queue lane of the designated ramp zone, or which workstation of the designated breakbulk area would be used during work, without the movement and dispatching rule.

2. The purpose of the simulation model is to illustrate the random movement of ULD between certain facilities and their processing at certain ramp zone or workstation area.

The ULD movement rule is described as follows:

The ULD moves from its origin to its destination via the shortest possible path, which is by observation the direct path between the origin and destination positions, in most cases.

Most of these ULDs would move along their shortest path, which is favored by the system. In case the designated workstation for this ULD is full during peak hour, the ULD will be put at the temporary storage space before it continues to move, based on the automatic decision from the central control system.

When a ULD movement involves with more than one vehicle due to the need of the intermediate storage, this ULD will be moved to the PCHS temporary storage by its first vehicle, then it will be picked up by the next vehicle to resume its movement from the intermediate storage to the next processing facility.

4.1.2.2 Vehicle scheduling rule: first in first out according to the ULD request sequence

The vehicle scheduling rule is incorporated into the simulation model as shown in Figure 4.1. Together with the ULD movement rule, it explains how the elevating transferring vehicles (ETV) move inside the PCHS system.

The traveling vehicle is called upon request whenever there is a ULD arriving at the pick-up point of the queue lane. The requests for the same vehicle will be processed according to their time sequences. Hence, the earlier job request for the vehicle will be processed prior to the later requests.

Whenever a ULD movement involves with two or more vehicles successively, all these related vehicles are waken up at the same moment. These vehicles will be called for service accordingly.

4.1.2.3 Checking team deployment policies at breakbulk workstations

To approximate the import cargo movements more accurately and comprehensively, it is necessary to include the breakbulk process at the workstation stage also. Hence, the breakbulk checking team deployment rule for simulation is incorporated into the model, as well.

The import process at breakbulk workstation involves with the human resources: checking team, and the equipment resources: breakbulk workstations.

Based on the cargo checking and breakbulking process, the policy to dispatch the workforce is carried out according to the following steps.

Step 1. Based on the flight assignment plan, the flights are assigned to the breakbulk workstation areas.

Step 2. Randomly assign the currently available checking teams to the arrival flights. The flights for a checking team are saved in a list. If it is a freighter flight, then two checking teams are assigned to this flight.

Step 3. Define a usage level for each checking team. Once the team is working, the usage level is set to 1, otherwise 0.

Step 4. When the ULD comes from PCHS, the simulation logic will check the current status of its corresponding checking team, if it is 1 (working), the ULD would stay in the buffer which is inside the PCHS; if it is 0 (idle), this ULD would start to arrive at the workstation.

Step 5. After finishing the breakbulk for all ULDs of a flight, the status of this checking team is set back to 0 again.

4.1.3 Input parameters

Proper input information is required for the simulation model based on the initial study.

1. Flight assignment plan: this assignment plan indicates which ramp zone and breakbulk workstation area are assigned to which flight. This information directs the ULD movements within the terminal.

2. Flight arrival time: the weekly flight schedule provides the arrival time for each incoming flight. Hence, it also provides the start processing time of the ULD on each flight at the PCHS. In the simulation model, the flight arrival time is a deterministic

value, while the ULD arrival time is a stochastic value which is derived from flight arrival time and is based on the empirical random distribution. The ULD arrival time for the same ULD varies from week to week.

3. Equipment specifications: the movement speed and capacity of an ETV, or other transferring equipments are the necessary information for the simulation, as well. Furthermore, information on the processing capacities of ramp zones and breakbulk workstation areas is also important for the simulation.

4. Aircraft types: this information provides the carrying capacity of each aircraft type. Therefore, the number of ULDs on the flights varies according to different aircraft types. In the simulation model, the number of ULDs is a random value which is generated based on the empirical distribution of the number of ULDs on some certain type of aircraft. The number of ULDs on the same flight changes from one week to another.

4.1.4 Performance measure

The simulation model is built as a tool to analyze the performances of the ULD movements in the PCHS system under various settings and configurations. Therefore by assessing those different settings and configurations, we can choose the one that offers the better favorable performance than others.

In order to compare the different designs, the main objective in the MIP model in Chapter 3, namely the overall flow time is chosen as the measurement for performance.

The overall flow time is the primary objective in the MIP model. Moreover, it is straightforward for management to understand, especially the decision makers. Besides, it is one of the most important concerns for the decision making for cargo terminals. In addition, other objectives discussed in Chapter 3 are less important for the decision making of cargo terminals, as they are created to further polish the solutions of the mixed-integer programming model. Therefore, the overall flow time is chosen as the measurement.

4.1.5 Model Implementation

This simulation model involves with different movement systems and different equipments. It is implemented in the AutoMod® 11.0 package. A brief introduction about AutoMod® is given in the first subsection. The description of the translation of simulation model into AutoMod® is provided in the second subsection.

4.1.5.1 Introduction on AutoMod®

AutoMod® provides users with a simulation environment that facilitates the modeling and analysis of logistics systems. It combines a simulation language with a graphically interfaced "simulator". It contains several systems and collections of multiple entities. An AutoMod model must have a process system, and can have one or more movement systems optionally. The main systems in AutoMod® are as follows.

Process system

The process system defines the logic that controls how products (loads) are processed in a simulation model. An AutoMod model has only one process system. The process system could allow user to define many entities within a model, such as resources (machines or operators) and queues (waiting lines) for those resources.

Movement system

A movement system contains components that can be used to simulate the movement of loads, such as the components in a conveyor system. A path mover system in AutoMod is a flexible path-based system which can be used to simulate automated guided vehicles (AGVs), fork trucks, or personnel who carry loads through a facility along a predetermined path. Any number and combinations of movement systems are accepted in an AutoMod model.

Movement system includes:

- Conveyor system
- Path mover system (AGV, fork truck, etc)
- Power & free (widely used in the automotive industry)
- AS/RS (Automated Storage/Retrieval systems)
- Bridge cranes
- Kinematics (used for robotic modeling)
- Tanks & pipes

Loads

Loads are the active entities in AutoMod and can be generated in many ways, including deterministic or random generation. They are used to represent physical entities that move through a system, such as the products, freights, or people. Each load has a user-defined description called a load type. Hence some useful attributes can be added and modified for a load type.

4.1.5.2 Model Implementation in Automod

This section explains how to implement the simulation model in AutoMod.

There are primarily three types of systems included in the model design for our problem using AutoMod, namely the processing system, the conveyor system, and the path mover system. A sample layout graph in AutoMod is given in Figure 4.6. Thorough definitions of these systems are presented as follows.

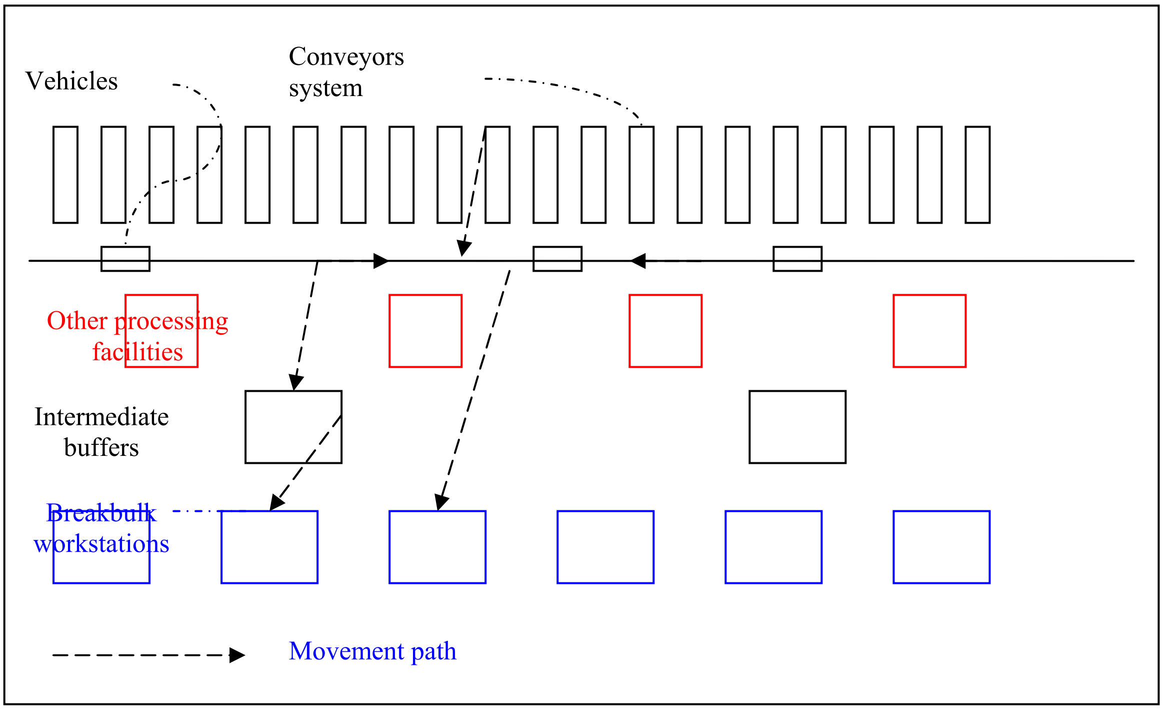

Figure 4.6 Sample layout of simulation model

Process system

The process system in this model represents the air cargo movement processes, such as the hoist movement process, the breakbulk process at breakbulk workstation area, etc. This process system includes the following components:

- **Load**

The Load object is defined for each ULD. The ULD movement within the PCHS is represented by the movements of the Load within the system.

- **Resource**

All the cargo handling facilities are modeled as a limited capacity resource entity with finite processing rate. Therefore, the ramp zone, breakbulk workstation area, and the hoist are all modeled as the Resource objects.

- **Queue**

For each Resource object, there must be a Queue object served as the queuing buffer for its processing. In addition, the intermediate storage spaces are also modeled as Queue entities.

Conveyor system

The ramp queue lanes at the ground level of the cargo terminal are defined as the conveyor system in the simulation model.

Path mover system

In this AutoMod model, the path mover system is employed to represent the movement of the ETV within the PCHS system, in which the ETV moves between the loading points and unloading points for the ULD to pick-up and drop-off.

4.2 Verification of the model

This AutoMod model needs proper verification procedures before being used as the simulation model. The verification procedures are performed as the following three steps:

1. The structure of the simulation model is verified against the actual layout of the physical terminal. As shown in Figure 4.6, the simulation model is consistent with the physical model and physical relationships between different entities in the physical terminal. Moreover, the simulation model describes the actual relations between different systems within the terminal. The graphical size of the AutoMod simulation

model is also proportional to the real physical size of the terminal and respective equipments.

2. The processes and objects in this simulation are also verified against the actual operations and entities in the inbound terminal. As stated in step 1, all the physical entities in the actual terminal have their respective modeling objects in the simulation model. Therefore, all the related operations which occur at some certain systems are all reflected as some movement processes, processing processes or queuing processes in the simulation model. The validity of the process modeling is further examined using the comparisons of simulation data against actual data, as stated in the simulation validation part.

3. Based on the previous two stages, the system and component parameters are further verified according to their corresponding specifications in reality. These parameters are calculated based on the data collection and equipment specifications. The processing rate, processing capacity, queuing capacity, length, width, as well as other characteristics are verified against their physical and geometrical attributes which are collected from the initial data collection and equipment specifications.

Based upon the above verification procedures, the AutoMod simulation model is verified against the physical system and ready to effectuate in the next stage.

4.3 Simulation setups and pilot runs

This section addresses some of the simulation setups and pilot runs based on the replications / deletions approach in Law and Kelton (2000) and the inputs of the simulation procedures.

4.3.1 Simulation run design

The performance measure of such a simulation is the overall flow time for all cargoes. Hence, the observation of the overall flow time for all cargoes is collected each time when the simulation clock passes a week (for the weekly simulation) or a day (for the daily simulation).

The simulation is designed as a non-terminating simulation, which means, the simulation clock evolves all the time until it is forced to stop. Therefore the length of warm-up period, the length of replications, and the number of replications should be determined based on initial pilot runs before the collection of data statistics from production runs.

In order to collect sufficient statistics for the observations of the concerned performance measure, we need to estimate the length of warm-up period based on the initial pilot runs. The warm-up process is the initial transient process of a simulation before the steady-state means of the performance measure is reached. After the deletion of the data from warm-up process, the steady-state statistics could be collected based on the replication runs. For example, if T_1, T_2, T_3, …, T_m, are the observations in one replication, then the best estimator for T is suggested to be

$$\overline{T}(m,l) = \frac{\sum_{j=l+1}^{m} T_j}{m-l}, \tag{4.1}$$

Where m is the total number of observations, and l is the number of observations during warm-up period.

As the observations in the beginning of the simulation may not be very representative of steady-state behaviors due to the initial conditions, it is better to eliminate these biases by deleting the data from initial observations. Law and Kelton (2000) suggested a straightforward approach to determine the length of warm-up period (l).

1. Make n replications each with length m for the simulation. Let T_{ij} denote the *j-th* observation in the *i-th* replication, where $i = 1, 2, 3, \ldots, n$, and $j = 1, 2, 3, \ldots, m$;

2. Let $\overline{T}_j$ denote the mean of the *j-th* observations from these n replications, hence we have:

$$\overline{T}_j = \frac{\sum_{i=1}^{n} T_{ij}}{n}, \text{ for } j = 1, 2, \ldots, m, \tag{4.2}$$

3. Let $\overline{T}_j(w)$ denote the moving average of $\overline{T}_j$ with w as the window, hence we have:

$$\overline{T}_j(w) = \begin{cases} \dfrac{\sum_{s=-w}^{w} \overline{T}_{j+s}}{2w+1} & j > w \\ , & , \text{ for } j = 1, 2, \ldots, m, \\ \dfrac{\sum_{s=-(j-1)}^{j-1} \overline{T}_{j+s}}{2j+1} & j \le w \end{cases} \tag{4.3}$$

Where w is a positive integer and $w \le \left\lfloor \frac{m}{2} \right\rfloor$.

4. The graphs of $\overline{T}_j(w)$ for different w are plotted with j on the x-axis, and the value of $\overline{T}_j(w)$ on the y-axis. Choose the suitable value of l so that the value of $\overline{T}_j(w)$ appears to converge, when the value of j exceeds l.

This is a trial-and-error process to determine the value of warm-up length l as the value of n might be very large in order to tackle the variability of $\overline{T}_j$. Moreover, the value of m is of consideration for the convergence of $\overline{T}_j(w)$, too. The value of m should also be much greater than the value of l in order to include more extraordinary and infrequent events.

After the suitable design of the simulation run based on above approach is achieved from these pilot runs, the production runs are then performed. Hence, a series of production runs were conducted to replicate the simulation for both the one-week and one-day problem. Suppose that we make n' replications each with length m' observations for production runs, where the value of m' is much larger than the value of warm-up period l determined by the above approach. The value of the *j-th* observation in the *i-th* replication run is still defined as T_{ij}. Let X_i denote the steady state mean for the *i-th* replication, then we have:

$$X_i = \frac{\sum_{j=l+1}^{m'} T_{ij}}{m' - l}, \text{ for } i = 1, 2, \ldots, n', \tag{4.4}$$

Since these X_i are independent identical distribution variables, their values can be estimated by their mean value $\overline{X}(n')$ with the approximate confidence level of (1 - α), using the following formula:

$$\overline{X}(n') \pm t_{n'-1,1-\alpha/2}\sqrt{\frac{S^2(n')}{n'}}, \tag{4.5}$$

Here in Equation 4.5, $\overline{X}(n')$ is the average of X_i for total *n'* replications (i = 1, 2, …, *n'*), and $S^2(n')$ is the variance of X_i for total *n'* replications.

The statistics results collected from production runs based on the replication / deletion approach are presented in the following sections.

4.3.2 Pilot runs for one-day simulation

The simulation runs are implemented in the AutoMod® 11.0 simulation environment, on a PC Pentium IV 2.60 GHz platform, with 512 MB build-in memory.

The pilot runs for one-day simulation with 100 observations in each replication are conducted for 5 replications. It was decided that the length of warm-up period could be l = 14 for these 13 different designs in the one-day problem with run length m = 100.

A sample of the graphs of the pilot runs for one-day simulation is given in Figure 4.7. The x-axis illustrates the sequence of the observations, and the y-axis stands for the overall flow time in the simulation for all cargo entities.

The sample average graph denotes the average of the performance measures in each replication run for each observation. The moving average denotes the average of the sample average values within a fixed number of continuous selections, with the selections of sample averages move forward along the *x*-axis.

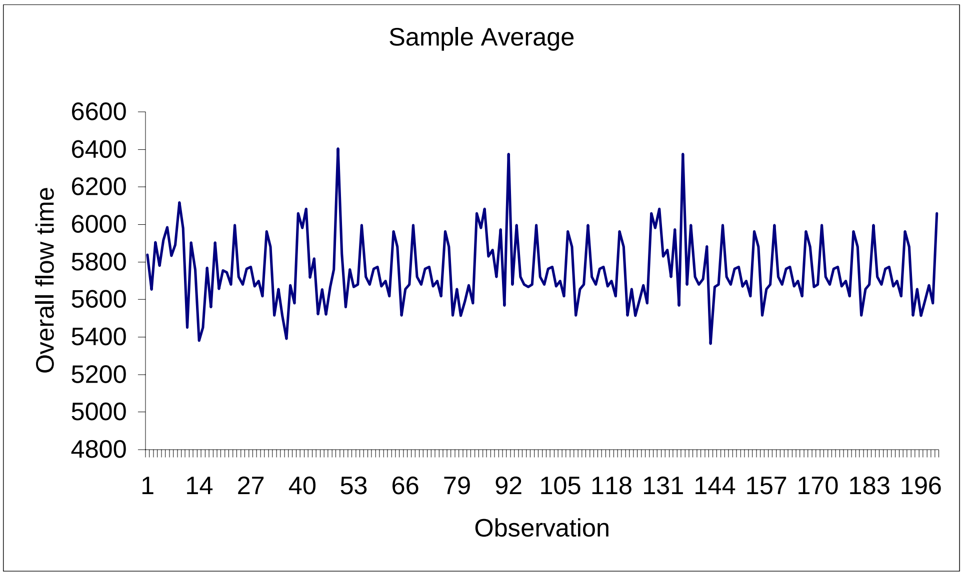

Figure 4.7 Sample average of replication runs (one-day)

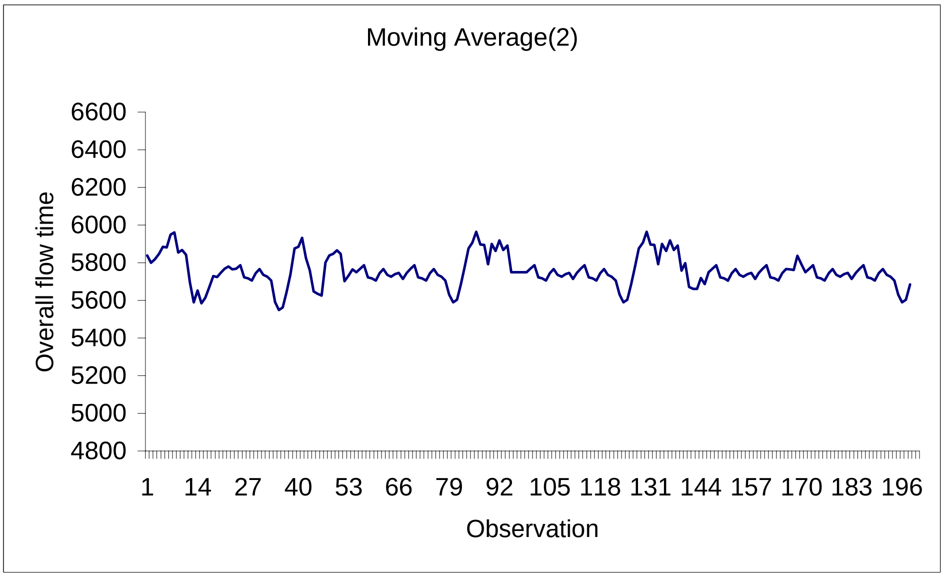

Figure 4.8 Moving average of replication runs (one-day, window size = 2)

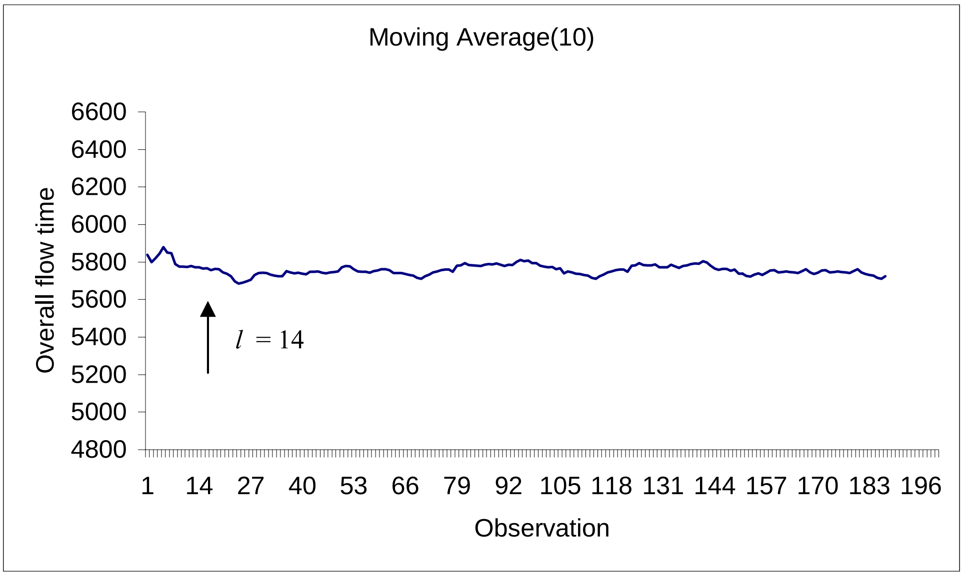

Figure 4.9 Moving average of replication runs(one-day, window size = 10)

As shown in Figure 4.7, the averaged values of overall flow time in these observations from the pilot runs are plotted. It is therefore preferable to further smooth the curve as the curve itself is quite unstable. Several values of the window size for moving average process are tested, and the most representative results are plotted in Figure 4.8 and Figure 4.9. As can be seen from the graph, when window size w = 2, it is still not sufficient to estimate the warm-up period l. By enlarging the window size w to 10, it is evident that the warm-up period length l can be chosen as 14, which is large enough to be the warm-up length for all the runs of different designs.

Therefore, the production runs for each design are to run for the number of replications n' = 200 with length m' = 500 and warm-up length l = 14. The product run results are given in Chapter 5.

4.4 Validation of the model

The purpose of this section is to validate the simulation model design based on actual data. It employs a three-step approach to perform this task.

In order to validate the simulation model, the simulation framework is run for the duration of one day only, with the actual one-day flight arrival information as the simulation inputs. The simulation outputs are collected and organized based on proper warm-up time. The data during warm-up period are not included in the output analysis. In fact, the purpose of gathering the overall flow time for each ULD object is to compare the simulated flow time with the actual flow time which is collected from the on-site study.

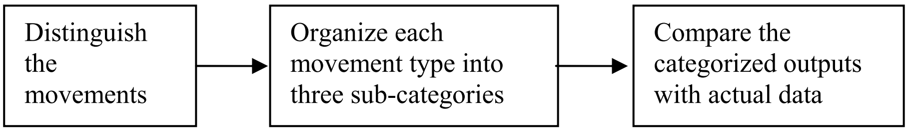

Figure 4.10 Three steps for validation

This validation for simulation is performed in three steps as shown in Figure 4.10:

The *first* step is to distinguish these output movement data by their different origin and destination positions in the system. Therefore, the movements are grouped according to their origin-destination pairs. The movements are different with each other in the sense that their starting positions and ending positions are different with each other. The origins of these movements are the different ramp zones in the terminal. The destination positions of these movements are the breakbulk workstations and the outbound terminal.

The *second* step is to further classify each movement data into three categories, according to the different queuing circumstances.

To differentiate the effects of possible queuing, the movement data are clustered into three categories based on their corresponding actual time durations. An estimate of the pure movement duration for each movement type was done beforehand based on the equipment specifications. The movements are split into 3 groups, namely Tier 1, Tier 2, and Tier 3. Details about each tier are given below:

Tier 1: almost no queuing, i.e., the corresponding actual duration is less than or equal to 1.5 times of the estimated pure movement duration;
Tier 2: light queuing, i.e., the corresponding actual duration is less than or equal to 3 times of the estimated pure movement duration, and excluding Tier 1;
Tier 3: heavy queuing, i.e., the corresponding actual duration is greater than 3 times of the estimated pure movement duration.

The *final* step is to compare the simulated movement data against their corresponding actual movement data under different sub-categories.

At this stage, all different movements are distinguished according to their origin-destination positions. Furthermore, each origin-destination category is divided into three sub-categories to illustrate different congestion levels for each origin-destination path. The simulated durations and the actual durations of the movements under different sub-categories are compared against each other. The simulation model validation result is then provided based on the statistical test of this comparison.

In order to compare the actual movement data with the simulation output data of various movement patterns, a *t*-test is conducted for the data under each category. A hypothesis test is performed to see whether there is significant difference in means between the simulation data and the actual data of each movement under each category. The null hypothesis and the alternate hypothesis are as follows:

H_0: there is no difference between the mean of simulation output data and the mean of actual data;
H_1: there is significant difference between the mean of simulation output data and the mean of actual data.

This hypothesis test is performed as a two-sample pooled *t*-test. Although the simulation tries to capture the exact movements of every single ULD and compare against its counterparts in reality, due to the uncertainty and human intervention, we are only interested in the pooled *t*-test results instead of the paired *t*-test.

The sample size of each *t*-test is listed in Table 4.1:

Table 4.1 Sample size of each hypothesis test

Movement category	Tier 1	Tier 2	Tier 3	Overall
Movement #1	2	16	65	83
Movement #2	0	0	6	6
Movement #3	0	0	1	1
Movement #4	0	1	1	2
Movement #5	7	6	15	28
Movement #6	0	4	13	17
Movement #7	1	3	7	11
Movement #8	10	13	43	66
Movement #9	1	0	2	3
Movement #10	8	11	40	59
Movement #11	1	2	4	7
Movement #12	10	16	1	27
Movement #13	7	17	59	83

The probability values of these *t*-tests are calculated based on *t*-tests results. The *p*-values of these hypothesis tests are given in Table 4.2.

Table 4.2 *p*-value of each hypothesis test

Movement category	Tier 1	Tier 2	Tier 3
Movement #1	-	0.201	0.245
Movement #2	-	-	0.270
Movement #3	-	-	-
Movement #4	-	-	-
Movement #5	0.271	0.039	0.211
Movement #6	-	0.131	0.310
Movement #7	-	0.032	0.324

Movement #8	0.302	0.011	0.293
Movement #9	-	-	-
Movement #10	0.324	0.289	0.292
Movement #11	-	-	0.153
Movement #12	0.291	0.504	-
Movement #13	0.227	0.189	0.251

* 1. The mark “-“ denotes that the sample size is too small to perform the statistical test.

As can be shown from Table 4.2, most of the *p*-values of these hypothesis tests are well above 0.05 (the common significance level). 3 *p*-values of the tests are less than 0.05, however, they are all greater than 0.01. If the confidence level is set to be 95%, only 3 test results imply that the means of the two samples are different. Therefore, if given the confidence level of 99% for all these tests, the test results suggest that the means of actual data are not different from the means of simulation outputs.

Hence, the validation for the simulation model is accomplished by the three steps.

5 Solution and Result Presentations

This chapter gives the optimization solutions of the MIP model, and the simulation results of the simulation model.

The solution approach to tackle the multiple-objective MIP problem is presented in Section 5.1.

In Section 5.2, the optimal assignment plans for the one-day problem are obtained by solving the multiple-objective MIP problem using the procedures described in Section 5.1. These efficient solutions are evaluated by the simulation model.

The same experiment is repeated for the one-week problem and the results are shown in Section 5.3.

5.1 Optimization procedures

The purpose of this section is to illustrate the approach to solve the multiple-objective MIP problem. An ε-Constraint approach is employed to explore the efficient solutions for this multiple-objective problem.

The original multiple objectives problem is:

$$\min_{x \in X} \left(f_1(x), \ldots, f_Q(x) \right). \tag{5.1}$$

The ε-Constraint method is different from weighted sum of objectives, instead it only minimizes one of the objectives, while the others are transformed to constraints. Given the original problem objectives:

$$\min_{x \in X} \left(f_1(x), \ldots, f_Q(x) \right), \tag{5.2}$$

The converted problem can be rewritten as:

$$\min_{x \in X} f_k(x), \text{ subject to } f_i(x) \le \varepsilon_i, \forall i = 1, \ldots Q, \; i \ne k \tag{5.3}$$

The ε serves as the reference value which is determined by the preferences of Decision Maker (DM). Thus it offers the flexibility to adjust the right-hand-side of the constraint value to adapt to the aspiration level of DM. Without loss of generality, fine-tuning the value of ε could ensure that the objective-turned-constraints are always closest to the desired value.

This multiple objective problem would not provide a single optimal solution but a series of non-dominated solutions. With the consideration of both solution quality and computation time, this ε-Constraint method is employed here to explore these pareto-optimal solutions.

We maintain one of the objectives as the unique objective function for optimization, and then adjusting the value of ε-Constraints for other objectives. Thus, a corresponding MIP model is constructed. By solving the MIP model, an efficient solution will be achieved.

Therefore, different sets of values for the ε-constraints would result in different solution designs for this problem. Here several sets of the values of ε were set according to their respective most desirable values, and the solutions under each design were obtained through the optimization of the mixed-integer program model. The major procedures are outlined here:

1. Relax the problem with one single objective only. Hence, the most desirable value of each objective was achieved (refer to the Table 5.2). Further on, they would serve as the basis values for future adjustment of the ε.

2. Convert the objective (3.2), objective (3.3), objective (3.4), and objective (3.5) as the ε-constraints, set the values of ε based on the percentage adjustment of their basis values. (refer to Table 5.1)

Table 5.1 Experiment designs

Objective sets	Adjustable ranges
Objective (3.2), Objective (3.3)	110%, 150%, 200%
Objective (3.4), Objective (3.5)	150%, 200%

The computation experiments based on the above analysis establishes the framework to explore some of the non-dominated solutions for this multiple objective mixed-integer program model. Computation procedures are deployed using solution package ILOG® CPLEX 8.0, on a PC Pentium IV 2.60 GHz platform, with 512 MB build-in memory. The output results for both one-week and one-day problem are further obtained within limited computation time.

5.2 Results and outputs for the one-day problem

The purpose of this section is to summarize the solutions results and the simulation outputs for the one-day problem.

5.2.1 MIP solution results for the one-day problem

The flights arrival data of the one-day problem is based on the actual flight arrival time during one single day. The cargo handling volume of each flight at the air cargo import terminal is measured in the number of ULD, which is also calculated from the actual flight information. The preset time interval in the model for this one-day problem is set as 5 minutes to better represent the cargo handling workload in detail. Hence, the workload is captured as in every 5-minute interval, and there are altogether 288 time intervals during the day.

This one-day problem consists of 63 flights of which 58 are passenger flights and 5 are freighter flights. The possible assignment plan tries to assign each flight to three ramp zones and four breakbulk workstation areas. Since there are 12 = 3 * 4 possible assignments for one single flight, the total number of flight assignment binary integer variables is 63 * 12 = 756. The time interval length equals 5 minutes, therefore there are total 288 time intervals for this one-day problem. Therefore, it can be seen from Equation 3.6 – 3.13, the total number of constraints for this problem is 63 + 288 * (3 + 4) + 288 * (3 * 2 + 4 * 3) +288 * (3 + 4) = 9729.

The ε-Constraint values for each objective-converted constraint in each design are the optimization results given itself as the single objective, with other objectives relaxed. In Table 5.2, the most desirable results for each objective are listed. The first column of this table denotes the optimal value of Objective (3.1). The remaining four columns present the optimal values of Objective (3.2), (3.3), (3.4), and (3.5), accordingly.

Table 5.2 Extreme values of each objective (one-day)

Objective (3.1)	Objective (3.2)	Objective (3.3)	Objective (3.4)	Objective (3.5)
1186.38	2.00	1.00	8.60	3.23

According to Table 5.1 in the previous section, there are 3 different levels for Objective (3.2) and (3.3), and 2 different levels for Objective (3.4) and (3.5). Hence, considering all the possible combinations, there are 3 * 3 * 2 * 2 = 36 different designs for the computation of results. Therefore, the respective values of each level of the objective set are illustrated in Table 5.3 with the first column denoting the respective objectives, and the remaining columns giving the corresponding values of different levels for each objective.

Table 5.3 The values of constraints for each setting (one-day)

Objectives	Respective values for each level		
(3.2)	2.20	3.00	4.00
(3.3)	1.10	1.50	2.00
(3.4)	12.97	17.20	
(3.5)	4.85	6.46	

The computation experiments take place after those values are initiated. The tolerance level of the optimization is set to 0.0005. The solution values of each objective for every solution are all recorded in Table 5.4 for this one-day problem. As shown in the table, the first column of each table gives the different settings of objective (3.2) and (3.3) based on Table 5.3, while the first row of each table presents the different settings of Objective (3.4) and (3.5).

Table 5.4 Solution results of each objective for each solution (one-day)

	Obj (3.4) ≤ 12.97, Obj (3.5) ≤ 4.85				
	Obj (3.1)	Obj (3.2)	Obj (3.3)	Obj (3.4)	Obj (3.5)
Obj (3.2) ≤ 2.20, Obj (3.3) ≤ 1.10	1846.14	2.00	1.00	12.80	4.60
Obj (3.2) ≤ 2.20, Obj (3.3) ≤ 1.50	1845.53	2.00	1.20	12.80	4.80
Obj (3.2) ≤ 2.20, Obj (3.3) ≤ 2.00	1845.53	2.00	1.20	12.80	4.80
Obj (3.2) ≤ 3.00, Obj (3.3) ≤ 1.10	1846.14	2.00	1.00	12.80	4.60
Obj (3.2) ≤ 3.00, Obj (3.3) ≤ 1.50	1845.53	2.00	1.20	12.80	4.80
Obj (3.2) ≤ 3.00, Obj (3.3) ≤ 2.00	1845.53	2.00	1.20	12.80	4.80
Obj (3.2) ≤ 4.00, Obj (3.3) ≤ 1.10	1846.14	2.00	1.00	12.80	4.60
Obj (3.2) ≤ 4.00, Obj (3.3) ≤ 1.50	1845.53	2.00	1.20	12.80	4.80
Obj (3.2) ≤ 4.00, Obj (3.3) ≤ 2.00	1845.53	2.00	1.20	12.80	4.80

	Obj (3.4) ≤ 12.97, Obj (3.5) ≤ 6.46				
	Obj (3.1)	Obj (3.2)	Obj (3.3)	Obj (3.4)	Obj (3.5)
Obj (3.2) ≤ 2.20, Obj (3.3) ≤ 1.10	1837.52	2.00	1.00	12.80	5.60
Obj (3.2) ≤ 2.20, Obj (3.3) ≤ 1.50	1831.74	2.00	1.20	12.80	6.00
Obj (3.2) ≤ 2.20, Obj (3.3) ≤ 2.00	1831.54	2.00	1.20	12.80	6.40
Obj (3.2) ≤ 3.00, Obj (3.3) ≤ 1.10	1837.52	2.00	1.00	12.80	5.60
Obj (3.2) ≤ 3.00, Obj (3.3) ≤ 1.50	1831.94	2.00	1.40	12.80	6.40
Obj (3.2) ≤ 3.00, Obj (3.3) ≤ 2.00	1831.79	2.00	1.40	12.80	6.40
Obj (3.2) ≤ 4.00, Obj (3.3) ≤ 1.10	1837.52	2.00	1.00	12.80	5.60
Obj (3.2) ≤ 4.00, Obj (3.3) ≤ 1.50	1831.79	2.00	1.40	12.80	6.40
Obj (3.2) ≤ 4.00, Obj (3.3) ≤ 2.00	1831.54	2.00	1.40	12.80	6.40

	Obj (3.4) ≤ 17.20, Obj (3.5) ≤ 4.85				
	Obj (3.1)	Obj (3.2)	Obj (3.3)	Obj (3.4)	Obj (3.5)
Obj (3.2) ≤ 2.20, Obj (3.3) ≤ 1.10	1756.07	2.00	1.00	17.20	4.60
Obj (3.2) ≤ 2.20, Obj (3.3) ≤ 1.50	1757.51	2.00	1.20	16.80	4.40
Obj (3.2) ≤ 2.20, Obj (3.3) ≤ 2.00	1757.51	2.00	1.20	16.80	4.40
Obj (3.2) ≤ 3.00, Obj (3.3) ≤ 1.10	1756.07	2.00	1.00	17.20	4.60
Obj (3.2) ≤ 3.00, Obj (3.3) ≤ 1.50	1757.51	2.00	1.20	16.80	4.40
Obj (3.2) ≤ 3.00, Obj (3.3) ≤ 2.00	1757.51	2.00	1.20	16.80	4.40
Obj (3.2) ≤ 4.00, Obj (3.3) ≤ 1.10	1756.07	2.00	1.00	17.20	4.60
Obj (3.2) ≤ 4.00, Obj (3.3) ≤ 1.50	1757.51	2.00	1.20	16.80	4.40
Obj (3.2) ≤ 4.00, Obj (3.3) ≤ 2.00	1740.86	2.00	1.20	17.00	4.80

	Obj (3.4) ≤ 17.20, Obj (3.5) ≤ 6.46				
	Obj (3.1)	Obj (3.2)	Obj (3.3)	Obj (3.4)	Obj (3.5)
Obj (3.2) ≤ 2.20, Obj (3.3) ≤ 1.10	1742.56	2.00	1.00	17.20	5.40
Obj (3.2) ≤ 2.20, Obj (3.3) ≤ 1.50	1719.81	2.00	1.20	17.20	6.40
Obj (3.2) ≤ 2.20, Obj (3.3) ≤ 2.00	1719.81	2.00	1.20	17.20	6.40
Obj (3.2) ≤ 3.00, Obj (3.3) ≤ 1.10	1742.56	2.00	1.00	17.20	5.40
Obj (3.2) ≤ 3.00, Obj (3.3) ≤ 1.50	1719.81	2.00	1.20	17.20	6.40
Obj (3.2) ≤ 3.00, Obj (3.3) ≤ 2.00	1719.81	2.00	1.20	17.20	6.40
Obj (3.2) ≤ 4.00, Obj (3.3) ≤ 1.10	1741.58	2.00	1.00	17.20	5.40
Obj (3.2) ≤ 4.00, Obj (3.3) ≤ 1.50	1719.81	2.00	1.20	17.20	6.40
Obj (3.2) ≤ 4.00, Obj (3.3) ≤ 2.00	1719.81	2.00	1.20	17.20	6.40

5.2.2 Simulation outputs for the one-day problem

After reviewing the solution results of the one-day MIP problem, it is discovered that some of these solutions are identical. Hence, there are 13 distinguished efficient solutions achieved from the one-day MIP problem, i.e., there are 13 different designs for the simulation runs.

Based upon the approach suggested in Section 4.3, 5 initial pilot runs for each design with 100 observations in each replication are conducted for the one-day simulation, in order to determine the length of warm-up period. It was decided that the length of warm-up period could be $l = 14$ for these 13 different designs with run length $m = 100$.

The production runs for each design and the original plan are then run for the number of replications $n' = 200$ with length $m' = 500$ and warm-up length $l = 14$. The statistics of the observations for all the simulation designs are presented with their means, standard deviations and 95% confidence intervals as follows:

Table 5.5 Simulation result statistics for one-day problem

	Performance measurements				Pure Traveling time in optimization	Other measures for load balancing and congestion			
Simulation design	Mean	Std. Dev.	95% CI Low	95% CI High	Obj (3.1)	Obj (3.2)	Obj (3.3)	Obj (3.4)	Obj (3.5)
12	5859.9	209.20	5854.06	5865.73	1719.81	2.00	1.20	17.20	6.40
10	6164.5	222.05	6158.27	6170.72	1740.86	2.00	1.20	17.00	4.80
13	5833.7	218.41	5827.71	5839.68	1741.58	2.83	1.00	23.49	5.40
11	5887.2	200.89	5881.54	5892.85	1742.56	2.00	1.00	17.20	5.40
8*	5832.1	205.66	5826.41	5837.79	1756.07	2.00	1.00	17.20	4.60
9	5879.5	210.24	5873.58	5885.41	1757.51	2.00	1.20	16.80	4.40
5	6173.7	219.19	6167.35	6180.04	1831.54	2.00	1.20	12.80	6.40
7	5879.2	212.37	5873.60	5884.79	1831.79	2.00	1.40	12.80	6.40
6	5939.1	216.55	5933.12	5945.07	1831.94	2.00	1.40	12.80	6.40
3	6063.9	212.68	6058.11	6069.69	1837.52	2.00	1.00	12.80	5.60
4	5890.5	214.66	5884.81	5896.19	1841.58	2.00	1.00	17.20	5.40
2	5849.9	212.74	5843.97	5855.82	1845.53	2.00	1.20	12.80	4.80
1	6077.2	220.00	6071.35	6083.04	1846.14	2.00	1.00	12.80	4.60
Original plan	6841.0	244.88	6834.05	6847.94	2081.94	2.264	1.00	43.58	64.2

* 1. The asterisk "*" denotes that the corresponding simulation design has the smallest mean value of the overall flow time.

The simulation results of the 13 different designs are shown in Table 5.5. The means and standard deviations of the overall flow time are collected with their 95% confidence intervals shown in the 4th and 5th column. Their corresponding values in MIP optimization are shown in columns 6 – 10. It is apparent to see that the overall flow time collected from simulation is much greater than the corresponding objective value in Obj (3.1) of pure traveling time.

5.2.3 Comments

The simulation results for the *one-day problem* suggest that the design #8 is the most desirable design for the one-day problem, since its overall flow time in simulation is the smallest among all 13 designs with the average overall flow time of 5832.1 minutes. This most preferable solution has its sample standard deviation of 205.66, with its 95% confidence interval [5826.41, 5837.79].

Its corresponding MIP solution suggests an efficient assignment plan with 14.7% decline in overall flow time as comparing to the simulation result for the original assignment plan of 6841.0 minutes.

A graphical display of these designs is given below in Figure 5.1, with the pure traveling time for all cargo entities in MIP optimization results on x-axis (Objective 3.1), and overall flow time in simulation for all cargo entities on y-axis. This overall flow time is the summation of the cargo flow time for all the cargo entities in the simulation. The mean value and confidence intervals of simulation statistics are displayed as High-Low-Mean values together as a dotted line.

As we can see from Figure 5.1, the most preferable efficient solution with the smallest overall flow time in simulation (design #8) does not necessarily stand for the design that has the lowest pure traveling time in the MIP optimization results. Such a discrepancy implies the possible existence of more severe congestion for other designs, thus supports our motive to approximate the congestion effects as other objective functions in the MIP problem.

The MIP provides the candidates pool for the evaluation by simulation. Thus the efficient solutions produced by MIP are meaningful and useful for the final evaluation by simulation. Without the selection pool suggested by the MIP solutions, it would be extremely hard to find a desirable design for the real problem.

By comparing the simulation outputs and optimization results of #8 with other designs, we can see that although this design does not have the smallest pure traveling time in optimization, it has the smallest value of Objective (3.2) and (3.3) and the second smallest value of Objective (3.4) and (3.5). Thus, its congestion effects as suggested in the MIP model could be possibly less than other designs. And this possibility is further proven in our simulation model because this design shows the smallest overall flow time in simulation.

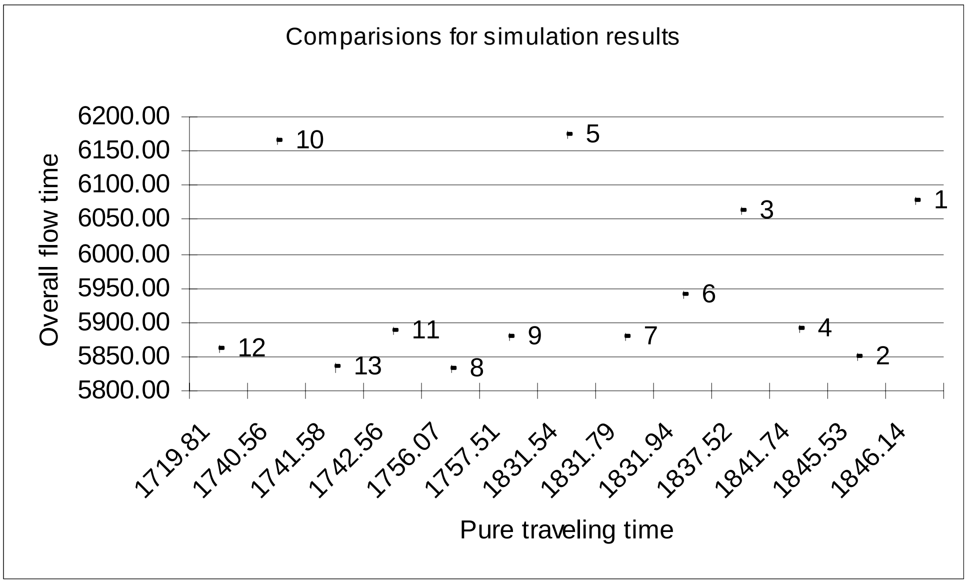

Figure 5.1 Comparisons for simulation results (one-day)

5.3 Results and outputs for the one-week problem

The purpose of this section is to summarize the solutions results and the simulation outputs for the one-week problem.

5.3.1 MIP solution results for the one-week problem

The most desirable results for each objective are given below in Table 5.6. Those results are the extreme values, which will be served later as the basis for constructing the constraints. The first column of this table denotes the optimal value of Objective (3.1). The remaining four columns present the optimal values of Objective (3.2), (3.3), (3.4), and (3.5), respectively.

Table 5.6 Extreme values of each objective (one-week)

Objective (3.1)	Objective (3.2)	Objective (3.3)	Objective (3.4)	Objective (3.5)
16920.4	1.033	0.683	7.983	3.183

Hence, the different levels for the ε-Constraints of the converted objectives are given in Table 5.7. As shown in the table, there are 3 different levels for Objective (3.2) and (3.3), and 2 different levels for Objective (3.4) and (3.5). Hence, considering all the possible combinations, there are 3 * 3 * 2 * 2 = 36 different designs for the computation of results. Therefore, the respective values of each level of the objective set are illustrated below in Table 5.7 with the first column denoting the respective objectives, and the remaining columns giving the corresponding values of different levels for each objective.

Table 5.7 The values of constraints for each setting (one-week)

Objectives	Respective values for each level		
(3.2)	1.137	1.549	2.061
(3.3)	0.750	1.025	1.367
(3.4)	11.973	15.966	
(3.5)	4.775	6.367	

The review time interval in the model for this one-week problem is set as 60 minutes to better represent the workload in detail as every one hour in one week. Therefore, the overall time horizon for this problem is 7 days, which contains exactly 24 * 7 = 168 time windows.

The flights arrival data of one-week problem is based on the recent flight schedule adopted by this airline company. As each flight can be possibly assigned to three ramp zones and four breakbulk workstation areas, there are 12 = 3 * 4 possible assignments for one single flight. This one-week problem consists of 151 flights of which 113 are passenger flights and 38 of which are freighter flights. Hence the total number of flight assignment binary integer variables is 151 * 12 = 1812. The time interval length equals one hour, so there are 168 intervals for this one-week problem. Therefore, it can be seen from Equation 3.6 – 3.13, the total number of constraints for this problem is 151 + 168 * (3 + 4) + 168 * (12 + 6) +168 * (3 + 4) = 5527.

Likewise, the mixed-integer-programming solution results for one-week problem are obtained by the same approach that is adopted for the one-day problem. The solution results for the one-week MIP problem were listed below in Table 5.8. Since there are 3 * 3 * 2 * 2 = 36 designs for the computation procedure, the result for each design is listed under each solution result's corresponding constraints. The tolerance level of the optimization is set at 0.005. The values of each objective for every solution are all recorded in this table for this one-week problem. As shown in the table, the first column of each table gives the different settings of objective (3.2) and (3.3) based on Table 5.7, while the first row of each table presents the different settings of Objective (3.4) and (3.5).

Table 5.8 Solution results of each objective for each solution (one-week)

	Obj (3.4) ≤ 11.973, Obj (3.5) ≤ 4.775				
	Obj (3.1)	Obj (3.2)	Obj (3.3)	Obj (3.4)	Obj (3.5)
Obj (3.2) ≤ 1.137, Obj (3.3) ≤ 0.750	-	-	-	-	-
Obj (3.2) ≤ 1.137, Obj (3.3) ≤ 1.025	34239.8	1.135	1.017	11.183	4.753
Obj (3.2) ≤ 1.137, Obj (3.3) ≤ 1.367	34234.0	1.135	1.017	11.183	4.753
Obj (3.2) ≤ 1.549, Obj (3.3) ≤ 0.750	-	-	-	-	-
Obj (3.2) ≤ 1.549, Obj (3.3) ≤ 1.025	34234.0	1.135	1.017	11.183	4.753
Obj (3.2) ≤ 1.549, Obj (3.3) ≤ 1.367	34213.6	1.135	1.017	11.183	4.753
Obj (3.2) ≤ 2.061, Obj (3.3) ≤ 0.750	-	-	-	-	-
Obj (3.2) ≤ 2. 061, Obj (3.3) ≤ 1.025	34208.0	1.135	1.017	11.183	4.753
Obj (3.2) ≤ 2.061, Obj (3.3) ≤ 1.367	34204.4	1.135	1.017	11.183	4.753

	Obj (3.4) ≤ 11.973, Obj (3.5) ≤ 6.367				
	Obj (3.1)	Obj (3.2)	Obj (3.3)	Obj (3.4)	Obj (3.5)
Obj (3.2) ≤ 1.137, Obj (3.3) ≤ 0.750	-	-	-	-	-
Obj (3.2) ≤ 1.137, Obj (3.3) ≤ 1.025	34239.8	1.135	1.017	11.183	4.753
Obj (3.2) ≤ 1.137, Obj (3.3) ≤ 1.367	34178.0	1.135	1.017	11.183	4.753
Obj (3.2) ≤ 1.549, Obj (3.3) ≤ 0.750	-	-	-	-	-
Obj (3.2) ≤ 1.549, Obj (3.3) ≤ 1.025	34231.2	1.135	1.017	11.183	4.753
Obj (3.2) ≤ 1.549, Obj (3.3) ≤ 1.367	33864.4	1.187	1.233	11.786	6.195
Obj (3.2) ≤ 2.061, Obj (3.3) ≤ 0.750	-	-	-	-	-
Obj (3.2) ≤ 2.061, Obj (3.3) ≤ 1.025	34208.0	1.135	1.017	11.183	4.753
Obj (3.2) ≤ 2.061, Obj (3.3) ≤ 1.367	33864.4	1.187	1.233	11.786	6.195

	Obj (3.4) ≤ 15.966, Obj (3.5) ≤ 4.775				
	Obj (3.1)	Obj (3.2)	Obj (3.3)	Obj (3.4)	Obj (3.5)
Obj (3.2) ≤ 1.137, Obj (3.3) ≤ 0.750	-	-	-	-	-
Obj (3.2) ≤ 1.137, Obj (3.3) ≤ 1.025	34234.0	1.135	1.017	11.183	4.753
Obj (3.2) ≤ 1.137, Obj (3.3) ≤ 1.367	34231.2	1.135	1.017	11.183	4.753
Obj (3.2) ≤ 1.549, Obj (3.3) ≤ 0.750	-	-	-	-	-
Obj (3.2) ≤ 1.549, Obj (3.3) ≤ 1.025	33128.0	1.450	0.983	12.417	4.696
Obj (3.2) ≤ 1.549, Obj (3.3) ≤ 1.367	33128.0	1.450	0.983	12.417	4.696
Obj (3.2) ≤ 2.061, Obj (3.3) ≤ 0.750	-	-	-	-	-
Obj (3.2) ≤ 2.061, Obj (3.3) ≤ 1.025	33128.0	1.450	0.983	12.417	4.696
Obj (3.2) ≤ 2.061, Obj (3.3) ≤ 1.367	33128.0	1.450	0.983	12.417	4.696

	Obj (3.13) ≤ 15.966, Obj (3.14) ≤ 6.367				
	Obj (3.1)	Obj (3.2)	Obj (3.3)	Obj (3.4)	Obj (3.5)
Obj (3.2) ≤ 1.137, Obj (3.3) ≤ 0.750	-	-	-	-	-
Obj (3.2) ≤ 1.137, Obj (3.3) ≤ 1.025	34234.0	1.135	1.017	11.183	4.753
Obj (3.2) ≤ 1.137, Obj (3.3) ≤ 1.367	33786.2	1.127	1.067	12.342	4.237
Obj (3.2) ≤ 1.549, Obj (3.3) ≤ 0.750	-	-	-	-	-
Obj (3.2) ≤ 1.549, Obj (3.3) ≤ 1.025	33128.0	1.450	0.983	12.417	4.696
Obj (3.2) ≤ 1.549, Obj (3.3) ≤ 1.367	33034.6	1.617	1.200	15.396	5.879
Obj (3.2) ≤ 2.061, Obj (3.3) ≤ 0.750	-	-	-	-	-
Obj (3.2) ≤ 2.061, Obj (3.3) ≤ 1.025	33128.0	1.450	0.983	12.417	4.696
Obj (3.2) ≤ 2.061, Obj (3.3) ≤ 1.367	33034.6	1.617	1.200	15.396	5.879

* 1. Some of the entries which are denoted as "-" suggest that the solution results could not exhibit the trend of convergence or even cause the exhaustion of computer memory within 48 hours.

It is observed that the tolerance level is set relatively higher as comparing to traditional requirements. The reason is that this problem is difficult to reach optimum due to its enormous size within maximum computation time of 48 hours. Furthermore, the purpose of the optimization process is to explore the solution space for suitable good enough solutions to serve as the inputs for the simulation model. Due to some approximations in estimating the coefficients for optimization model, the variation of coefficients may significantly affect the degree of difficulty of the optimization process. Therefore, a proper tolerance level is selected, and given the current tolerance level, a series of equally good solutions in terms of these objectives are generated. The exact performance of each of the efficient solutions is to be tested by the simulation for more accurate choice of the possible design.

5.3.2 Simulation outputs for the one-week problem

After reviewing the solution results of the one-week MIP problem, it is discovered that some of these solutions are identical. Hence, there are 11 distinguished efficient solutions for the one-week MIP problem.

Based upon the approach suggested in Section 4.3, 5 initial pilot runs for each design with 200 observations in each replication are conducted for the one-week simulation, in order to decide the length of warm-up period. It was decided that the length of warm-up period could be $l = 42$ for these 11 different designs with run length $m = 200$.

The production runs for each design and the original plan are then run for the number of replications $n' = 50$ with length $m' = 200$ and warm-up length $l = 42$. The statistics of the observations for all the simulation designs are presented with their means, standard deviations and 95% confidence intervals as in Table 5.9:

Table 5.9 Simulation result statistics for one-week problem

Simulation design	Performance measurements				Pure Traveling time in optimization	Other measures for load balancing and congestion			
	Mean	Std. Dev.	95% CI Low	95% CI High	Obj (3.1)	Obj (3.2)	Obj (3.3)	Obj (3.4)	Obj (3.5)
6	58249	2039.4	57674.12	58837.70	33034.6	1.62	1.20	15.40	5.88
10	59448	2118.9	58838.18	60064.49	33128.0	1.45	0.98	12.42	4.70
4	71996	2553.7	71270.19	72721.51	33786.2	1.13	1.07	12.34	4.24
9	55991	2032.2	55415.54	56563.87	33864.4	1.19	1.23	11.79	6.20
5	67275	2393.5	66614.38	67936.61	34178.0	1.14	1.02	11.18	4.75
7	56940	2040.8	56346.42	57541.26	34204.4	1.14	1.02	11.18	4.75
1	67507	2366.6	66850.35	68177.52	34208.0	1.14	1.02	11.18	4.75
11	57927	1977.7	57349.47	58495.04	34213.6	1.14	1.02	11.18	4.75
8	55238	2067.0	54677.47	55809.99	34231.2	1.14	1.02	11.18	4.75
3*	54733	1940.3	54215.69	55260.99	34234.0	1.14	1.02	11.18	4.75
2	60317	2169.5	59741.47	60906.68	34239.8	1.14	1.02	11.18	4.75
original	71185	2577.2	70437.83	71926.96	57770.8	2.29	2.24	20.29	49.86

* 1. The asterisk “*“ denotes that the corresponding simulation design has the smallest mean value of the overall flow time.

The simulation results of the 11 different designs are shown in Table 5.9. The means and standard deviations of the overall flow time are collected with their 95% confidence intervals shown in the 4th and 5th column. Their corresponding values in MIP optimization are shown in columns 6 – 10.

5.3.3 Comments

As shown in Table 5.9, the simulation design #3 displays the smallest average overall flow time among the 11 efficient designs for the assignment plan. This most desirable design provides the sample mean of 54733 minutes for the overall flow time, with sample standard deviation of 1940.3 minutes, given the 95% confidence interval of [54215.69, 55260.99]. Again, we observe that the overall flow time collected in simulation is much greater than the corresponding objective value of objective (3.1) in Table 5.9 for the one-week MIP problem.

Likewise, as can be seen from the results of *one-week problem*, it also suggests that the overall flow time in simulation includes both the pure traveling time and the cargo processing time at the facilities. And furthermore, it implies that the cargo processing time accounts for a relatively large portion in the overall flow time due to the congestion.

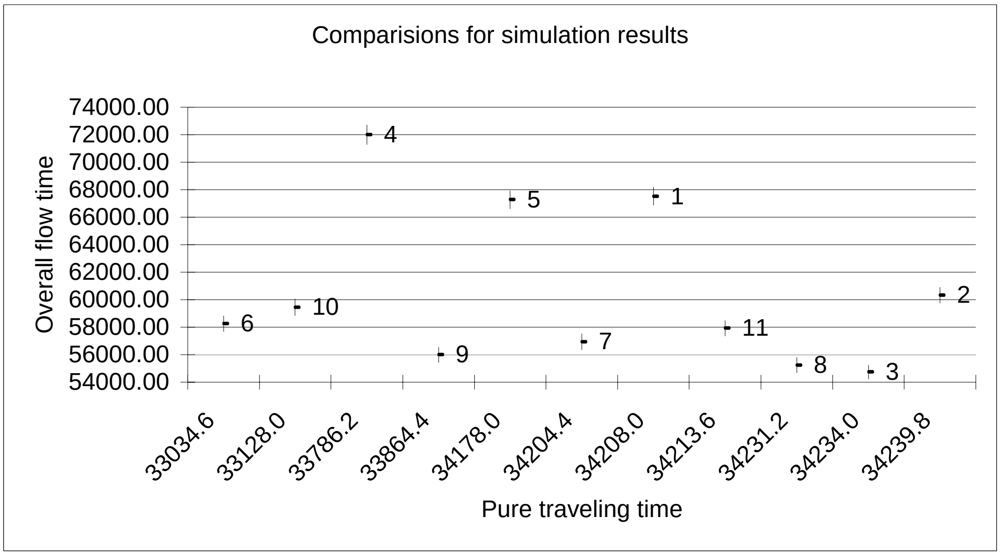

Figure 5.2 Comparisons for simulation results (one-week)

The most desirable solution design #3 provides the smallest overall flow time. Since the overall flow time in simulation considers the possible congestion, this design in Figure 5.2 with the shortest overall flow time in simulation is not the one with the shortest pure traveling time in MIP optimization results. However, this design does have the smallest values of Obj (3.2) and Obj (3.4), and the second smallest values of Obj (3.3) and Obj (3.5) in optimization results. Thus, this is the best design in simulation, which goes in line with our observation from the results for one-day problem.

As suggested in Section 5.2.3, the MIP provides the candidates pool for the evaluation by simulation for the one-week problem. Without the selection pool suggested by MIP solution, it would be extremely hard to find a desirable design for the real problem, as the potential designs are too numerous if without the solutions from MIP. Thus the efficient solutions produced by MIP are proved to be useful for the final evaluation by simulation, since they save the computational efforts for the time-consuming simulation.

This best efficient solution demonstrates its overall flow time in simulation of 54733 minutes, which is a 23.1% improvement from 71185 minutes of the original assignment plan for one-week problem.

5.4 Conclusion

As it is recognized that the flight schedule duplicates itself every week, it would be good to address only the one-week problem for tactical or strategic decision-making. Nevertheless, the results from the one-day assignment could also provide an alternative from the perspective of operational planning for the one-week planning. It is a common practice that the flights with the identical flight number should have the same assignment no matter on which day of a week. Therefore, the weekly problem provides a “stronger” perspective for the flights assignment than the daily one.

In the weekly problem, if considering all the possible assignment of the flight to ramp zone and the flight to workstation area combinations, there will be a total of $(3 * 4)^{151}$ $= 12^{151} = 9.044e+162$ possible assignments pending for evaluation, for this one-week problem. Therefore, the selection of 11 different efficient solutions by MIP out of the entire potential selection pool is a huge reduction of possible computational efforts for further evaluations of the assignment designs. Similarly, the same rationale also applies to the daily problem.

As suggested in the comments for the simulation results, the MIP provides the candidates pool for the simulation. As discussed, the potential designs are too numerous if without the efficient solutions from MIP, therefore the computational time is extremely reduced by the trimmed selection pool suggested by the MIP solutions. Hence, there are a total of 13 distinguished solutions selected from MIP for the one-day simulation, and 11 distinguished solutions from MIP for the one-week simulation. These efficient solutions from MIP are able to save the computational efforts for the final evaluation by the simulation, since they bring fewer candidates as the input of efficient designs for the simulation.

It is concluded that, the weekly problem is relatively harder to reach optimum than the daily problem in the MIP stage. In this paragraph, the different settings of tolerance levels for these two problems are investigated. A more comprehensive display of the comparisons between the one-day problem and one-week problem is given in Table 5.10. The first two columns show the time horizons and time intervals for both problems. The third column illustrates the total number of constraints, while the fourth and fifth columns demonstrate the total number of flights and total number of binary variables. The last column gives their tolerance levels. It can be seen that, although the one-day problem has more constraints than the one-week problem, it has fewer binary variables and thus the tolerance level for optimization is set more stringent according to the hardness of the problems.

Table 5.10 The comparison between one-day and one-week problems

	Time horizon / Time interval	# Time intervals	# Total constraints	# Flights to assign	#Assignment variables	Tolerance level
One-day problem	1 day / 5 min	288	9, 729	63	756	0.0005
One-week problem	1 week / 1 hour	168	5, 527	151	1, 812	0.005

In addition, it is reasonable to see such a significant deviation between the overall flow time in the simulation and the pure traveling time in objective (3.1) in the MIP optimization results, since the overall flow time in the simulation contains both the traveling time and the cargo processing time at the facilities, with the possible queuing effects.

Overall, the MIP solutions provide a series of non-dominated solutions for this multiple objective problem. They are all efficient plans for the weekly freight handling assignment planning. Thus, the operation efficiency in the air cargo terminal could be further improved by applying some of these new assignment plans.

Furthermore, the simulation results for the performance of the weekly plans are collected. In this regard, the simulation could provide a much more thorough measurement of the system performance than the MIP model.

Hence, by applying the optimization methodology, the decision maker will firstly assign the incoming flights according to their distinguished flight numbers to different ramp zones, and then to different breakbulk workstation areas, based on the best design we choose. Consequently, the re-design of the cargo dispatching plan in this air cargo terminal is finished.

6 Conclusions and Future Research

The chapter concludes this research study on an air cargo inbound terminal. Conclusions are made in this chapter from an overall perspective. Furthermore, possible future research studies about this problem are projected.

6.1 Conclusions

In general, this research work involves with the tactical planning issue of an international air cargo terminal which serves for a regional air transportation hub. In this study of the air freight terminal, the main problem is addressed based on piloting survey and data analysis. Then, the mathematical modeling technique is applied to resolve the problem using a multi-objective mixed-integer programming model. The stochastic simulation method is later employed to further identify the best decision choice. Finally, the recommendation given the various objectives and constraints is presented based on the results and outputs. More details about these findings are given in the following paragraphs.

First, based on the data analysis and on-site observations, our pilot study shows that the current inefficiencies such as the cargo flow time and so on are mainly caused by the current weekly assignment of the flight to the facility, in which flights are assigned to various ramp zones, and breakbulk workstation areas according to their flight numbers. Hence, a hybrid approach which involves the cooperation of the mathematical model and the simulation model is suggested.

Second, the problem is hard to formulate as a precise mathematical model without any approximations. A mixed-integer programming model is applied to formulate this assignment problem with certain approximations. In order to account in the different concerns of the problem, it is revised as a multiple objective mixed-integer programming model. A simple deterministic model to estimate the timing and size of the cargo handling process is also proposed and adopted to calculate the coefficients for the mixed-integer programming model. Following the ε-optimal approach, both the short-term (daily) and long-term (weekly) planning problems are tackled by this model. With the help of CPLEX®, a series of non-dominated solutions are generated, and thus these solutions will be utilized for the inputs of the simulation in the next stage.

Third, to address the complexity of the system and the stochasticity of the cargo movement patterns, a simulation model for the system is then developed. This simulation concentrates on the ULD (Unit Load Device) movement within the inbound cargo terminal, and tries to evaluate the performances of the available designs. The simulation model also attempts to approximate the transferring vehicle and load dispatching rules within the PCHS (Pallet Container Holding System) and the workforce deployment procedures. This simulation model is developed, verified, and validated. The available assignment plans extracted from the optimization solutions are examined in the simulation. The performance measure, i.e., the overall flow time for all the cargos is obtained for each efficient assignment plan. These assignment plans are evaluated with appropriate simulation experimental designs. The statistics of the performance measurements of are presented for further decision-making.

6.2 Future Research

Due to some practical constraints for the research study, there exist several aspects for future improvements on this research work.

First, it could be more enlightening if the optimization for the MIP formulation of the problem could produce more efficient solutions. Due to the existing restrictions, there are a total of 36 designs generated from the efficient solutions. More efficient solutions could be provided in the future given the advent of technology.

Second, such an optimization along with simulation approach would be more reliable if the scope of the simulation model could cover the entire cargo terminal. In doing so, the cargo flow time would be better captured since the entire life cycle of the cargo processing in the cargo terminal will be described in the simulation model.

Third, thinking from an overall perspective, it is possible to address the manpower planning problem in this study as well. The checking and breakbulking workforce is an indispensable part of the terminal operations. Thus, by accounting in more dimensions about the terminal management, the research results will be more consistent and useful.

Last, the terminal currently operates on a priority processing basis. These priorities are determined principally by the time sequences of the jobs; however, it is possible to override these priority settings by human interventions. The current research does not account in these complexities of priorities and human interventions. Therefore, it is also beneficial to project the future research on the priority based terminal operations.

This priority based approach in conjunction with the previous optimization plus simulation modeling approach, would develop a more intelligent and customer-friendly decision support systems for the decision makers.

Bibliography

Amiouny, S. V., Bartholdi, J. J., III., Vate, J. H. V., Zhang, J., 1992, *Operations Research*, 40(2), 238-246

Bartholdi, J. J., III, Gue, K. R., 2000, Reducing labor costs in an LTL corssdocking terminal, *Operations Research*, 48(6), 823-832

Berrada, M., Stecke, K. E., 1986, A branch and bound approach for machine load balancing in flexible manufacturing systems, *Management Science*, 32(10), 1316-1335

Bish, E. K., 2003, A multiple-crane-constrained scheduling problem in a container terminal, *European Journal of Operational Research*, 144, 83-107

Co, C. G., Tanchoco, J. M. A., 1991, A review of research on AGVS vehicle management, *Engineering Cost and Production Economics*, 21, 25-42

Gambardella, L. M., Rizzoli, A. E., Zaffalon, M., 1998, Simulation and planning of an intermodal container terminal, Special issue SIMULATION on harbor and maritime simulation

Gue, K. R., 1999, The effects of trailer scheduling on the layout of freight terminal, *Transportation Science*, 33(4), 419-428

Houghton, E., Portougal, V., 1997, Trade-offs in JIT production planning for multi-stage systems: balancing work-load variations and WIP inventories, *International Transactions on Operational Research*, 4(5/6), 315-32

Khouja, M., Conrad, R., 1995, Balancing the assignment of customer groups among employees: zero-one goal programming and heuristic approaches, *International Journal of Operations & Production Management*, 15(3), 76-85

Kim, K. H., Park, K. T., 2003, A note on a dynamic space-allocation method for outbound containers, *European Journal of Operational Research*, 148, 92-101

King, R. E., Wilson, C. M., 1991, A review of AGV system design and scheduling, *Production Planning and Control*, 2, 44-51

Law, A. M., Kelton, W. D., 2000, Simulation Modeling and Analysis, McGraw-Hill

Marco, J. G., Salmi, R. E., 2002, A simulation tool to determine warehouse efficiencies and storage allocations, Proceedings of the 2002 winter simulation conference

Nozick, L. K. , Morlok, E. K., 1997, A Model For Medium-Term Operations Planning in An Intermodal Rail-Truck Terminal, *Transportation Research-A*, 31(2), 91-107

Preston, P., Kozan, E., 2001, An approach to determine storage locations of containers at seaport terminal, *Computers & Operations Research*, 28, 983-995

Sawik, T., 2002, Monolithic vs. hierarchical balancing and scheduling of a flexible assembly line, *European Journal of Operational Research*, 143, 11-124

Taniguchi, E., Noritake, M., Yamada, T., Izumitani, T., 1999, Optimal size and location planning of public logistics terminals, *Transportation Research-E*, 35, 207-222

Tsui, L. Y., Chang, C. H., 1990, A microcomputer based decision support tool for assigning dock doors in freight yards, *Computer & Industrial Engineering*, 19, 309-312

Tsui, L. Y., Chang, C. H., 1992, An optimal solution to a dock door assignment problem, *Computer & Industrial Engineering*, 23, 283-286

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

Wilson, J. M., 1992, Approaches to machine load balancing in flexible manufacturing systems, *Journal of Operational Research Society*, 43(5), 415-423

Yun, W. Y., Choi, Y. S., 1999, A simulation model for container-terminal operation analysis using an object-oriented approach, *International Journal of Production Research*, 59, 221-230