A STUDY ON LOCATION SET COVERING MODELS WITH TWO APPLICATIONS

ZHOU YUAN

(M.Eng., Huazhong University of Science and Technology)

Supervisor: Dr. HUANG BORAY

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Department of Industrial and Systems Engineering
National University of Singapore

2014

Declaration

I hereby declare that the thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

August 11, 2015

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincerest gratitude to my supervisor, Dr. Huang Boray, for his valuable guidance and continuous help during my research career. His amiable personality, passion for research, creative ideas and studiousness make him a respected advisor and always motivate me to explore the truth in the world. Working with him is a pleasant experience which will also be a great fortune in my whole life.

I also want to thank Dr. Hung Hui-Chih for his guidance during my first one and a half years of study in National University of Singapore. Moreover, his efforts and contributions on this dissertation are highly appreciated as well. Besides, I particularly want to thank Prof. Gong Dah-Chuan for his help on my second research topic and his encouragement during my pursuit too. Special thanks are also given to A/Prof. Ng Kien Ming and A/Prof. Ng Tsan Sheng, Adam for their valuable comments and help on this thesis.

I am also very grateful to our reading group members, Chen Liangpeng, Jing Huayi, Jing Lei, and Pourghaderi Ahmad Reza, for their inspiring suggestions on my research as well as their encouragement during the hardest times in my doctoral pursuit. Moreover, I have to thank some undergraduate students in National University of Singapore, Zhou Fang, Zhang Zhongcheng, Tan Wee Kiat Kenneth and Chen Shuyu for sharing their ideas on my research.

I also have to thank the department of Industrial and Systems Engineering in National University of Singapore for offering me the research scholarship, NUS libraries for providing abundant resources, without which this thesis would never have been finished. Gratitude is also given to all other faculty members and staffs in the department of Industrial and Systems Engineering, especially the members in former Simulation Lab, for their help and supports over these years.

Last but not least, I would like to thank my families, especially my husband Zhang Sizhe and my daughter Zhang Jinyu for their love, support and encouragement. I dedicate this thesis to them, without whom this work would never have been finished.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii

TABLE OF CONTENTS v

SUMMARY ix

LIST OF TABLES xiii

LIST OF FIGURES xv

LIST OF ABBREVIATIONS xvi

1 INTRODUCTION 1

1.1 Application in EMS systems 3

1.2 Application in GrSC network design 9

1.3 Structure of the thesis 13

2 LITERATURE REVIEW 17

2.1 Review on the LSC models and the availability issue 18

2.2 Review on the MCL models and the availability issue 26

2.3 Review on the GrSC network design 31

2.4 Review on the carbon tariff mechanism 35

3 PROBLEM STATEMENT FOR THE MODEL WITH DEMAND ASSIGNMENT 40

3.1 The failure probability with assignment and its upper bound 40

3.1.1 The failure probability with assignment 41

3.1.2 An upper bound of the failure probability with assignment 45

3.2 The advantages of assignment 47

3.2.1 Two station case 47

3.2.2 Three station case 49

3.2.3 Comparison between FP^Q and FP^{AU} 52

3.3 The traditional failure probability and its estimates 55

3.3.1 Two station case: one vehicle located at each station 57

3.3.2 Two station case: two vehicles located at each station 60

3.3.3 Three station case: one vehicle located at each station 60

3.3.4 Comparison of the traditional failure probability and its estimates 62

3.4 Summary 68

4 MATHEMATICAL MODEL WITH DEMAND ASSIGNMENT 69

4.1 Mathematical model 69

4.2 Linearization transformation 71

4.3 Heuristic 72

4.4 Summary 79

5 EXPERIMENTAL RESULTS AND DISCUSSIONS ON THE T-LSCPA 81

5.1 Case study 1: 55-node problem 81

5.2 Case study 2: real case study for a known city S 92
5.3 Summary and discussions 94

6 PROBLEM STATEMENT FOR THE GrSC NETWORK DESIGN 97

6.1 Main issues in the GrSC network design 97
6.2 Relationship of modules and options 99
6.3 Technology selection and transportation mode selection . . . 100
6.4 Description of how carbon tariffs are imposed 101
6.5 Summary 106

7 MODEL CONSTRUCTION FOR THE GrSC NETWORK DESIGN 107

7.1 List of notations 107
7.2 Mathematical model 111
7.3 Model analysis 116
7.3.1 Transformation to a pure linear model 117
7.3.2 The impacts of carbon tariffs on technology selection 119
7.3.3 The impacts of carbon tariffs on total emissions . . . 123
7.4 Summary 123

8 EXPERIMENTAL RESULTS AND DISCUSSIONS ON THE MMGSC 125

8.1 Numerical experimental study and results 125
8.1.1 The impacts of carbon tariff on production quantity decision and technology selection 127
8.1.2 The impacts on total costs and emissions 133

8.1.3 The impacts of supply range on the total costs with and without carbon tariffs 135
8.1.4 Factory's willingness to join the member society . . . 136
8.2 Summary and discussions 138

9 CONCLUSIONS AND FUTURE RESEARCH 140

9.1 Summary of research results 140
9.2 Limitations of this study and future research 144

BIBLIOGRAPHY 146

APPENDICES 159

A Proof of Lemma 1 159
B Proof of Proposition 1 160
C Proof of Proposition 2 161
D Transition Rate Matrix for the two station case: $z = 2$, shown in Table D.1 163
E Transition Rate Matrix for the three station case 163
E.1 3 station case, $z = 1$, $o = 0$, shown in Table E.1 . . . 163
E.2 3 station case, $z = 1$, $0 < o \leq 0.346$, shown in Table E.2 163
E.3 3 station case, $z = 1$, $0.346 < o \leq 0.973$, shown in Table E.3 163
F Function "find the threshold workload" 168
G Proof of the equivalence of Model *T-LSCPA* and *LSCPA* . . 169
H The 55-node problem, shown in Table H.1 171
I Modification to address infeasibility 173

SUMMARY

Facility location problems concern the optimal placement of facilities. Two related types of models are widely studied to solve the facility location problems. The first one is the Maximal Covering Location (MCL) models, which maximize the demand covered with a stated service level, while the fleet size is fixed. The other type is called the Location Set Covering (LSC) models, which seek the least number of facilities or total costs such that some service requirements can be achieved. This thesis provides a review on both types of models and focuses on two applications of LSC models in two typical practical fields: Emergency Medical Service (EMS) systems and Green Supply Chain (GrSC) systems, in order to provide insights for the policy-makers or managers. In particular, the insights obtained from study of Part I are applied in Part II.

Part I of this study is motivated to apply LSC models to EMS systems with a goal to minimize the overall costs under a service requirement, such as determining the number of ambulances required in Singapore where the population is rapidly growing and ageing. To meet such requirement, the assessment of server workload and availability is often a major concern. It is interesting to note that most existing LSC models do not assign demand to servers or stations. In other words, all the demand in the coverage of each service station count in the workload of the station, without consideration

of the overlaps. Even if demand assignment is applied in some articles, few of them investigate the benefits of doing so. Our study in part I fill in this gap by comparing the failure probability as well as fleet size required with and without demand assignment. And the results show that it is beneficial to pre-assign demand to servers or stations in terms of saving costs and meeting service requirements.

Part II of the thesis studies another application of LSC models in GrSC design. With the growing awareness of global warming, many companies worldwide are improving their supply chain sustainability, under pressure from the government or their own shareholders. However, not all countries around the world place equal emphasis on this. For example, some countries still have not implemented any carbon dioxide emission regulations to address this problem. Carbon regulations in only sub-global areas may result in even higher global emissions because of carbon leakage. A possible approach to cope with the carbon leakage problem is to impose carbon tariffs on the goods from unregulated countries. In order for supply chain design, lateral transshipment is widely adopted by many researchers. However, inspired by the insights obtained from the study in Part I, we apply similar pre-assignment into the model by introducing a series of decision variables to connect the three-tier facilities, in order to build our model. Furthermore, we propose a tricky way to model the problem as a mixed integer program to make it can be solved by standard methods. The managerial insights from a real case study in Part II of this thesis also help to shed light on those who want to study the impacts of carbon tariff in practice.

Keywords: emergency medical service; location set covering; demand allocation; workload overestimation; green supply chain design; carbon leakage; carbon tariff

LIST OF TABLES

Table 3.1 The transition rate matrix (2 stations, $z = 1$) 58

Table 5.1 The computation time for the 27 instances 82

Table 5.2 Cost saving compared with the QRLSCP 85

Table 5.3 The basis solutions ($p_1 = 90\%$, β=1) 91

Table 5.4 Solution sensitivity analysis to p_1 91

Table 5.5 Solution sensitivity analysis to $\lambda_j s_{ij}$ 91

Table 5.6 Comparison under the real case study 93

Table 6.1 Relationship of modules and options of a certain product 100

Table 8.1 Sales distribution of Type I notebook 126

Table 8.2 The network under analysis 127

Table 8.3 Technology level selection under different carbon tariff rates 128

Table 8.4 The suppliers' capacities 135

Table D.1 The transition rate matrix (2 stations, $z = 2$) 164

Table E.1 The transition rate matrix (3 stations, $z = 1$, $o = 0$) . 165

Table E.2 The transition rate matrix (3 stations, $z = 1$, $0 < o \leq 0.346$) 166

Table E.3 The transition rate matrix (3 stations, $z = 1$, $0.346 < o \leq 0.973$ 167

Table H.1 The detailed information of the 55-node problem . . . 171

LIST OF FIGURES

Figure 1.1 Structure of the thesis 14

Figure 3.1 Relation between the workload and the required number of vehicles for different failure probabilities in loss systems 44

Figure 3.2 The example of the two station case 48

Figure 3.3 The three examples for the three station case 51

Figure 3.4 Comparison between FP^{Q} and FP^{AU} in the two cases 54

Figure 3.5 Switch-over curves for FP^{AU} and FP^{Q} 56

Figure 3.6 Comparison of fp and its estimates (2 station case, $z = 1$) 63

Figure 3.7 Comparison of fp and its estimates (2 station case, $z = 2$) 64

Figure 3.8 Comparison of fp and its estimates (3 station case, $z = 1$) 65

Figure 3.9 Worst case of $fp - FP^{AU}$ when $o \to 1$ 67

Figure 6.1 The three-tier supply chain network 103

Figure 6.2 The multi-tier supply chain network 106

Figure 7.1 The graphs of $NB(t)$ 121

Figure 8.1 The flow of products in the network 132

Figure 8.2 Minimum production rate to operate a factory 133

Figure 8.3 Total costs and emissions for different values of γ . . 134

Figure 8.4 Cost reduction for different supply ranges 136

Figure 8.5 Threshold cost that Company G's factory in China prefers to join the member society 138

LIST OF ABBREVIATIONS

MCL	Maximal Covering Location
LSC	Location Set Covering
EMS	Emergency Medical Service
GrSC	Green Supply Chain
ED	Emergency Department
DC	Distribution Centre
KPI	Key Performance Indicator
EU	European Union
DLSCP	Deterministic Location Set Covering Problem
PLSCP	Probabilistic Location Set Covering Problem
BPLSCP	Binomial Probabilistic Location Set Covering Problem
PRLSCP	Poisson Reliability Location Set Covering Problem
QRLSCP	Queueing Reliability Location Set Covering Problem
BRLSCP	Binomial Reliability Location Set Covering Problem
MEXCLP	Maximal Expected Covering Location Problem
MALP	Maximum Availability Location Problem
PASTA	Poisson Arrivals See Time Average
CTMC	Continuous Time Markov Chain
LSCPA	Location Set Covering Problem with Assignment
MINLP	Mixed Integer and Non Linear Program

CDF	Cumulative Distribution Function
MMGSC	Mathematical Model for the Green Supply Chain

Chapter 1 INTRODUCTION

Facility location problems usually produce solutions concerning the optimal placement of facilities by solving mathematical models. The location models are widely used in public service and industry fields. The traditional applications include locating the ambulances or determining the fleet size of the Emergency Medical Service (EMS) vehicles in health care systems (see, Ball and Lin 1993, Daskin and Dean 2005, Beraldi and Bruni 2009); positioning the fire company and ladder trucks for the Fire department (see, Walker 1974, Plane and Hendrick 1977, Schreuder 1981); siting public schools, bus stops or recycling centers for the government (see, Daskin 1995, Gleason 1975, Ye et al. 2011); guiding the practitioners in the supply chain to make decisions regarding the site selection, such as Distribution Centers (DCs), manufacturing factories and storage facilities (see, Shen et al. 2003, Chaabane et al. 2012, Hwang 2004). Besides, Berman et al. (2009) introduce other applications of the facility location models in installation of warning sirens, cell phone towers, light towers and outdoor gas heaters. For LSC models one can be refer to Schilling et al. (1993) and Brotcorne et al. (2003) for excellent reviews.

Two related types of models are studied to solve facility location problems. The first one is the Maximal Covering Location (MCL) models. These models maximize the demand covered within a stated service level,

while the ambulance fleet size is fixed. Usually, the MCL models assume that the fixed number of facilities is insufficient to cover all demand within the stated service level, due to the budget limitation. Otherwise, the problem would become trivial. The other type of facility location models is the Location Set Covering (LSC) models. The earlier LSC models seek the least number of facilities or overall costs with a constraint that all demand can be covered by at least one facility within a stated service level. Later, the researchers take the *availability* issue into account. That is, once a vehicle is dispatched, it cannot respond to the incoming requests until the current service has been finished. The constraint in the earlier LSC models becomes to a probabilistic one: the probability that a node can be immediately responded by an idle vehicle in its neighbourhood should exceed some stated level. The two types of models above-mentioned can be applied in different situations. The LSC models are more appropriate when the overall demand in a region changes a lot; the demand pattern significantly fluctuates over a day of the week or time of the day; or when a flexible capacity is to be decided.

This thesis mainly discusses the applications of the LSC models in two fields: EMS systems and Green Supply Chain (GrSC) network design. The insights obtained from study of Part I are applied in Part II. The model for the former application focuses on exploring the advantages and disadvantages of demand assignment. It is interesting to note that most existing LSC models do not assign demand to servers or stations. Even if demand assignment is applied in some articles (see, Shen et al. 2003), few of them investigate the effects of doing so. Our study in part I fill in this gap by

comparing the failure probability as well as fleet size required with and without demand assignment. And the results show that it is beneficial to pre-assign demand to servers or stations in terms of saving costs and meeting service requirements. The insights obtained from study of this part are applied in Part II.

Part II of this thesis mainly examines the impacts of carbon tariffs, an adjustment measure at the border in the context of climate policy, which will be further discussed in Section 1.2, on the facility location, technology selection, and production quantity decisions. In order for supply chain design, lateral transshipment is widely adopted by many researchers (see, Wee and Dada 2005). However, inspired by the insights obtained from the study in Part I, we apply similar pre-assignment into the model by introducing a series of decision variables to connect the three-tier facilities. In terms of modelling, we offer a novel way to incorporate carbon tariff in practical supply chain design and a tricky way to formulate the problem as a mixed integer program. The program then can be solved using standard methods. The managerial insights in this part also help to shed light on those who want to study the impacts of carbon tariff in practice.

1.1. Application in EMS systems

The EMS systems deal with a large number of patients every year, and the ambulance location problem is one of the critical issues. It is reported that about 114 million visits to Emergency Departments (EDs) occurred every year in the United States, and 16 millions of these patients are responded by ambulances (see, Institute of Medicine of the National Academies 2006).

They also state that the response time within which the ambulances arrive at the emergency calls is highly variable, and this variability has much to do with the geography. Thus, how to efficiently and effectively determine the geographical deployment of ambulances plays an important role in EMS systems. One of the major goals in the ambulance location problem is to achieve a balance between an adequate service level and the reduction of the total costs.

To tackle this issue, a variety of Location Set Covering (LSC) models are proposed with a goal to minimize the overall cost under a service requirement. Compared with other location problems like Maximal Covering Location (MCL) models where a fixed number of vehicles are given, one unique feature of LSC problems is the optimization of the ambulance fleet size, which is especially applicable to many cities with rapid growth or ageing. For example, the population is growing and ageing in Singapore quickly in the present years. As reported by Lai (2012), the population in Singapore reaches 5.08 million in June, 2010, reflecting a growth by about 25% within the last 10 years. The aging tendency in Singapore is also serious as well. It is reported that the median age of the Singapore residents rises from 34.0 in 2000 to 37.4 in 2010 (see, Lai 2012). The growing and aging population in Singapore results in a higher demand pattern for the ambulance service, which brings a challenge to its EMS system. Thus, how to locate the ambulance stations and redetermine a suitable fleet size is a critical issue for the Singapore EMS system. On the other hand, Ong et al. (2009) state that the emergency call demand pattern significantly fluctuates over the day of the week or the time of the day in Singapore, and this is

also the case in many areas (see, Rajagopalan et al. 2008, Ong et al. 2009). This fluctuation enforces EMS systems to deploy flexible ambulance capacities for different demand patterns. For example, the systems can maintain a fixed number of facilities and hire additional private facilities for higher demand pattern. In this case, the LSC models also show their advantage in deciding the ambulance deployment and fleet size for each time period when a specific service level is to be guaranteed. As the fleet is variable, the LSC models can help to assign the resource or staff for different time periods as well.

To meet the service requirement, an important step in the LSC problem is to assess the *availability* of the vehicles or ambulances. Once a vehicle is dispatched, it cannot respond to the incoming requests (i.e., the vehicle is *unavailable*) until the current service is finished. A common definition of availability is the probability that each demand node $j \in J$ can be immediately responded by an idle vehicle in the neighbourhood $S_{COV}(j)$. $S_{COV}(j)$ is determined by either the maximal travel distance or the maximal response time. This availability is also referred to as the *reliability* (see, ReVelle and Hogan 1988, Revelle and Hogan 1989, Borrás and Pastor 2002), which is interchangeably used with the availability in our presentation. Ball and Lin (1993) consider the *failure probability*, which equals to 1 − availability. The availability or the failure probability is a common requirement for EMS systems in practice. For example, it is required in North America that the EMS vehicles should respond to 90% of all the highest priority calls within 8 min, and 90% of all calls within 9 min and 15 min for the urban and rural areas, respectively (see, De Maio et al.

2003, Fitch 2005). That means, 90% of the calls have to be responded by available ambulances within the required time.

For the following discussion, we first define *traditional failure probability* as the probability that a demand cannot immediately find any idle vehicle in its neighbourhood, under the assumption that any incoming demand would be served by the closest available vehicle and lost if all stations in its neighbourhood run out of vehicles. The traditional failure probability is the performance target for the practitioners and also widely numerically investigated by researchers (see, Ball and Lin 1993, Borrás and Pastor 2002, Baron et al. 2009).

The availability, however, is quite difficult to assess. Some early works do not touch this issue in the location problem (see, Toregas et al. 1971, Aly and White 1978). They just simply assume that the vehicles are always available as long as the demand is in the coverage area. Subsequent models start to address the availability issue by making assumptive workload for each vehicle or station, which are referred to in Section 2.1. The reason is that, in order to obtain the probability that a vehicle is available, the workload for each vehicle or station has to be known in advance. However, it is very difficult to find the actual workload and the availability for each vehicle or station without a specific vehicle dispatching or demand allocation policy (e.g., closest available vehicle first.)

Interestingly, we notice that most of the location models in the existing literature do not assign each demand to a specific station (see, Toregas et al. 1971, Chapman and White 1974, Aly and White 1978, Ball and Lin 1993, Borrás and Pastor 2002). The location models without assignment

are also prevailing in the MCL problems as well (see, Church and ReVelle 1974, Daskin 1983, Berman and Krass 2002). Besides, Restrepo et al. (2009) minimize the lost demand with a fixed fleet size in the absence of the assignment. There may be two intuitive reasons not to allocate demand in the models. Firstly, the analytical model may become quite large with assignment due to the increased number of decision variables. Secondly, as each demand node can be served by multiple stations, it appears that fewer vehicles are required to meet the availability requirement due to the risk pooling effect. Similar arguments can be found in inventory literatures, where lateral transshipment is adopted to pool inventory and save cost (see, Wee and Dada 2005, Sošić 2006, Yang and Qin 2007, Paterson et al. 2011).

Without assignment, as mentioned earlier, the workload for each station would be difficult to estimate, as each demand node can be served by any surrounding vehicle. This makes assessment of the availability a tough task. Some models (see, Ball and Lin 1993, Borrás and Pastor 2002) approximate the workload for each station as the sum of all workload in the coverage of the station. But such approximation would inevitably overestimate the actual workload for each station, particularly when the workload through the neighbourhoods or regions is highly overlapped. For example, the urban area is usually highly overlapped as the demand in this area can be covered by multiple stations simultaneously.

A possible solution to the workload overestimation problem is the assignment (allocation) of demand. For example, Marianov and Serra (1998, 2002) study a location-allocation problem by assuming the demand allocated to a particular station would queue at the station. The service time fol-

lows an exponential distribution depending only on the stations. Toro-Díaz et al. (2013) also propose a model incorporating location and dispatching decisions, where a given number of vehicles are dispatched according to a pre-determined demand-server preference list. The model is quite complicated to solve as it is an NP-hard problem. With a specific dispatching policy (server preference list), Budge et al. (2009) propose an algorithm to approximate the actual dispatch probabilities. Other research on vehicle dispatching under a fixed fleet size includes Mayorga et al. (2013) and McLay and Mayorga (2013). Different from these literatures which usually take the allocation or dispatching as a policy to be optimized, Part I of this thesis investigates the potential benefit of demand assignment. We assume generally-distributed service times which depend on *both* demand node and the station allocated to the node. It matches the practice as a stationed vehicle usually takes different times to serve different demand nodes. A novel availability constraint and heuristic are also proposed, which yields good results even when the assignment policy is not strictly followed.

There are some two-stage stochastic programming models proposed to tackle the availability issue (see, Beraldi and Bruni 2009, Snyder 2006), where the demand is assigned for each scenario in the second stage, after locating ambulances in the first stage. However, the assignment policy is hard to implemented in practice as it depends on specific scenarios, which are usually uncertain in advance. We also notice that some location-inventory models conduct demand assignment. For example, Shen et al. (2003) introduce the assignment variables into their model, to assign the retailers to the Distribution Centres (DCs) instead of considering all retail-

ers in the coverage of each DC. In their paper, the optimal inventory level at each DC cannot be identified without decisions of demand assignment.

To summarize, the motivation of the first study is to explore the advantages and disadvantages of locating the ambulances and assigning the demand at the same time. The discussion in this is not only beneficial to the researchers in location research field, but also provides justifications for the location-inventory research work.

1.2. Application in GrSC network design

How to locate the assembling factories and decide the production quantity of each factory has been studied for a long time. However, with the growing attention paid to the global warming crisis during the recent decade, many companies face challenges to re-design their supply chain network under different carbon reduction policies, e.g., cap-and-trade; carbon tax; carbon price; etc. This is also called "Green Supply Chain network design" (see, Beamon 1999, Srivastava 2007, Wang et al. 2011). This means that the managers are more concerned about the environmental impacts when they are designing the supply chain, instead of only maximizing the economic profits. Part II of this thesis is to examine the impact of one policy, carbon tariff, on the network design decisions. Conceptually, carbon tariff is the tax placed on the goods that are imported from unregulated countries or regions, i.e., the countries or regions which have not adopted feasible carbon regulations, to limit their carbon emissions in production.

Firstly, we explain why the carbon tariff is required in the context of climate policy. There is growing consensus that global warming, if unchecked,

will bring a grave threat to the world. The increasing carbon dioxide (CO2) emissions are responsible for global warming (see, Solomon 2007). It is estimated by Enkvist et al. (2007) that CO2 emissions are about 40 billion tons in 2002, and this number is expected to expand to 58 billion tons by 2030. In order to mitigate explosively increasing carbon emissions, many countries are taking initiatives to curb the total amount of their CO2 emissions. For example, the European Union Emission Trading System (EU-ETS) proposes a "cap-and-trade" policy to manage CO2 emissions. According to this policy, each company is issued certain credits to emit CO2 based on its own operation scale, industry and so on. The companies should adjust their operations and reduce their emissions to meet the "cap". If the emissions of a company are below its cap, it has extra credits to be traded with other companies or in the market. On the other hand, if a company emits more CO2 than the cap, it should pay for these excessive emissions. China also sets a clear target to reduce the CO2 emissions by 10% in its 11th five year developing plan (see, Wang et al. 2011), as well. In addition, the Australian government reports that $1 billion in funds are spent by manufacturing companies on improving energy efficiency and reducing pollution. Moreover, over 220 clean technology projects have been launched at manufacturing plants around the country (see, Australian Government 2013).

As a result of the regulations from the governments and increasing concerns from the shareholders, companies around the world are also devoting great efforts to reduce their carbon footprints (see, Benjaafar et al. 2013). Sustainable supply chain network design becomes a new challenge

to many companies, either locally or internationally. They are trying to design greener products by choosing low emission raw materials; investing on more environment-friendly production technologies; minimizing CO2 emissions in transportation, and so forth. For example, Wang et al. (2011) state that the global procurement center of IBM, which is located in China, adds some environment-related indices in its Key Performance Indicators (KPIs). Our research is also motivated by a real case problem from Company G, a world class notebook company. The managers in Company G are interested in achieving a balance between their business and the environment under carbon reduction requirements.

However, the outcomes of carbon regulations may not be as good as the prediction. Not all countries or regions around the world are equivalently active in taking initiatives to combat climate change, which is called *unilateral* (or *sub-global*) carbon control policies. Drake (2011) points out that although carbon legislations are intended to reduce total emissions, the outcomes may just shift production to unregulated countries or regions (For simplicity, those countries or regions adopting carbon regulations, e.g., cap-and-trade, carbon tax, carbon price, etc., are referred to as *member society*; and those without such regulations are referred to as *non-member society* in this thesis). There are two reasons for these outcomes. Firstly, the companies based in the non-member countries would be more competitive due to the cost advantage. Secondly, the companies in the member countries have incentives to shift their plants to the non-member countries to enjoy the cost advantage. For example, as indicated in Drake (2011), the cost advantage in the non-member countries may almost double after

the carbon regulations are put into practice, making domestic production no longer a possible choice.

Shifting production to non-member countries or regions cannot curb the global CO2 emissions. This would result in "*carbon leakage*": an increase of CO2 emissions of one country as a result of the carbon reduction in another country where a strict carbon regulation is in place. Furthermore, without carbon regulation in non-member society, there will be no incentive for those companies to reduce emissions. Hence, the carbon regulations may potentially increase the overall CO2 emissions. On the other hand, as stated above, such unilateral control policies would endanger the domestic industries of member countries due to the cost advantages in unregulated regions. This is the reason why the European Union (EU) claims that it will take "appropriate measures" in regard to the rest of the world which does not match Europe's carbon control standard in the EU summit agreement in 2008 (see, Corcoran 2008).

One possible approach for the member countries to address the carbon leakage problem is to impose *carbon tariff* on the goods flowing from the non-member countries. As stated above, carbon tariff is the tax placed on the goods that are imported from non-member countries to limit their CO2 emissions in production. Proponents of a carbon tariff policy believe that such a method can protect domestic industries in the member countries, motivate the non-member countries to legislate carbon policies and ultimately curb global emissions. On the other hand, there are also some opponents of carbon tariff mechanism who doubt the effectiveness of such measures.

Part II of this thesis focuses on exploring the impacts of carbon tariff imposition on the global supply chain, and answers the questions of concern to the managers from Company G. In particular, after the carbon tariffs are imposed, will the factories in non-member countries be motivated to reduce their emissions? Is it worthwhile to invest on high technology equipment to reduce carbon emissions? Shall factories be set up in member countries only? Under what conditions, should production be shifted from non-member society to member society (or vice versa)? Or, is there any market growth opportunity which benefits a factory and encourages its country to move from a non-member status to a member status? In order to answer such questions, a LSC model (a mixed integer programming model) is constructed in the context of a GrSC network, and then applied in the Company G's case to conduct a complete experiment. This study will provide a series of interesting managerial insights.

1.3. Structure of the thesis

In summary, the thesis, which consists of two parts, focuses on two applications of the LSC models. In particular, Part I (chapters 3, 4 and 5) is about the application in the EMS systems, where the model is mainly studied to investigate the pros and cons of demand assignment. Part II of the thesis (chapters 6, 7 and 8) is about the application in the GrSC network design, where the model is proposed to examine the impacts of carbon tariff imposition on the supply chain design. The structure of the thesis is shown in Figure 1.1.

Chapter 2 first provides a comprehensive review on the existing studies

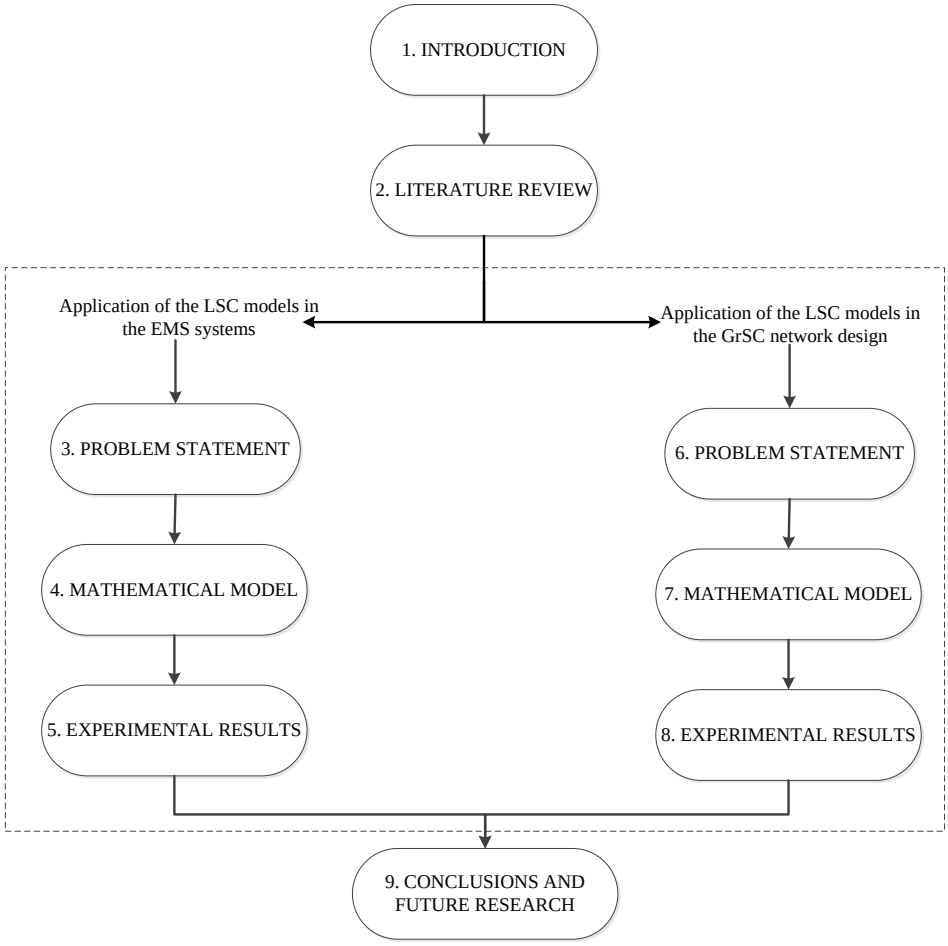

Figure 1.1.. Structure of the thesis

on the LSC models as well as the MCL models. Next, the overviews on the GrSC network design and carbon tariff mechanism are also presented.

Part I of this thesis includes chapters 3, 4 and 5. Chapter 3 states the problem in the EMS systems, and tries to explore the advantages and disadvantages of demand assignment. In this chapter, we define the failure probability with assignment and derive an upper bound of such failure probability which can be more easily incorporated into mathematical models. It is observed that when the overlap among the coverage areas is significant, or when the demand is covered by more stations, our model with such upper bound is superior to some previous models. Furthermore, this chapter

also shows the evidence that using such upper bound in the availability constraint would make the results viable in practice.

Chapter 4 develops the mathematical model which uses the upper bound derived in Chapter 3 in the availability constraint for the demand assignment problem; and also proposes a heuristic with good performances as well as low computational burden.

To further verify the results as analysed in chapters 3 and 4, Chapter 5 runs a series of numerical examples and observe plenty of interesting results. We note that the results of our model consistently outperform the results of the existing models. It is also demonstrated that the solutions of our heuristic are insensitive, which further supports the application of our model.

Part II of this thesis includes chapters 6, 7 and 8. In Chapter 6 the problem in the GrSC network design with carbon tariffs considered is well defined, including the main issues regarding the decision making and how the carbon tariffs are imposed through the supply chain.

Chapter 7 constructs the mathematical model for the problem in the GrSC network design, where pre-assignment is applied inspired by the insights obtained from Part I. This chapter first lists the objective function and all the constraints, and then transforms the non-linear terms in the model to be linear to make the model eventually become a mixed integer linear program.

In Chapter 8, the model proposed in Chapter 7 is applied in a real case from Company G, a pioneer in the electronic product industry, for a complete experimental study. We are interested in finding the actual

conditions where the introduction of carbon tariffs would force the firms in non-member countries to take actions to curb the emissions; the facility's willingness to join in the member society; the impacts of carbon tariffs on the total costs and emissions, and the impacts of supply range on the total costs with and without carbon tariffs.

The final chapter, Chapter 9 concludes this thesis and provides some future potential research directions.

Chapter 2 LITERATURE REVIEW

The studies of facility location problems have a long history. Savas (1969) provides the standard steps to analyse the EMS systems: (1) studying and defining the objective function; (2) finding possible alternatives to reach the objective; (3) identifying explicit criteria to evaluate the alternatives; (4) selecting the best alternative. After this work, many researchers have studied a variety of models for the facility location problem. In addition, both analytical methods and heuristics have been developed to approach the optimal solutions. Schilling et al. (1993) and Brotcorne et al. (2003) also present reviews on the models for facility location problem.

This chapter presents an overview on literature which is related to studies on applications of the LSC models in EMS systems and GrSC systems. Section 2.1 is a complete survey on the LSC models and the methods applied in those models to address the availability issue. Note that the availability is defined in Section 1.1, which is the probability that each demand node, say $j \in J$, can be immediately served by an available facility in its neighbourhood ($S_{COV}(j)$). After that, a review on the MCL models is offered in Section 2.2, where the methods applied in the MCL models to address the availability are discussed as well. sections 2.3 and 2.4 review two important topics which are relevant to Part II of this thesis: GrSC network design and carbon tariff mechanism.

2.1. Review on the LSC models and the availability issue

The first LSC model is proposed by Toregas et al. (1971), referred to as the following Deterministic LSC Problem (DLSCP).

$$DLSCP \qquad \min \sum_{i \in I} x_i \tag{2.1}$$

$$s.t. \sum_{i \in S_{COV}(j)} x_i \geq 1, \qquad \forall j \in J, \tag{2.2}$$

$$x_i \in \{0,1\}, \qquad \forall i \in I, \tag{2.3}$$

where the notations are listed as follows.

I = the set of candidate stations;

J = the set of demand nodes;

$S_{COV}(j)$ = the set of candidate stations which can cover Node j. In the DLSCP, $S_{COV}(j) = \{i \in I | L_{ij} \leq L\}$, where L_{ij} is the travel distance between i and j, L is the upper bound of the distance allowed;

$$x_i = \begin{cases} 1, & \text{if candidate station } i \text{ is selected;} \\ 0, & \text{otherwise.} \end{cases}$$

As objective (2.1) suggests, the objective of this model is to minimize the number of vehicles. Constraint (2.2) shows that all demand nodes are covered by at least one vehicle in its neighbourhood, i.e., $S_{COV}(j)$. Constraint (2.3) means that at most one vehicle can be located at one candidate station.

If each demand node denotes a region and the demand is randomly distributed within each region, the DLSCP is extended to the probabilistic

models. As indicated in Aly and White (1978), Chapman and White (1974) present a pilot version of the probabilistic LSC models by treating the response time as random variables with a given distribution. Following upon Chapman and White's work, Aly and White (1978) provide a variant of the probabilistic LSC model. The formulation of Aly and White's model, namely Probabilistic LSC Problem (PLSCP), is identical with the DLSCP except that $S_{COV}(j)$ is redefined as $\{i \in I | Pr(T_{ij} \leq \overline{t_i}) \geq p_{0i}\}$, where T_{ij} is a random variable of the response time; $\overline{t_i}$ is the maximal response time limit allowed at Station i; and p_{0i} is the required service level for Station i. Note that the randomness in this model only exists in the travel time.

Nonetheless, both DLSCP and the PLSCP do not address the availability issue of the vehicles at all. As discussed in Section 1.1, in order to assess the availability, the workload for each vehicle or station have to be known beforehand. However, it is hard to find the workload for each vehicle or station without a specific dispatching policy. This is just the dilemma which makes the availability quite difficult to be assessed accurately.

As indicated in Section 1.1, we define the *traditional failure probability*, denoted as fp, as the probability that a demand cannot immediately find any idle vehicle within its neighbourhood S_{COV}, under the assumption that any incoming demand would be served by the closest available vehicle and lost if all stations in its neighbourhood S_{COV} run out of vehicles. The traditional failure probability is a common performance target for the practitioners and also numerically investigated by many researchers (see, Ball and Lin 1993, Baron et al. 2009). To our best knowledge, how to assess the traditional failure probability is still an open question to the researchers.

The difficulties mainly lie in several aspects. To begin with, the vehicles at neighbouring stations operate dependently with each other. As we know, the coverage of the stations is usually overlapped, and the demand within the overlap can be responded by any nearby vehicle. Thus, the actual demand served by each vehicle is unknown in advance. Furthermore, as mentioned in Sorensen and Church (2010), such dependency is non-linear. The service requirement, such as the availability, is non-linearly dependent on the inputs like the fleet size. Last but not least, if we consider the states of all the vehicles under discussion, the size grows exponentially due to the combinatorial structure (see, Sorensen and Church 2010), which aggravates the computational burden.

Some simplifying assumptions are therefore invoked to address the issue above-mentioned. The first attempt is proposed by Daskin (1983) in the MCL problems, known as the uniform system-wide busy fraction assumption. The detailed review on this model will be referred in the next section. Considering the different workload through different regions, ReVelle and Hogan (1988) introduce the local busy fraction assumption into the LSC problems, for which the busy fractions are allowed to be different through different regions. Revelle and Hogan's model is referred to as Binomial PLSC Problem (BPLSCP) in Borrás and Pastor (2002). The local busy fraction in the neighbourhood of j is $b_j = \frac{\sum_{k \in N_j} \rho_k}{\sum_{i \in M_j} x_i}$, where ρ_k is the workload from Node k, N_j and M_j are the set of demand nodes and the set of candidate stations in the neighbourhood of j, respectively. Accordingly, the reliability at j is $q_j = 1 - (b_j)^{\sum_{i \in M_j} x_i}$. In this model, Constraint (2.2)

is refined to (2.4).

$$1 - (b_j)^{\sum_{i \in M_j} x_i} \geq p_j, \quad \forall j \in J, \tag{2.4}$$

where p_j is the target reliability in region j. Constraint (2.4) can also be revised to its linear equivalence as follows.

$$\sum_{i \in M_j} x_i \geq \lceil \frac{\ln(1 - p_j)}{\ln b_j} \rceil, \quad \forall j \in J.$$

Revelle and Hogan assume that for each demand node, all vehicles in its neighbourhood only serve the demand in the same region. In addition, the workload in the neighbourhood of each demand node (e.g., j), which can be calculated in advance, is equally dispatched to all vehicles nearby. Hence, the reliability at j is evaluated based on the prior known workload for any specific number of vehicles. Nevertheless, such assumption may not hold when there are vehicles located in the intersection of multiple neighbourhoods, as shown in Section 3.2. The reliability of these neighbourhoods are actually dependent on each other as they share some vehicles.

Researchers are then motivated to find other approaches to address the availability. Ball and Lin (1993) create an upper bound for the traditional failure probability. In their model, multiple vehicles are allowed to be located at one station. Ball and Lin's model is referred to as Poisson Reliability LSC Problem (PRLSCP) in Borrás and Pastor (2002).

$$PRLSCP \qquad \min \sum_{i \in I} \sum_{1 \le k \le u_i} W_{ik} z_{ik} \tag{2.5}$$

$$s.t.\ \Pi_{i \in S_{COV}(j)} \Pi_{1 \le k \le u_i} [Pr(D_i \ge k)]^{z_{ik}} \ \le \ 1 - p_1, \quad \forall j \in J, \tag{2.6}$$

$$\sum_{1 \le k \le u_i} z_{ik} \ \le \ 1, \qquad \forall i \in I,$$

$$z_{ik} \ \in \ \{0, 1\}, \quad \forall i \in I, \forall k,$$

where

W_{ik} = the cost of locating k vehicles at Station i;

D_i = the aggregate demand within the coverage area of Station i during the upper bound of the service time;

u_i = maximal number of vehicles allowed to be located at Station i;

$$z_{ik} = \begin{cases} 1, & \text{if } k \text{ vehicles are located at station } i; \\ 0, & \text{otherwise.} \end{cases}$$

The objective of this model, as formulated by expression (2.5), is to minimize the total cost. The left hand side of Constraint (2.6) is an upper bound of the failure probability for Node j, where the term $Pr(D(i) \ge k)$ is an upper bound of the failure probability at Station i if k vehicles are located. In Ball and Lin's model, the workload for each station is approximated as the sum of all workload in the coverage of the station.

When the demand for a station is known and follows a Poisson distribution, the Erlang loss formula (see, Erlang 1917) can be used to assess

the failure probability given the assumption that a call is lost without immediate service (see, Alsalloum and Rand 2006, Restrepo et al. 2009). This assumption (the loss system assumption) is applicable to the EMS systems where the patients not being immediately served will be handled by an alternative agent for service (see, Restrepo et al. 2009, Budge et al. 2009). Using the Erlang loss formula, Borrás and Pastor (2002) study the Queueing Reliability LSC Problem (QRLSCP), by replacing the availability constraint in PRLSCP (Constraint (2.6)) to the following inequality.

$$\frac{{\rho_i}^{z_i}/z_i!}{\sum_{k=0}^{z_i} {\rho_i}^k/k!} \leq 1 - p_1, \quad \forall i \in I, \tag{2.7}$$

where $\rho_i = \sum_{j \in S_{PT}(i)} \lambda_j \overline{s}$ is the workload, i.e., the sum of multiplications of the demand arrival rate λ_j and the upper bound of the service time s for all demand nodes in the coverage area of Station i. z_i is the number of vehicles located at Station i. Hence, Station i operates as an $M/G/z_i/z_i$ queue, with a number of servers but no buffer spaces for incoming demand. We denote the left hand side of (2.7) as FP_i^Q, which is also an upper bound of the traditional failure probability for each demand node within the coverage of Station i. The QRLSCP always requires no more vehicles than the PRLSCP as FP_i^Q is a tighter bound than the left hand side of Inequality (2.6).

Borrás and Pastor (2002) also study the Binomial Reliability LSC Problem (BRLSCP). In this model, the failure probabilities at all stations are independently estimated and the reliability at each station is evaluated in advance for any specific fleet size. In the BRLSCP, the availability con-

straint is replaced by Inequality (2.8).

$$(\frac{\rho_i^{'}}{z_i})^{z_i} \leq 1 - p_1, \quad \forall i \in I, \tag{2.8}$$

where $\rho_i^{'} = \sum_{j \in S_{PT}(i)} \lambda_j s$ and s is the average service time. As an estimate of the traditional failure probability, $(\frac{\rho_i}{z_i})^{z_i}$ is denoted as FP_i^B in this thesis. According to Borrás and Pastor (2002), the BRLSCP further reduces the number of required vehicles compared to the QRLSCP. However, FP_i^B is not necessarily an upper bound of the traditional failure probability. As shown in Section 3.3.4, the left hand side of Inequality (2.8), FP_i^B, may underestimate the traditional failure probabilities for most demand nodes.

Besides, Beraldi and Bruni (2009) discuss a two stage stochastic program to tackle the availability issue. In the second stage, the demand is assigned for a specific scenario. As a consequence, the assignment cannot be implemented since it depends on some specific scenario, which is usually unknown in advance. Other two stage stochastic programming models proposed for the problem can be referred to in Snyder (2006). There are also some location-inventory models which is relevant to our research, which conduct demand assignment though without justification. For example, Shen et al. (2003) introduce the assignment variables into their model, to assign the retailers to the Distribution Centres (DCs) instead of considering all retailers in the coverage of each DC. In their paper, the optimal inventory level at each DC cannot be identified without decisions of demand assignment. The contribution of our work (in part I) is to explore the benefits of demand assignment by both analytical and numerical studies. Then the policy and insights from this work can be applied to other

fields (e.g., II of this thesis.

Baron et al. (2009) propose their models to ensure the availability at each station, referred to as BBKK1 and BBKK2 in their paper. They also show that BBKK2 leads to quite marginal improvement over BBKK1, but is significantly harder to solve. Hence, only the BBKK1 is discussed in our study.

$$BBKK1 \qquad \min \sum_{i \in I} m(\rho_i) y_i \tag{2.9}$$

$$s.t. \sum_{i \in S_{PT}(j)} y_i \geq 1, \qquad \forall j \in J, \tag{2.10}$$

$$y_i \in \{0,1\}, \qquad \forall i \in I. \tag{2.11}$$

where $m(\rho_i)$ is the least number of required vehicles to ensure the availability at Station i if the total workload in the coverage of Station i is ρ_i. The value of $m(\rho_i)$ is calculated with Erlang loss formula. y_i is a binary variable, denoting whether Station i is activated. We note that Baron et al. (2009) also consider all demand in the coverage of each station to calculate the number of vehicles in need. As explained earlier, the required number of vehicles will be inevitably overestimated if the neighbourhoods or regions are heavily overlapped. As $m(\rho_i)$ is calculated using Erlang loss formula, BBKK1 actually estimates the availability equivalently as the QRLSCP. The difference between the BBKK1 and the QRLSCP lies in the decision on which stations should be activated. In Part I of this thesis, we analytically compare our method with FP^Q in sections 3.2 and 3.3, and numerically compare our model with the BBKK1 in Section 5.1, as currently BBKK1 is the best model ensuring the target failure probability is

guaranteed to our best knowledge.

2.2. Review on the MCL models and the availability issue

As a parallel of the LSC models, this section also provides a review on the MCL models. In particular, the methods adopted in those models to address the availability issue are discussed.

Church and ReVelle (1974) build up the first MCL model, referred to as the Maximum Covering Location Problem (MCLP), to demonstrate how to locate a given number of facilities to maximize the demand covered within a target service distance.

$$MCLP \qquad \max \sum_{j \in J} (D_j Y_j) \tag{2.12}$$

$$s.t. \sum_{i \in N_j} x_i \ \geq \ Y_j, \qquad \forall j \in J, \tag{2.13}$$

$$\sum_{i \in I} x_i \ = \ K, \tag{2.14}$$

$$x_i, Y_j \ \in \ \{0, 1\}, \qquad \forall i \in I, \ \forall j \in J, \tag{2.15}$$

where

D_j = the demand at Node j;

K = the given fixed number of vehicles to be located;

$$Y_j = \begin{cases} 1, & \text{if node } j \text{ can be covered by at least one vehicle within a stated service distance.} \\ 0, & \text{otherwise.} \end{cases}$$

In this model, the authors assume that there is always at least one vehicle

available once a new demand arrives. That is, like the earlier LSC models (see, Toregas et al. 1971, Aly and White 1978), this model also does not take into consideration system congestion or the demands would be lost due to no vehicle available when they arrive. In other words, the availability is not taken into account.

When taking system congestion into consideration, Daskin (1983) further develops the work of Church and ReVelle (1974) by assuming all facilities in the whole system are equally busy, and studies how to locate a fixed number of facilities to maximize the expected percentage of demand covered instead. Daskin's model is referred to as the Maximal Expected Covering Location Problem (MEXCLP). To present his model, we first need to list the following notations below.

b_s = the system-wide busy fraction of each facility, which is assumed to be a known parameter;

q_k = service reliability at any call node if k ambulances locate within the target area of this node. Since the busy probability of each facility is b_s, $q_k = 1 - (b_s)^k$;

$$Y_{j,k} = \begin{cases} 1, & \text{if node } j \text{ is covered by } k \text{ vehicles} \\ 0, & \text{otherwise.} \end{cases}$$

The MEXCLP is constucted as follows.

$$MEXCLP \quad \max \sum_{j \in J} \sum_{k=1}^{K} D_j q_k Y_{j,k} \tag{2.16}$$

$$s.t. \sum_{i \in N_j} x_i - \sum_{k=1}^{K} k Y_{j,k} \geq 0, \qquad \forall j \in J, \tag{2.17}$$

$$\sum_{k=1}^{K} Y_{j,k} \leq 1, \tag{2.18}$$

$$\sum_{i \in I} x_i = K, \tag{2.19}$$

$$x_i, Y_{j,k} \in \{0,1\}, \quad \forall i \in I, \ \forall j \in J, \ \forall k, \tag{2.20}$$

As stated previously, the author assumes that the total workload is equally shared by all vehicles in the system. Hence, all vehicles share a uniform busy fraction, i.e., $b_s = \frac{\sum_{j \in J} \rho_j}{K}$, where ρ_j is the workload from Node j and K is the predetermined given fleet size. The reliability at the demand node $j \in J$ can therefore be calculated as $q_k = 1 - (b_s)^k$ if k vehicles are located in set $S_{COV}(j)$. The equally shared workload implies that the demand is randomly dispatched to all vehicles in the system, which may not be applicable to most practical cases.

The underlying assumption in Daskin's model does not take into account the differences among regions, which is the common case in EMS systems. Thus, the solutions from this model may lead to unfairness for different regions, e.g., varied sanctification level for urban areas and rural areas. Later, ReVelle and Hogan (1989) relax the system-wide busy fraction and assume busy fractions are different in various regions, then construct a

model for Maximum Availability Location Problem (MALP).

$$MALP \qquad \max \sum_{j \in J} D_j Y_j \tag{2.21}$$

$$s.t. \sum_{i \in N_j} x_i - s_j Y_j \ \geq \ 0, \qquad \forall j \in J, \tag{2.22}$$

$$\sum_{i \in I} x_i \ = \ K, \tag{2.23}$$

$$x_i, Y_j \ \in \ \{0,1\}, \qquad \forall i \in I, \ \forall j \in J, \tag{2.24}$$

where

s_j = minimum number of facilities required to ensure that Node j can be covered with a probability of p_1;

$$Y_j = \begin{cases} 1, & \text{if node } j \text{ can be covered with a probability no less than } p_1; \\ 0, & \text{otherwise.} \end{cases}$$

The objective of this model is to maximize the coverage within a target response time with p_1-reliability when the number of facilities is given. Like the assumption in ReVelle and Hogan (1988), the authors also treat neighbouring areas as independent and isolated. That is, all areas are locally constrained, the demand within one area can only be served by the facilities within the same area, and the facilities within one area can only be assigned to serve the calls within the same area too. However, as stated in the previous section, this assumption may not be applicable in practice as the neighbourhoods are always dependent on each other as they share some vehicles.

With the local busy fraction concepts, Sorensen and Church (2010) ex-

tend the MEXCLP model to a new one, referred to as Local Reliability-based MEXCLP (LR-MEXCLP). They also present the advantages of their model over MALP and MEXCLP by numerical studies. Berman and Krass (2002) also generalize the basic MCLP with the coverage degree being a non-increasing step function of the distance.

There are some variants of the MCL models. Erkut et al. (2008) build their model to maximize the expected number of patients who survive when the number of ambulances is fixed. In addition, Restrepo et al. (2009) seek to minimize the number of lost demand with a fixed fleet size. In their paper, the authors adopt Erlang loss Formula to find the failure probability at each station.

Because of the possible great computational efforts on solving the integer programming, many researchers are investigating efficient heuristics for the models (see, Farahani et al. 2014, Pereira et al. 2015). They also carry out extensive computational experiments to assess the quality of their heuristics.

As a conclusion of the existing LSC models and MCL models, different models estimate the workload in different ways: Daskin (1983) assumes all workload is equally shared by all vehicles; Revelle and Hogan (1988, 1989) assume the workload in each neighbourhood is equally shared by the vehicles in the same neighbourhood. In addition, some other models (Ball and Lin's model, the QRLSCP and the BRLSCP) approximate the workload for each station as the sum of all workload in the coverage of the station. As discussed in Section 1.1, such approximation would overestimate the actual workload for each station, particularly when the workload through

the neighbourhoods or regions is highly overlapped. Our study in Part I of this thesis is motivated to explore the benefits of the demand assignment in addressing such overestimation problem.

2.3. Review on the GrSC network design

The traditional supply chain network models mainly focus on maximizing long-term economic profits. Goetschalckx et al. (2002) provide a review of such models. Goetschalcks and Fleischmann (2008) also study the strategic planning and design of the supply chain to maximize economic benefits over a long period of time. The supply chain network design problem is a comprehensive strategic decision problem which needs to be optimized for long-term efficient operations of the entire supply chain. It is studied to determine a portfolio of configuration parameters such as the number, location, capacity of facilities.

However, recently with the growing attention paid to the global warming crisis, supply chain managers cannot only focus on economic profits any more. They have to take the environmental issues into account. A concept called "Green Supply Chain Management (GrSCM)" emerges (see, Beamon 1999, Srivastava 2007, Wang et al. 2011, Diabat and Al-Salem 2015, Tognetti et al. 2015). It means that, the managers are more concerned about the environmental impacts when they are designing the supply chain, instead of only maximizing economic profits, as was the case before. For example, Abdallah et al. (2013) indicate that industrial sectors are facing challenges to integrate environmentally related decisions into supply chain network design and logistics activities. Martí et al. (2015) propose a

model to achieve a balance between operations and environmental impacts, which arising from the interaction between different-stage supply chain processes such as procurement, assembling, transportation and inventory management.

Srivastava (2007), Dekker et al. (2012) both provide comprehensive reviews on GrSCM. According to Srivastava (2007), two types of "greenness are considered by current researchers: green design for products (see, Kuo et al. 2001, Abdallah et al. 2013) and green operations. Our study falls in the second category, including sustainable manufacturing and remanufacturing, recycling logistics network design and waste disposal. A great deal of research has been done on this topic. Fleischmann et al. (2001) consider some recovery centers and compare the economic costs between the reverse logistics network with the traditional open-loop networks. Savaskan et al. (2004) also study the most effective reverse channel structure for collecting used products from consumers in a closed-loop network. In addition, Wang et al. (2011) consider that companies can invest in several technology levels with different environmental impacts, and propose a multi-objective optimization model to investigate the trade-off between the overall logistic costs and total CO2 emissions. They find that enlarging the capacity of the network or increasing the supply to the factories can realize CO2 emission reduction. Chaabane et al. (2012) propose a mixed integer linear programming model to design sustainable supply chain under the "cap-and-trade" scheme, which combines the life cycle assessment principle and the traditional material balance at each node of the network. The model is applied in the aluminium industry and the conclusion is that efficient car-

bon management policies will indeed achieve carbon reduction targets in a cost-effective manner. Jaber et al. (2013) discuss the impacts of different carbon policies on the production decisions of manufacturers. Chen and Hao (2014) investigates the pricing and production policies for competing firms with carbon emissions tax policy. As reported by O'Connell and Stutz (2010), Dell unveils the carbon footprint of its products and is also committed to being environmentally responsible by reducing their carbon footprint.

How to manage the inventory under various carbon regulations is also a concern for companies. Hua et al. (2011) use a simple inventory model to study how firms manage their inventories under the "cap-and-trade" mechanism, and examine the impacts of carbon trade, carbon cap and carbon price on the firm's decisions, total costs and emissions. Benjaafar et al. (2013) also study this problem via a series of inventory models followed by numerical experiments as well. They observe that it is possible to reduce the CO2 emissions without significantly increasing the costs, by adjusting order quantities. Regarding Benjaafar et al.'s research, Chen et al. (2013) provide the analytical support for such observations. Another conclusion from their models is that even a small tax imposed on the CO2 emissions can motivate changes on decisions including procurement, inventory management, and deployment of production facilities to significantly reduce emissions.

An important issue in green supply chains is the transportation mode selection, especially for international trading. As stated in Cristea et al. (2013), in North America, international transport constitutes about 67 %

of its export related emissions, and more than 80 % of machinery export emissions are also due to international transport. They also point out that the costs and carbon emissions would differ a lot for different transportation modes. The data from Company G also demonstrate that the emission of different transportation modes varies a lot. For example, the ratio of emissions by sea versus by air is about 1:40. Therefore, many green supply chain models (see, Ramudhin et al. 2010, Chaabane et al. 2012) take transportation mode selection into consideration when developing their models.

Most of those works focus on the trade-off between total cost and the carbon footprint just before the final products are prepared. For example, the model proposed by Chaabane et al. (2012) does not take into account the emissions when the products are put into use. In addition, many such green supply chain models do not consider raw material selection when designing the networks. As indicated in O'Connell and Stutz (2010), however, the carbon footprint in the use phase is a major contributor to the overall emissions, especially in the computer industry. Through interviewing with managers from Company G, it is also confirmed that the carbon footprint in the use phase is a key factor of the total emissions through the lifetime of a notebook. Furthermore, they state that the CO2 emissions during usage mainly depend on the selection of raw materials, but are little related to the production procedure.

In our study, the raw materials used to assemble the products are referred to as *modules*, and we also consider several *options* (different purchasing prices and different emission levels, but with equivalent functions) that can be selected for each module. In particular, each option of a module

has its own unit purchasing price, *cradle-to-gate CO2 emission* and expected CO2 emission in future usage. Note that the cradle-to-gate CO2 emission of a module refers to the emission from extraction or acquisition of raw materials to the time when the module is ready to be sold (see, British Standards Institute 2008). These two emissions may increase the carbon tariff imposed when exported to member society, because: (1) Usually higher purchasing price means lower cradle-to-gate CO2 emissions and lower expected CO2 emissions in future usage. With environmental impact taken into consideration, factories would not just ask for low price modules as they traditionally do. Instead, they struggle to balance the purchasing costs and the CO2 emissions, as more carbon emissions may lead to a higher carbon tariff when exported to the end markets; (2) On the other hand, we take the emissions during usage phase into account, which may also bring about some carbon tariff. That is, when the final products arrive to the end markets, they will estimate these emissions and impose a certain amount of tariff. A detailed description of how carbon tariffs are imposed will be further discussed in Section 6.4.

2.4. Review on the carbon tariff mechanism

As explained in Section 1.2, unilateral (or sub-global) carbon control policies to curb CO2 emissions would lead to two possible consequences: overall emission increase due to carbon leakage; and the danger to domestic industries in member countries that take more stringent regulations on carbon reduction, which is called the *distortion in competitiveness* (see, Persson 2010). Houser et al. (2008) conclude that the effectiveness of a carbon tariff

mechanism varies a lot for different industries and firms, and such mechanism does not benefit all US industries. Some other studies (see, Reinaud 2008, Gielen and Moriguchi 2002, Demailly et al. 2005) quantitatively assess the scale of carbon leakage and the magnitude of the distortion in competitiveness.

The best approach to deal with the above concerns is to achieve an international agreement that imposes an equivalent carbon price on all emitters. However, achieving such a unanimous agreement at present is not an easy task (see, Zhou et al. 2010). Houser et al. (2008) also indicate that response measures to address the above-mentioned problem should focus on three targets: (1) reducing the production cost for domestic producers when complying with the regulations; (2) imposing equivalent costs on foreign producers from non-member societies at the border; and (3) encouraging non-member countries to impose equivalent costs on their industries directly.

A response measure which has been widely discussed recently to guard against carbon leakage and to level the playing field for domestic industries is to impose a similar penalty on the emissions of goods from non-member countries through *carbon border adjustment*, such as a carbon tariff mechanism. However, the debate on this mechanism is still ongoing. As indicated in Drake (2011), when EU member states voted to add a border adjustment to the EU-ETS, both Britain and the Netherlands publicly opposed such a proposal. The situation is similar in the US, according to Waxman and Markey (2009) and Broder (2009), where a bill which includes a border adjustment passes successfully through the House of Representatives, but

the President criticizes it.

Whether the carbon tariff mechanism can effectively deal with carbon leakage and the distortion in competitiveness is still an open question now. Firstly, Böhringer et al. (2013) states that developing countries argue that it in unfair to force them to reduce their emissions without any compensation in the near future, as enough emission quota will help them for their economic growth. Secondly, there are also different attitudes on the efficiency of such a mechanism. Proponents contend that in the absence of a global agreement, the carbon tariff mechanism, although still hypothetical, is necessary and would prove to be effective in tackling the concerns under discussion. As indicated in Zhou et al. (2010), as leakage results from the higher carbon costs the producers in member countries face, it can be avoided if the imported goods are adjusted for the carbon cost differences. Gros et al. (2010) also advocate such a mechanism because they think it transfers carbon price, at least partially, to the non-member society. Drake (2011) examines how carbon tariffs affect the market share distribution, global emissions, and so forth.

On the other hand, opponents of a carbon tariff mechanism doubt the effectiveness of such measures. They argue that such adjustments only build a trade barrier and are anti-competitive. Moreover, they claim that there would be many challenges when implementing such a mechanism. For example, how would the carbon tariff rates for foreign goods coming from different regions be determined? How would the carbon footprints of the imported goods which heavily depend on the whole life-cycle be estimated? It is even stated in Institute of Public Affairs (2009) that carbon tariffs

are "costly, ineffectual and protectionist". They claim that the expense to assess the carbon component and the new regulations and administration required to implement the carbon tariff mechanism is far in excess of the benefits brought by the adoption of the mechanism. Moore (2011) also thinks that the carbon tariff mechanism is not viable due to the administrative difficulties. Dong et al. (2015) study the impacts on carbon tariff on China's export and find that, carbon tariff imposition may be ineffective in reducing the domestic emissions as well as the global emissions.

Discussions on the carbon tariff can also be found in Sindico (2008), Cosbey (2008), Kuik and Hofkes (2010), Van Asselt and Brewer (2010), Monjon and Quirion (2010) and Monjon and Quirion (2011).

With regards to the debate on whether the carbon tariff should be imposed, our work is proposed to provide managerial insights to those designers by investigating the impacts of the carbon tariff mechanism on the supply chain design. In particular, a detailed description of such mechanism will be offered in Section 6.4. Inspired by the insights obtained from the study in Part I, we apply similar pre-assignment into the model by introducing a series of decision variables to connect the three-tier facilities. In terms of modelling, we offer a novel way to incorporate carbon tariff and a tricky way to formulate the problem as a mixed integer program. The program then can be solved using standard methods. The managerial insights generated in this part answer the questions the managers are concerned with, as stated in Section 1.2. In other words, our work helps to shed light on those who want to study the impacts of carbon tariff in practice.

PART I

Application of the LSC models in the EMS systems

Chapter 3 PROBLEM STATEMENT FOR THE MODEL WITH DEMAND ASSIGNMENT

As stated in Section 1.1, our thesis studies the LSC model in the EMS systems to explore the pros and cons of demand assignment. This chapter clearly states the problem in order to perform the model construction in the next chapter. In particular, Section 3.1 defines another type of failure probability, referred to as *failure probability with assignment*, which is different from the traditional failure probability defined in Section 1.1. Based on this new failure probability, the advantages of demand assignment are discussed in Section 3.2 by two simple case studies. Finally, Section 3.3 compares the failure probability with assignment with the traditional one, in order to demonstrate whether the solutions would be practical by incorporating this new failure probability into the mathematical model.

3.1. The failure probability with assignment and its upper bound

In order to discuss the benefits of assignment, we first study another type of failure probability, failure probability with assignment (denoted as FP^A), whose definition is different from the traditional one.

Recall that as defined in Section 1.1, the traditional failure probability is the probability that a demand does not see any idle vehicle in the

neighbourhood, which is widely the typical performance target for the practitioners. The underlying assumption for the traditional failure probability is that each demand could find any idle vehicle in its neighbourhood to be served, and only be lost if all stations in its neighbourhood are busy.

On the other hand, we define another term, referred to as failure probability with assignment alternatively. With demand assignment assumption, if Node j is assigned to Station i, the demand from Node j will be lost if Station i runs out of vehicles. (Note that this assumption is proposed to derive our insights on demand assignment, while in reality Node j will find any other available vehicles nearby for service.) Thus, FP_j^A is defined as the probability that a demand from Node j does not see any idle vehicle at Station i (with assignment of Node j to i in advance). This definition of failure probability is analogous with the percentage of unsatisfied retailers orders in the location-inventory model of Shen et al. (2003). The demand arrivals are assumed to follow Poisson distributions. Then according to the Poisson Arrivals See Time Average (PASTA) (Wolff (1982)), the failure probability with assignment of each demand node is equal to the probability that the assigned station runs out of vehicles. That is, $FP_j^A = FP_i^A$ if node j is assigned to station i. The differences between the two failure probabilities above-mentioned will be further investigated in Section 3.3.

3.1.1. The failure probability with assignment

As mentioned above, the demand arrivals at each node $j \in J$ are assumed to follow a Poisson distribution with rate λ_j. The service times s_{ij} are assumed to be continuous random variables with any general distribution-

s. Moreover, s_{ij} are dependent on the locations of *both* the demand and the assigned station. These assumptions for the service time are reasonable because the travel time, a critical component in the service time, is usually stochastic and does not follow any frequently-used distribution. In addition, the travel time and the time required for the vehicle from a call node to the hospital are dependent on the locations of the demand and the assigned station, respectively. Furthermore, the assignment variables y_{ij} are assumed to be binary in our model.

$$y_{ij} = \begin{cases} 1, & \text{if a vehicle at station } i \text{ is assigned to serve the demand} \\ & \text{at node } j; \\ 0, & \text{otherwise.} \end{cases}$$

Similar to the QRLSCP, the failure probability with assignment can also be derived using the Erlang loss formula with workload ρ_i^A which only considers the demand nodes assigned to station i, instead of all nodes in its coverage area. Lemma 1 shows how ρ_i^A is obtained. Note that $S_{PT}(i)$ is the set of demand nodes within the coverage area of station i, i.e., $S_{PT}(i) = \{j \in J | L_{ij} \leq \text{target coverage radius}\}$. In this study, $S_{PT}(i) = \{j \in J,$ where $t_{ij} \leq \bar{t}\}$, t_{ij} and $\bar{t}$ are the response time for the demand at node j which is served by station i and the maximum response time allowed, respectively.

Lemma 1. *With assignment of the demand, the average workload assigned to station i (including the possibly lost demand part) is as shown by the following equation.*

$$\rho_i^A = \sum_{j \in S_{PT}(i)} \frac{\lambda_j}{\mu_{ij}} y_{ij},$$

Where μ_{ij} is defined as the service rate when the demand originated from node j is served by vehicles in station i. In particular, $\mu_{ij} = 1/s_{ij}$.

The proof is shown in Appendix A. Now the constraint (2.7) is reformulated as follows.

$$\frac{(\rho_i^A)^{z_i}/z_i!}{\sum_{k=0}^{z_i}(\rho_i^A)^k/k!} \leq 1 - p_1, \quad \forall i \in I, \tag{3.1}$$

where the left hand side of Inequality (3.1) is denoted as FP_i^A. As mentioned above, ρ_i^A only includes the demand assigned to station i. On the other hand, ρ_i includes all demand within the coverage area of station i, i.e., $\rho_i = \lambda_i/\mu_i = \sum_{j \in S_{PT}(i)} \frac{\lambda_j}{\mu_{ij}}$.

From Inequality (3.1), we know that FP_i^A increases with parameter ρ_i^A. We also recall that FP_i^Q is defined as the left hand side of Inequality (2.8) in Section 2.1, with the same form of FP_i^A except the parameter ρ_i^A is substituted by ρ_i. On the other hand, ρ_i^A is a subset of ρ_i, indicating that $\rho_i^A \leq \rho_i$. Then we can conclude that $FP_i^A \leq FP_i^Q$, where the equality holds when the coverage areas of all stations do not overlap at all.

A further observation of the Erlang loss formula is as shown in Figure 3.1. Note that the real figure should look like step functions rather than several lines. Take $FP = 5\%$ for example, there should be a horizontal line at 6 when workload (ρ) lies in 2.22-2.96. That means, when the workload lies in this range, the required number of vehicles is 6 to ensure the target service level. We just show Figure 3.1 in order to make it easier to distinguish. The partial derivative $\frac{\partial z}{\partial \rho}$ (or $\frac{\partial z}{\partial \rho^A}$): $\forall i \in I$ is positive and decreasing with the failure probability FP^Q (or FP^A), respectively. This implies that, as the ratio $\frac{\partial z}{\partial \rho}$ (or $\frac{\partial z}{\partial \rho^A}$) is positive, more vehicles are needed when the workload

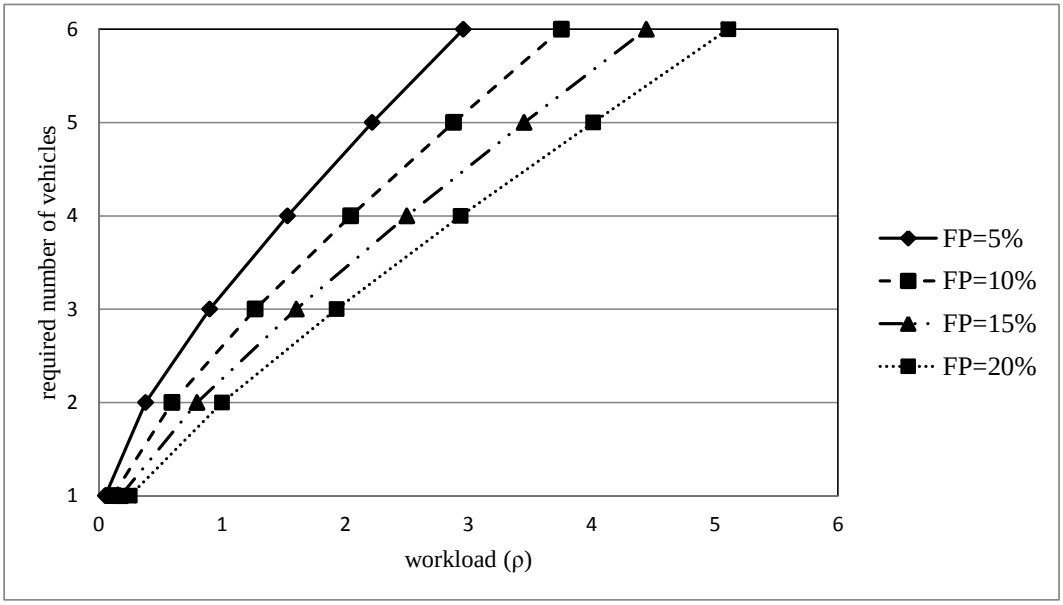

Figure 3.1.. Relation between the workload and the required number of vehicles for different failure probabilities in loss systems

increases. In other words, the potential shortcoming of the QRLSCP and other location models (e.g., Ball and Lin's model, the BRLSCP) is that all demand nodes in the coverage area of each station are included to estimate the workload of the station, even though the nodes can also be served by other stations. By assigning demand to stations, the workload ρ (or ρ^A) for each station in the Erlang loss formula is reduced, hence yielding fewer vehicles to meet the failure probability requirement. The impact of assigning becomes more significant under a lower failure probability (higher reliability) requirement as $\frac{\partial z}{\partial \rho}$ (or $\frac{\partial z}{\partial \rho^A}$) decreases with the failure probability. For instance, according to Figure 3.1, 3 more vehicles (3 to 6) have to be located under $FP = 5\%$ when the workload increases from 1 to 3, in contrast with only 2 more vehicles (2 to 4) under $FP = 20\%$.

3.1.2. An upper bound of the failure probability with assignment

Although assigning demand may result in fewer vehicles required, how to analytically solve the models with constraints of FP^A is still an open question. As indicated in Restrepo et al. (2009), the Erlang loss function (the left hand side of Inequality (3.1)) is not a jointly convex function of (ρ_i^A, z_i). ρ_i^A also contains an unknown decision variable y_{ij}, which makes the optimization problem almost unsolvable in practice. An upper bound of FP_i^A is proposed as follows.

Proposition 1.

$$FP_i^A \leq Pr(\sum_{j \in S_{PT}(i)} D_j(\frac{1}{\mu_{ij}}) y_{ij} \geq z_i), \quad \forall i \in I, \tag{3.2}$$

where $D_j(\frac{1}{\mu_{ij}})$ denotes the number of calls occurred at node j during the time length $\frac{1}{\mu_{ij}}$, and it follows a Poisson distribution with mean $\frac{\lambda_j}{\mu_{ij}}$, i.e., $D_j(\frac{1}{\mu_{ij}}) \sim POISSON(\frac{\lambda_j}{\mu_{ij}})$.

The right hand side of inequalities (3.2) is denoted as FP_i^{AU}. In order to prove Proposition 1, the following lemma is required. The detailed proof of Proposition 1 can be found in Appendix B.

Lemma 2. *For any demand node $j \in J$ with arrival rate λ_j,*

$$Pr(D_j(s) = N) = Pr(D_j^{'}(1) = N), \tag{3.3}$$

where s is a certain time length and $D_j^{'}(1) \sim POISSON(\lambda_j s)$.

According to the assumption that the arrivals at node j follow a Poisson distribution, and by the definition of Poisson distribution, it is easy to see that Lemma 2 holds.

With Lemma 2, the demand from different nodes with different service rates can be uniformized to that originated from only one node. For the right hand side of inequalities (3.2), we need to sum up the total demand covered by each station during their respective average service time. According to Equation (3.3) in Lemma 2 and the additive property of Poisson distribution, $Pr(\sum_{j\in S_{PT}(i)} D_j(\frac{1}{\mu_{ij}})y_{ij} \geq z_i)$ can be uniformized to $Pr(\widehat{D}_i \geq z_i)$, where $\widehat{D}_i$ is a Poisson random variable denoting the arrivals within one unit of time with rate $\rho_i^A = \sum_{j\in S_{PT}(i)} \frac{\lambda_j}{\mu_{ij}} y_{ij}, \forall i \in I$.

Constraint (3.1) can thus be replaced by the following Inequality (3.4) according to Proposition 1.

$$Pr(\sum_{j\in S_{PT}(i)} D_j(\frac{1}{\mu_{ij}})y_{ij} \geq z_i) \leq 1 - p_1, \quad \forall i \in I. \tag{3.4}$$

The solutions of the model with Constraint (3.4), in turn, will guarantee that the failure probability with assignment FP_i^A does not exceed $1 - p_1$.

In addition, the following proposition shows the asymptotical properties of FP_i^{AU} when either $\rho_i^A \to 0$ or $\rho_i^A \to \infty$.

Proposition 2. $\forall i \in I$,

$$\lim_{\rho^A\to 0} FP^{AU} - FP^A = 0; \qquad \lim_{\rho^A\to 0} \frac{FP^{AU}}{FP^A} = 1$$

$$\lim_{\rho^A\to \infty} FP^{AU} - FP^A = 0; \qquad \lim_{\rho^A\to \infty} \frac{FP^{AU}}{FP^A} = 1$$

The proof can be found in Appendix C. According to Proposition 2, our upper bound FP^{AU} asymptotically approaches to the failure probability with assignment FP^A whether when ρ^A is quite large or extremely small.

3.2. The advantages of assignment

As discussed in Section 1.1, the demand assignment appears to be unfavorable as the benefit of risk pooling is lost. However, without assigning demand to stations, the actual workload of the stations is hard to obtain. Most existing models (e.g., PRLSCP, QRLSCP, BRLSCP and BBKK1) thus consider all the demand in the coverage area of each station as its workload. This results in the problem of workload overestimation and may yield an inferior solution. Note that both FP^Q (in Constraint (2.7)) and FP^{AU} (in constraints (3.4)) are upper bounds of the failure probability with assignment FP^A. This section compares the results of using FP^Q and FP^{AU} in the availability constraints under the following two simple cases. The comparison provides further information with respect to the benefits of assignment.

3.2.1. Two station case

The two station case is constructed as in Figure 3.2. The symmetrical system contains three demand nodes A, B and C with demand following Poisson distributions of arrival rates λ_A, λ_B and λ_C, respectively, and $\lambda_A = \lambda_B$. Only A and B are candidate stations with equal coverage areas, and we can only locate vehicles at nodes A and B. Node C is in the coverage area of both stations and can be covered by either one. To simplify our analysis and also ensure the results we want can be generated, the service time for all the demand is identically assumed to be 1. The overlap fraction is defined as $o = \frac{\lambda_C}{\lambda_A+\lambda_C} = \frac{\lambda_C}{\lambda_B+\lambda_C}$. In order to investigate the impact of the overlap while the demand rate in the coverage of each station remains the same,

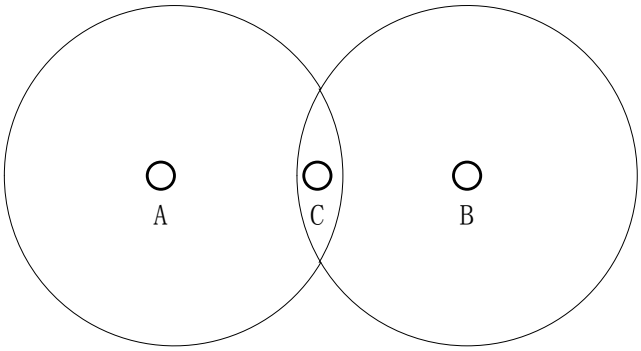

Figure 3.2.. The example of the two station case

we also assume that the total demand arrival rates within the coverage of A or B is a constant ρ, that is $(\lambda_A + \lambda_C) = (\lambda_B + \lambda_C) = \rho$ no matter how much λ_C changes. As the service time is 1, the total workload within the coverage of each station is also ρ. We also conduct similar analysis when fixing the demand rate $\lambda_A + \lambda_C + \lambda_B$ and changing λ_C only, in order to get different overlap fractions. In this way, the workload within the coverage of each station increases as the overlap fraction increases. It is observed that the results are in compliance with the results shown in this case study.

The workload for each station is always ρ in the QRLSCP regardless of the overlap fraction o. FP^Q can be obtained from the left hand side of Inequality (2.7).

Regarding FP^{AU}, it is assumed that each demand is assigned to the nearer station, and the demand at node C is equally assigned to the two candidate stations. (We note "equally assigned" as an assumption to derive the corresponding failure probabilities and required vehicles, and the next section will demonstrate the results perform well under the case in reality: a demand will find any available vehicle nearby for service). Thus, the demand assigned to station A (or B) is $\lambda_A + 0.5\lambda_C$ (or $\lambda_B + 0.5\lambda_C$). In

other words, the demand assigned to station A (or B) is $\rho(1-0.5o)$. FP^{AU} is then obtained as follows.

$$FP^{AU} = Pr(R \geq z), \quad \text{where } R \sim POISSON(\rho(1-0.5o)). \tag{3.5}$$

3.2.2. Three station case

A similar study is conducted for the three station case. In this case, there are three stations A, B and C located at the vertex of an equilateral triangle, within which the demand is uniformly distributed. Here the demand is assumed to be continuously distributed just to make the presentation and figures clear. (The corresponding problem where the demand is discrete is also studied, i.e., one demand node in A, B, and C, respectively; one demand node within the overlapped area of A and B, A and C, B and C, respectively; and one demand node within the overlapped area of A, B and C, which is similar with the two station case example above. We find that the results are identical with the continuous scenario.) Furthermore, the coverage areas of the stations are equal fan-shaped areas with center angle of 60 degrees (e.g., $W(A)$ for station A in Figure 3.3(a)). Similar to the two station case, the demand rate in the coverage of each station is assumed to be ρ, and the service rate is still one per unit time. With assignment, each demand is assigned to the nearest station.

Further explanation for the motivation of studying the three station case problem is that, even if we locate the stations inside the triangle instead of just at the vertex of the triangle as our assumption above, which is equivalent to assuming that the demand is distributed over the whole circle

coverage of each station. The results in terms of the impacts of demand assignment are identical with our current analysis. The reason why we only consider the demand inside the triangle is to make a better comparison with the previous two station case study. In particular, with the same overlap scale, will less fleet size be required? How the failure probability will change in the three station case as discussed in Section 3.3?

There are three ways that the coverage of the stations overlaps:
(a) all demand points in the triangle are covered by at most one station;
(b) the demand points within the triangle are covered by either one or two stations;
and (c) the demand points within the triangle are covered by up to three stations.
Figure 3.3 shows the special examples of these three situations. FP^Q and FP^{AU} are then compared in these three examples as follows.

- Example (a):

  As shown in Figure 3.3(a), $W(A)$, $W(B)$ and $W(C)$, adjacent with each other, represent the coverage areas of stations A, B and C, respectively. The workload for each station with assignment $\rho^A = \rho$, FP^{AU} is thus calculated according to Equation (3.5) with $o = 0$. On the other hand, FP^Q is always obtained from the left hand side of Inequality (2.7) in all these three examples.

- Example (b):

  As shown in Figure 3.3(b), $W(A)$, $W(B)$ and $W(C)$ represent the areas which can *only* be covered by stations A, B and C, respectively. $W(AB)$ represents the area which can be covered by either station A

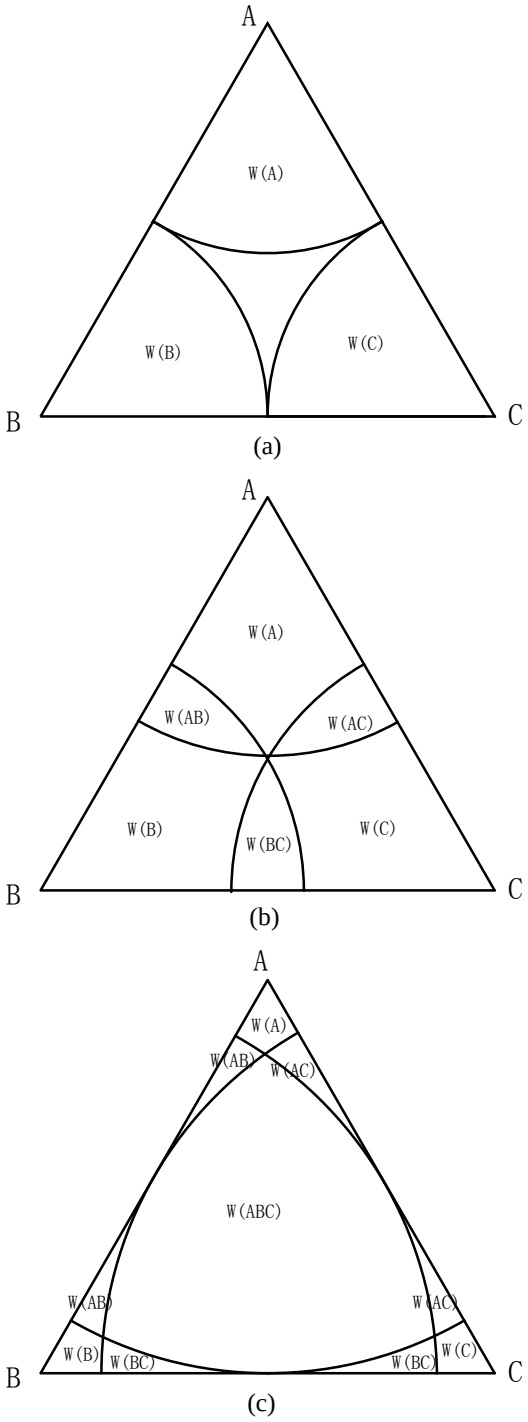

Figure 3.3.. The three examples for the three station case

or B. So are $W(AC)$ and $W(BC)$, respectively. The overlap fraction is

$$o = \frac{W(AB) + W(AC)}{W(A) + W(AB) + W(AC)} = 0.346.$$

In this example, $\rho^A = \rho(1 - 0.5o)$, which is the same as that of the two station case.

- Example (c):

  As shown in Figure 3.3(c), the area $W(ABC)$ can be covered by *all* the stations, while other areas are covered by either one or two stations. In this example, the overlap fraction is calculated as

$$o = \frac{W(AB) + W(AC) + W(ABC)}{W(A) + W(AB) + W(AC) + W(ABC)} = 0.973.$$

  Since each demand point is assigned to the nearest station, ρ^A can be found to be $\frac{2}{\sqrt{3\pi}}\rho$. As a result,

$$FP^{AU} = Pr(R' \geq z), \quad \text{where } R' \sim POISSON(\frac{2}{\sqrt{3\pi}}\rho).$$

  Note that FP^{AU} in the three station case is different from the one in the two station case only when some areas are covered by all the stations.

3.2.3. Comparison between FP^Q and FP^{AU}

We compare FP^Q and FP^{AU} in the two cases discussed in Section 3.2.1 and 3.2.2 with different numbers of vehicles at each station. As indicated before, both FP^Q and FP^{AU} are the upper bounds of the failure probability

with assignment FP^A. Figure 3.4(a) ($\rho = 0.8$) and Figure 3.4(b) ($\rho = 2.3$) show the failure probability estimates FP^Q and FP^{AU}, versus different numbers of vehicles. As presented in these two figures, when the overlap does not exist, i.e., $o = 0$, $FP^Q = FP^A$. As FP^{AU} is an upper bound of FP^A, $FP^{AU} \geq FP^A$. Thus, more vehicles may be needed to meet the failure probability requirement using FP^{AU}, whereas with the increase of the overlap, fewer vehicles are needed using FP^{AU}. In other words, when the overlap is small, it is better to adopt the QRLSCP, i.e., using FP^Q in the availability constraint. However, when the overlap increases, the advantage of using FP^{AU} in the availability constraint increases. Take Figure 3.4(a) as an example, with the overlap fraction of 0.346, 3 vehicles are required if the failure probability target is 15% when using FP^Q in the constraint. In contrast, only 2 vehicles are needed when using FP^{AU} in the constraint. Another observation is that, the advantage of using FP^{AU} is larger when the overlapped area is covered by more stations. As shown in these two figures, when the overlap fraction is 0.973, FP^{AU} in the three station case is always lower than the one in the two station case for any number of vehicles.

Figure 3.5(a) ($\rho = 0.8$) and Figure 3.5(b) ($\rho = 2.3$) show the switch-over curves for FP^Q and FP^{AU}. If the required failure probability and the overlap fraction fall in the upper right area of the curves (area M), using FP^{AU} may yield a better solution with fewer vehicles needed. On the other hand, if the required failure probability and the overlap fraction fall in the lower left area of the curves (area N), using FP^Q may yield a better solution with fewer vehicles needed. Another interesting observation

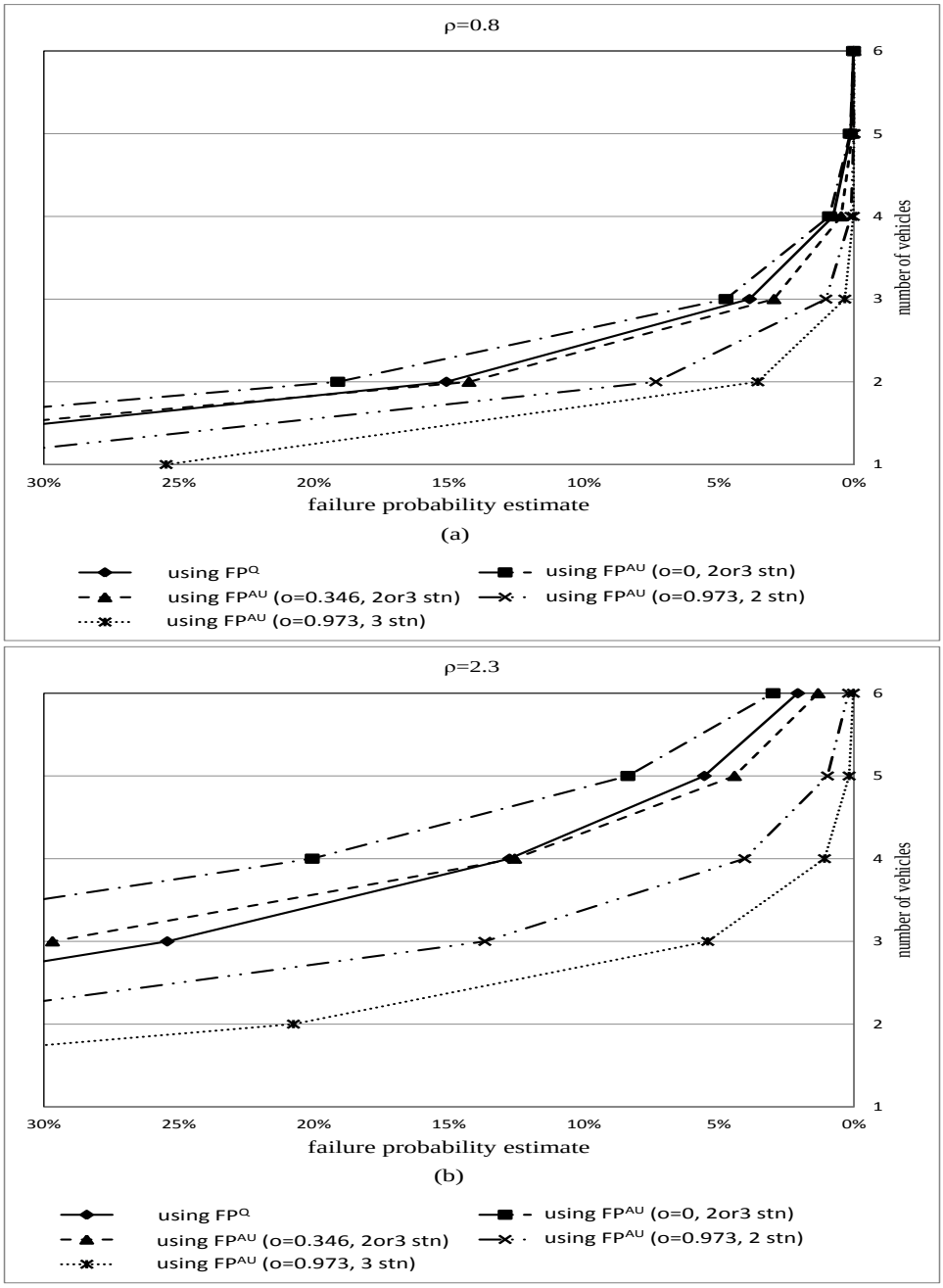

Figure 3.4.. Comparison between FP^Q and FP^{AU} in the two cases

is that, the threshold of the overlap fraction where using FP^{AU} has a better performance increases with the target failure probability. With the same overlap fraction, FP^{AU} will be more favorable when the required failure probability is lower. As can be noticed in Figure 3.5(a), when the overlap fraction is 0.3, using FP^{AU} may result in an inferior solution if the failure probability target is higher than 17%, while it results in a better solution with the commonly used requirement 10%.

It is worth to observe from Figure 3.4 that, the gap between FP^{AU} line when the overlap is 0 and FP^{Q} line is small, and such a gap decreases when the failure probability requirement decreases. This implies that, although using FP^{Q} may require fewer vehicles when the overlap fraction is low, the benefit is not much, i.e., it never exceeds one vehicle even in the worst case when the overlap fraction is 0. Furthermore, such a benefit is likely to decrease when the failure probability requirement decreases. On the other hand, when the overlap fraction is significant or the overlapped area is covered by more stations, using FP^{AU} may have a much better performance than using FP^{Q}. For example, with respect to the failure probability requirement 5% in Figure 3.4(b), 6 vehicles are required when using FP^{Q}, while 4 vehicles are enough when using FP^{AU} in the three station case.

3.3. The traditional failure probability and its estimates

In Section 3.1, we introduce a concept FP^{A} based on the failure probability with assignment, and develop the estimate FP^{AU} which is an upper bound of FP^{A}. We notice that FP^{A} is not always the traditional failure probability in practice, which is the probability that a demand does not see any

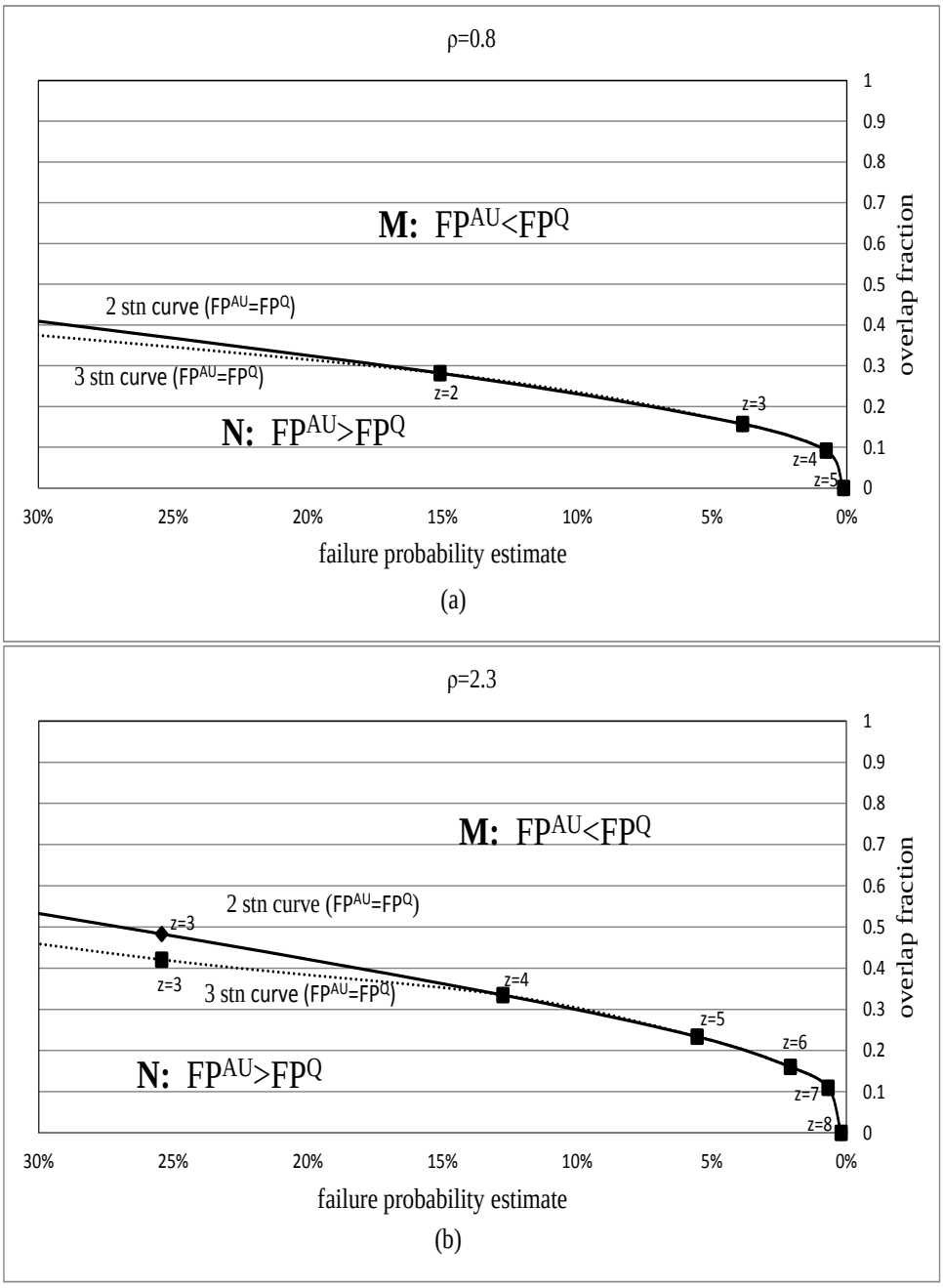

Figure 3.5.. Switch-over curves for FP^{AU} and FP^{Q}

idle vehicle in the neighbourhood as indicated previously. The traditional failure probability widely used in practice is denoted as fp, which is the performance target for the practitioners. Section 3.2 discusses the advantages of using FP^{AU} as an estimate of fp. In this section, we will explore the properties of FP^{A}, FP^{AU} as well as FP^{Q} (used in the QRLSCP) and FP^{B} (used in the BRLSCP), and their relationships with fp. Note that FP^{Q} and FP^{AU} are upper bounds of fp and FP^{A}, respectively.

3.3.1. Two station case: one vehicle located at each station

Our study starts with the simple two station case in Section 3.2.1, where the traditional failure probability can be obtained through Markov Chain analysis. It is assumed that only one vehicle is located at both stations A and B, i.e., $z_A = z_B = 1$. The demand arrival rate within the coverage area of each station is λ. The service time is assumed to be exponential with a rate of 1. Another assumption is that the demand is served by the nearer one if both vehicles are idle.

Consider a Continuous Time Markov Chain (CTMC) with a state space $S = \{ab\},\ a, b = 0$ or 1, where

S_1 = $\{00\}$, denotes the state that both of the two vehicles are idle;

S_2 = $\{01\}$, denotes the state that the vehicle at station A is idle, and the vehicle at station B is occupied;

S_3 = $\{10\}$, denotes the state that the vehicle at station A is occupied, and the vehicle at station B is idle;

Table 3.1.. The transition rate matrix (2 stations, $z = 1$)

$z = 1$	S_1	S_2	S_3	S_4
S_1	$-\lambda(2-o)$	$\lambda(1-0.5o)$	$\lambda(1-0.5o)$	0
S_2	1	$-(\lambda+1)$	0	λ
S_3	1	0	$-(\lambda+1)$	λ
S_4	0	1	1	-2

S_4 = $\{11\}$, denotes the state that both of the two vehicles are occupied.

Then the transition rate matrix is formulated as Table 3.1, which can be explained as follows. The transition rate from state S_1 to S_2 is the demand arrival rate along AC, i.e., $\lambda(1-0.5o)$, since the calls are served by the nearer vehicles when both of the vehicles are available. Similarly, the transition rate from state S_1 to S_3 is also $\lambda(1-0.5o)$. In terms of state S_2, the transition rate to state S_1 is the service rate 1, for the call serving by the vehicle at station B; and the transition rate to state S_4 is the demand arrival rate within the coverage of station A, as the vehicle at station B is already occupied. Likewise, the transition rates for state S_3 can be explained symmetrically. Moreover, both the transition rates from state S_4 to S_2 and S_3 are the service rate 1.

The stationary probability of each state is computed as follows.

$$\begin{aligned} P_{S_1} &= \frac{2}{2\lambda^2 - \lambda^2 o - 2\lambda o + 4\lambda + 2} \\ P_{S_2} &= \frac{\lambda(2-o)}{2\lambda^2 - \lambda^2 o - 2\lambda o + 4\lambda + 2} \\ P_{S_3} &= \frac{\lambda(2-o)}{2\lambda^2 - \lambda^2 o - 2\lambda o + 4\lambda + 2} \\ P_{S_4} &= \frac{\lambda^2(2-o)}{2\lambda^2 - \lambda^2 o - 2\lambda o + 4\lambda + 2}. \end{aligned}$$

The traditional failure probability for station A is $P_{S_3} + P_{S_4}$, which is equivalent to the traditional failure probability for the demand from AC (which is only covered by station A). Symmetrically, fp at station B is equal to fp at station A. Moreover, the failure probability of the demand from the overlapped area (the intersection part of the two circles in Figure 3.2, denoted as $Area_o$) is P_{S_4}. It is obvious that, the traditional failure probability of the demand from $Area_o$ is always lower than that of the demand from AC. As fp is a universal requirement for all demand nodes, in the rest of the study, the traditional failure probability of the demand nodes which can only be covered by one station is discussed.

Under the two station case with $z = 1$, our estimate FP^{AU} is formulated as Equation (3.6).

$$FP^{AU} = 1 - \exp[-\lambda(1 - 0.5o)]. \tag{3.6}$$

It can be shown that FP^{AU} is always higher than the traditional failure probability for the demand nodes from the overlapped area $Area_o$, i.e., $FP^{AU} > P_{S_4}$.

3.3.2. Two station case: two vehicles located at each station

With similar assumptions as the previous example, we only increase the number of vehicles located at each station to 2, i.e., $z_A = z_B = 2$. Correspondingly, the CTMC under this condition should be reconstructed through nine different states, i.e., $S = \{ab\}$, $a, b = 0, 1$ or 2, where $a = 0, 1$ and 2 represent that 0, 1 and 2 vehicles at station A are occupied, respectively. The meaning of b is explained similarly for station B.

The transition rate matrix can be derived similarly as the previous example, and the stationary probabilities are accordingly calculated. The detailed matrix is available in Appendix D. In this case, if the stationary probability of each corresponding state is denoted as P_{ab}, the traditional failure probability for station A is $fp = P_{20} + P_{21} + P_{22}$.

Under the two station case with $z = 2$, our estimate is formulated as Equation (3.7).

$$FP^{AU} = 1 - [\lambda(1 - 0.5o) + 1] \exp[-\lambda(1 - 0.5o)]. \tag{3.7}$$

3.3.3. Three station case: one vehicle located at each station

Next, we further study the traditional failure probability under the three station case with only one vehicle located at each station. With similar assumptions above-mentioned, the CTMC under the three station case should be constructed through eight different states, i.e., $S = \{abc\}$, $a, b, c =$ 0 or 1, where a, b and c represent the states of the vehicles at stations A, B and C, respectively. Furthermore, the states of the vehicles are indicated as: 0 for idle and 1 for occupied.

According to the discussion in the previous section, there are three ways the coverage of the stations overlaps under the three station case (see, Figure 3.3). As a result, the failure probabilities for the three different situations have to be assessed separately.

- Situation (a): $o = 0$ (e.g., Figure 3.3(a)).

  Under this situation, each station operates independently and the traditional failure probability $fp = FP^A$. Therefore, our estimate FP^{AU} is a consistent upper bound of fp according to Proposition 1. The detailed transition matrix can be referred in Appendix E.1.

- Situation (b): $0 < o \leq 0.346$ (e.g., Figure 3.3(b)).

  The detailed transition matrix is available in Appendix E.2. Under this situation, the traditional failure probability for station A is $fp = P_{100} + P_{101} + P_{110} + P_{111}$. Furthermore, our estimate $FP^{AU} = 1 - \exp[-\lambda(1 - 0.5o)]$.

- Situation (c): $0.346 < o \leq 0.973$ (e.g., Figure 3.3(c)).

  The detailed transition matrix is also available in Appendix E.3. Under this situation, the overlapped areas may be covered by two stations or three stations, which has to be addressed differently. Hence, the transition rates cannot be derived directly. The demand arrival rates in the areas $W(ABC)$, $W(AB)$ and $W(A)$ are denoted as a, b and c, respectively. The stationary probability of each state can be obtained in terms of a, b and c. In addition, the overlap fraction is reformulated with respect to a, b and c as follows.

$$o = \frac{a + 4b}{a + 4b + c}.$$

Furthermore, the demand arrival rate in the coverage area of each station is $\lambda = a + 4b + c$. Note that, the values of a, b and c are uniquely determined for a specific λ using the geometry theory.

On the other hand, with assignment, the workload assigned to station A is $\frac{a}{3} + 2b + c$, which is equal to that of stations B and C. Then our estimate FP^{AU} can be calculated using Equation (3.8).

$$FP^{AU} = 1 - \exp[-(a/3 + 2b + c)]. \tag{3.8}$$

3.3.4. Comparison of the traditional failure probability and its estimates

Figure 3.6, 3.7 and 3.8 show the comparison of the traditional failure probability fp and its several estimates FP^A, FP^{AU}, FP^Q and FP^B in different cases above-mentioned. As can be seen, FP^A undershoots the traditional failure probability fp for the demand nodes only covered by one station in the two station case with $z = 1$, almost coincides with fp in the two station case with $z = 2$, and overshoots fp in the three station case with $z = 1$. Especially as an upper bound of FP^A, FP^{AU} almost consistently overshoots the traditional failure probability fp in the cases under our study. The only exception here is in the two station case with $z = 1$, where FP^{AU} undershoots fp when the overlap fraction approaches to 1. However, FP^{AU} more consistently overestimates fp when the number of vehicles or stations increases (see, Figure 3.7 and 3.8).

Another point to be mentioned is that, using FP^Q (or FP^B) may require fewer vehicles than using FP^{AU} when FP^Q (or FP^B) is lower than FP^{AU},

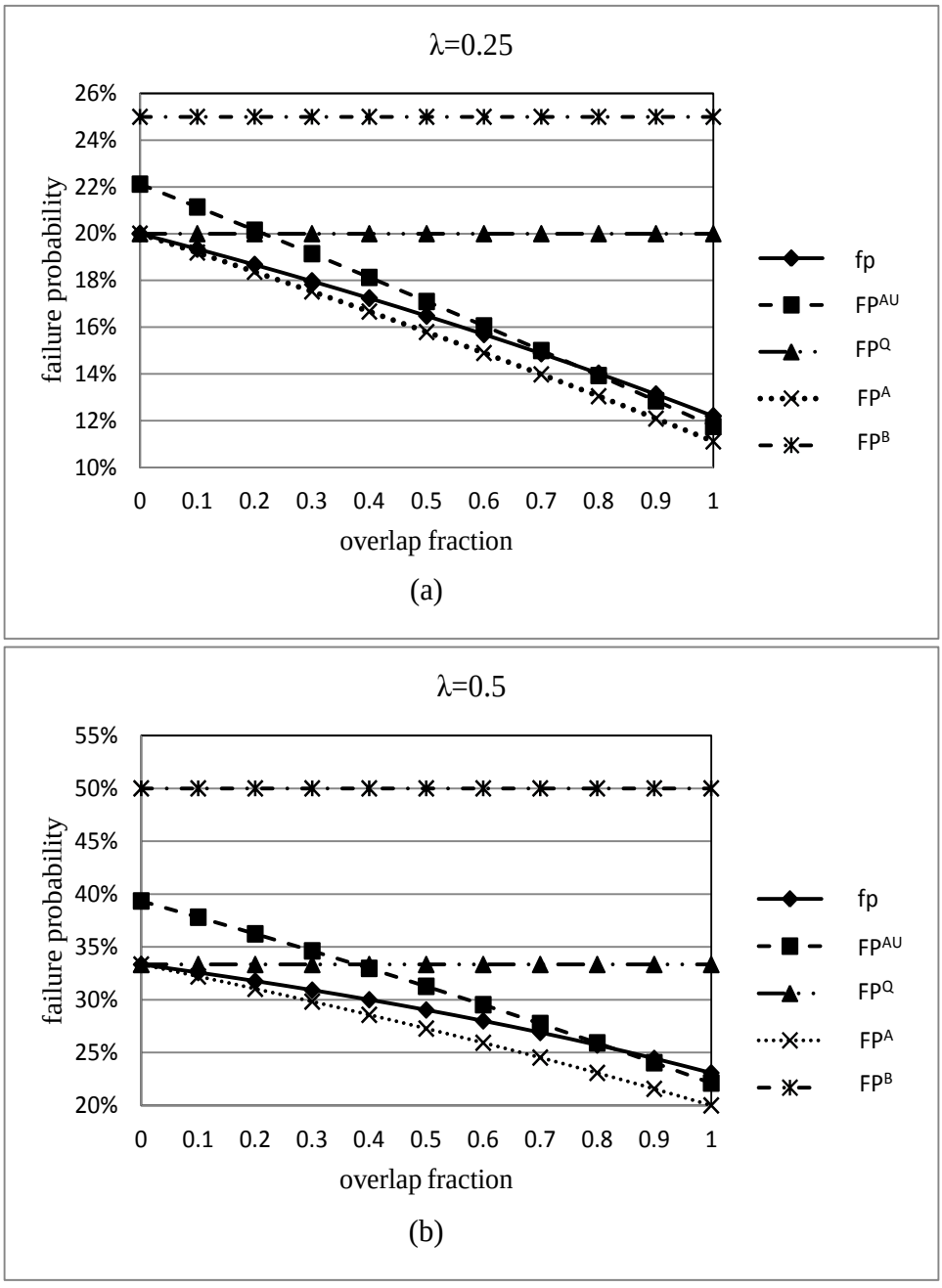

Figure 3.6.. Comparison of fp and its estimates (2 station case, $z = 1$)

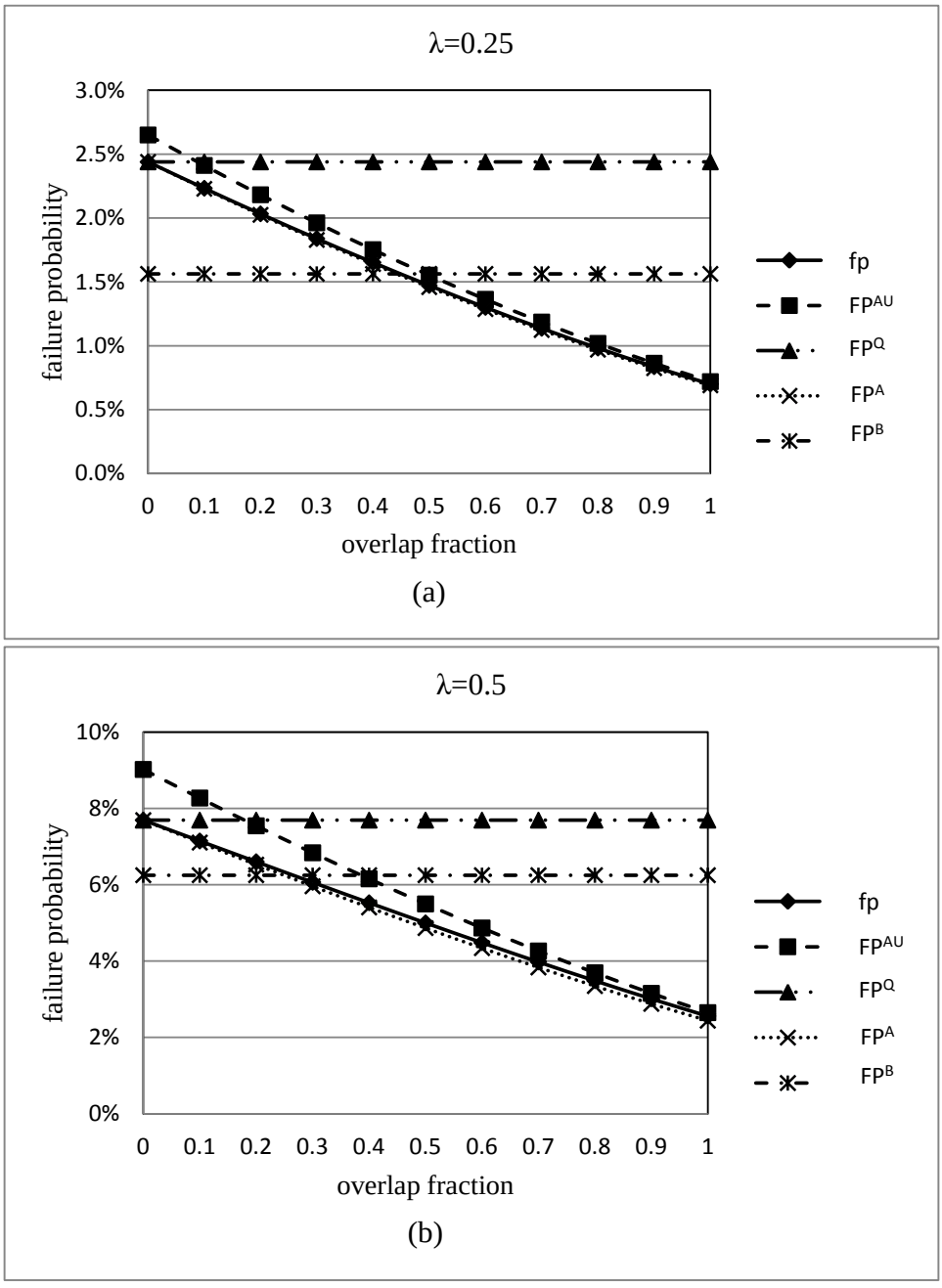

Figure 3.7.. Comparison of fp and its estimates (2 station case, $z = 2$)

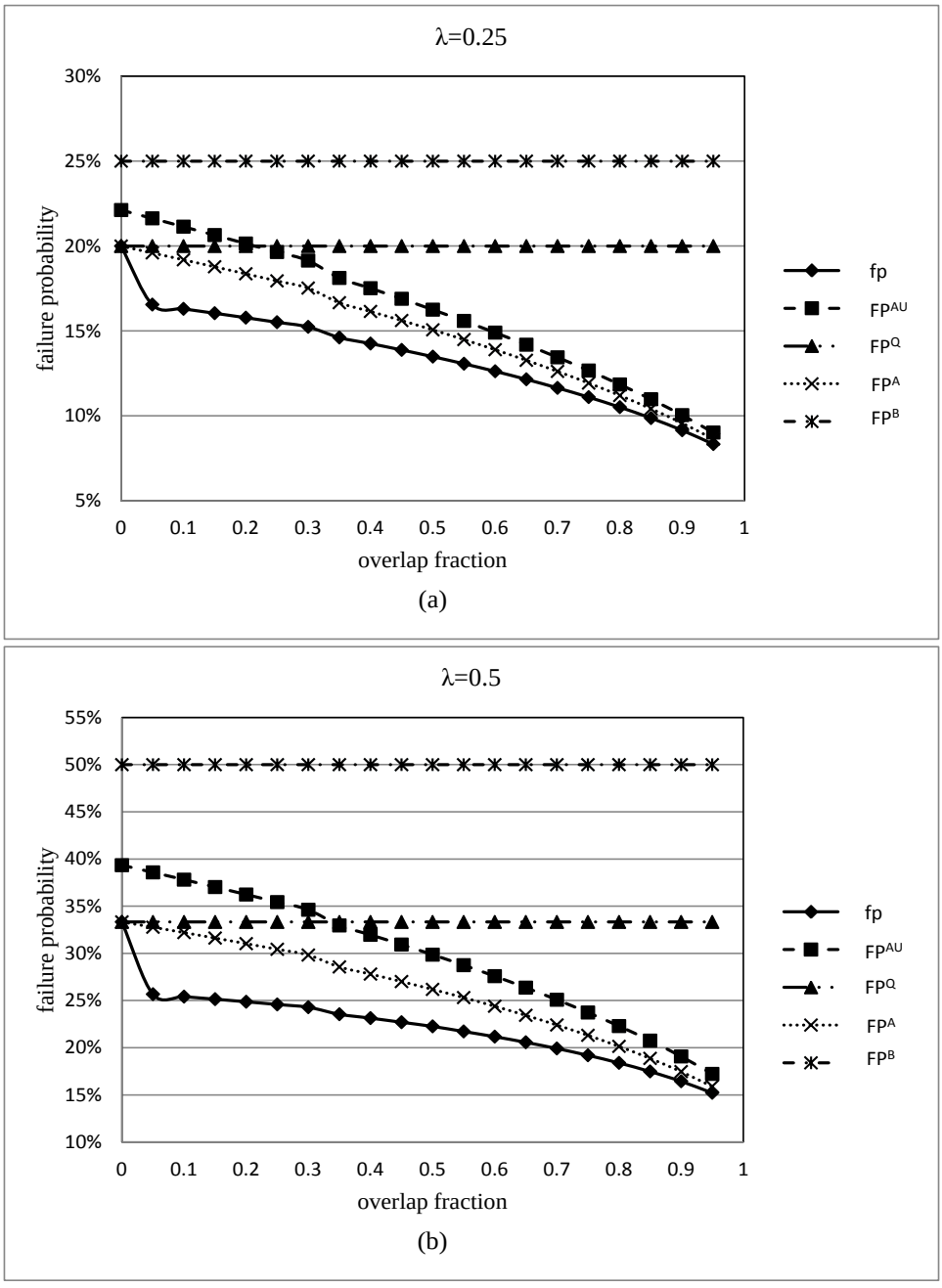

Figure 3.8.. Comparison of fp and its estimates (3 station case, $z = 1$)

respectively. This usually occurs under the condition when the overlap fraction is small. The comparison between the results of using FP^Q and FP^{AU} has been discussed in Section 3.2. In terms of FP^B, although it may result in fewer vehicles required, it may also undershoot the traditional failure probability fp for the demand nodes only covered by one station when the overlap is small. Note that when the overlap is small, most demand nodes can only be served by one station. In other words, in this case, the solutions of the set covering problem using FP^B in the availability constraint may not meet the failure probability requirement for most demand nodes.

In the three cases under our study, our estimate FP^{AU} consistently overshoots the traditional failure probability, where the only exception is in the two station case with only one vehicle located at each station. We then examine the worst case of the undershooting of FP^{AU}. We define the difference between FP^{AU} and fp as function $G = FP^{AU} - fp$. It can be shown that function G is decreasing with the overlap fraction o. Hence, FP^{AU} undershoots fp the most when $o \to 1$. Function G in the worst case of undershooting can be formulated as follows.

$$\begin{aligned} \lim_{o \to 1} G &= 1 - \exp[-0.5\lambda] - (P_{S_2} + P_{S_4}) \\ &= 1 - \exp[-0.5\lambda] - \frac{\lambda^2 + \lambda}{\lambda^2 + 2\lambda + 2}. \end{aligned}$$

Note that, when $o \to 1$, $\lambda_A \to 0$ (see, Figure 3.2), implying that the failure probabilities of very few demand nodes are undershot by FP^{AU}.

Figure 3.9 shows the maximum amount FP^{AU} undershoots fp in the two station case with $z = 1$. For different traditional failure probabilities, the

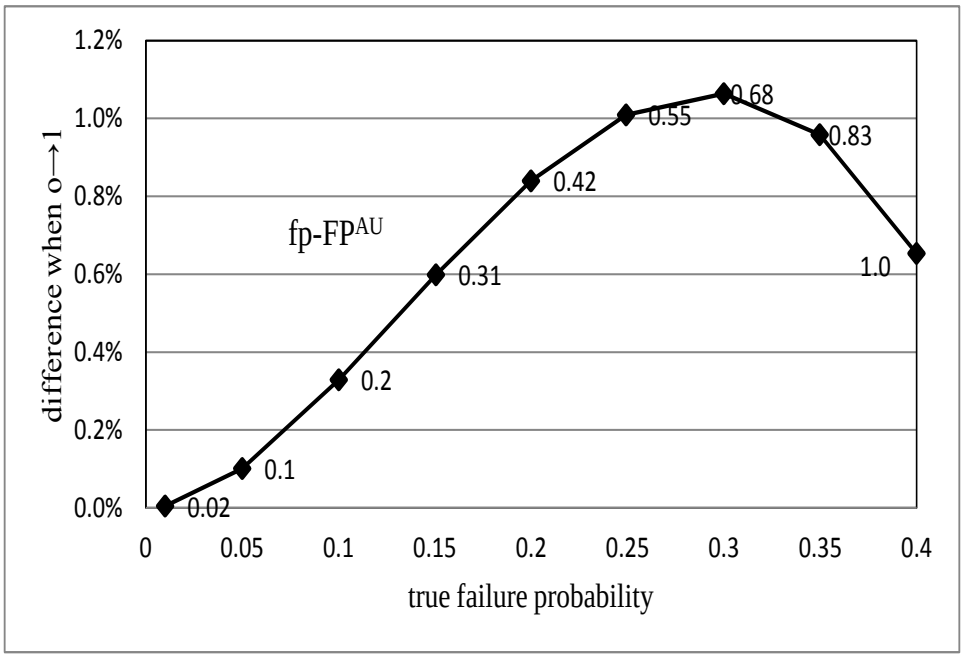

Figure 3.9.. Worst case of $fp - FP^{AU}$ when $o \to 1$

corresponding demand arrival rate and then the differences of FP^{AU} and fp are calculated. The numbers close to each node denote the corresponding demand rates. It is observed that FP^{AU} undershoots fp by no more than 1.06% (for very few demand nodes as indicated previously). In addition, we notice that such an amount (1.06%) that FP^{AU} undershoots fp occurs when the traditional failure probability is about 30%. For the commonly used 10% failure probability requirement, FP^{AU} only undershoots fp by 0.33%.

In short, FP^{AU} is almost a consistent upper bound of the traditional failure probability. The only exception is in the two station case with only one vehicle at each station. However, FP^{AU} undershoots the traditional failure probability by no more than 1.06% when the overlap fraction approaches to 1. In addition, for the commonly used 10% failure probability requirement, the undershooting is even nearly negligible. Furthermore, FP^{AU} is

a more consistent overestimate when the number of vehicles or stations increases. This conservative property is critical when FP^{AU} is adopted in the availability constraint of the set covering problem.

3.4. Summary

This chapter states the problem in the EMS systems, which focuses on exploring the pros and cons of demand assignment.

Section 3.1 defines a failure probability FP^A which is used to derive our model with assignment. We also show that FP^A is always less than the failure probability used in the QRLSCP. That means, using FP^A in the model will always lead to fewer vehicles required. Moreover, we derive an upper bound for FP^A, which is denoted as FP^{AU}, and demonstrate that this bound asymptotically approaches to FP^A whether the workload is quite large or extremely small.

Next, we investigate the advantages of demand assignment by two simple case studies, namely, the two station case and the three station case. It is found that when the overlap among the coverage areas is significant, or when the demand is covered by more stations, the model with FP^{AU} requires fewer vehicles than the QRLSCP.

In Section 3.2, FP^{AU} is demonstrated to be almost a consistent upper bound of the traditional failure probability, only except for the two station case with only one vehicle at each station. However, the maximum underestimation of FP^{AU} compared with the traditional failure probability is no more than 1.06%. This important conclusion indicates that the results of using FP^{AU} in the availability constraint would be viable in practice.

Chapter 4 MATHEMATICAL MODEL WITH DEMAND ASSIGNMENT

In the previous chapter, we demonstrate that FP^{AU} is almost a consistent upper bound of the traditional failure probability fp. Also, the maximum underestimation of FP^{AU} compared with the traditional failure probability is no more than 1.06%, which can be easily adjusted. This important conclusion supports us to use FP^{AU} in the availability constraint. This chapter first constructs the mathematical model based on the analysis in the previous chapter, and then develop a tricky approach to transform the model to a linear one, which can be solved by current solvers like $CPLEX$.

4.1. Mathematical model

According to the analysis in sections 3.1, 3.2 and 3.3, we can build up a LSC model using FP^{AU} in the availability constraint. This new model is referred to as Location Set Covering Problem with Assignment (LSCPA). As indicated previously, in most cases we can be almost sure that the required availability is achieved, and provide better solutions. The LSCPA

is proposed as follows.

$$
\begin{aligned}
LSCPA \qquad \min_{x_i, z_{i,k}, y_{ij}} c &= \sum_{i \in I} (c_1 x_i + c_2 z_i) && (4.1)\\
s.t.\ Pr\Big(\sum_{j \in S_{PT}(i)} D_j \big(\frac{1}{\mu_{ij}}\big) y_{ij} \geq z_i \Big) &\leq 1 - p_1, \forall\ i \in I, && (4.2)\\
z_i &\leq u_i x_i, \forall\ i \in I, && (4.3)\\
\sum_{i \in S_{COV}(j)} y_{ij} &= 1, \forall\ j \in J, && (4.4)\\
x_i, y_{ij} &\in \{0,1\}, \forall\ i \in I, \forall\ j \in J,\\
z_i, &\in \mathbf{Z}^+ \cup \{0\}, \forall\ i \in I,
\end{aligned}
$$

where c_1 and c_2, both assumed to be constants, are the unit costs for each station and each vehicle, respectively. x_i is the binary variable denoting whether Station i is selected.

In the LSCPA, x_i, the assignment variable y_{ij}, and the fleet size at each station z_i are the decision variables. Objective function (4.1) is to minimize the total cost, which is the sum of the station cost $\sum_{i \in I}(c_1 x_i)$ and the maintenance cost of the vehicle fleet $\sum_{i \in I}(c_2 z_i)$. As discussed previously, the left hand side of Inequality (4.2) is almost a consistent upper bound of the traditional failure probability, and so using (4.2) as a constraint, in turn, will guarantee that the traditional failure probability is below the requirement $1 - p_1$. Inequality (4.3) restricts the maximum number of vehicles that can be located at station i cannot exceed u_i. Equation (4.4) is a general assignment rule, which means that each demand node is assigned to only one station when y_{ij} is binary. Note that when y_{ij} is relaxed to be a real number in the next chapter, this constraint would mean that the

sum of the fractions of each demand assigned to all possible stations is 1.

4.2. Linearization transformation

As discussed in Section 1.1, the introduction of the assignment variable y_{ij} may result in a large model size. In addition, the LSCPA itself is a Mixed Integer and Non Linear Program (MINLP), and thus the computational time grows exponentially with the model size. Additionally, computation of the Cumulative Distribution Function (CDF) of a Poisson distribution is required in Constraint (4.2). The models with such a constraint cannot be addressed by current solvers (e.g., $CPLEX$, $BARON$) because of the unknown decision variable y_{ij} inside the expression of the Poisson parameter, i.e., ρ_i^A. That is, according to Section 3.1.2, $Pr(\sum_{j \in S_{PT}(i)} D_j(\frac{1}{\mu_{ij}})y_{ij} \geq z_i)$ can be uniformized to $Pr(\widehat{D}_i \geq z_i)$, where $\widehat{D}_i$ is a Poisson random variable denoting the arrivals within one unit of time with rate $\rho_i^A = \sum_{j \in S_{PT}(i)} \frac{\lambda_j}{\mu_{ij}} y_{ij}$, where y_{ij} is unknown decision variable, which makes the model cannot be solved by standard solvers.

We then develop a tricky approach to make the model to become a linear one, which are then solvable by current solvers like $CPLEX$. The basic idea is that we compute the maximum workload which can be served by each station with a specific fleet size in advance. In particular, let w_{ik} denote the maximum workload which can be served by Station i if k vehicles are located there, it can be calculated before solving the model as in Appendix F. Another binary decision variable z_{ik} is introduced to represent whether or not k vehicles are located at Station i. Then the model $LSCPA$ can be

transformed to the following $T-LSCPA$.

$$
\begin{aligned}
T-LSCPA \qquad \min_{x_i,z_{ik},y_{ij}} c &= \sum_{i\in I}\sum_{k=0}^{u_i}(c_1x_i + c_2kz_{ik}) \\
s.t. \sum_{j\in S_{PT}(i)} \frac{\lambda_j}{\mu_{ij}} y_{ij} &\leq \sum_{k=1}^{u_i} z_{ik}w_{ik}, \forall\ i \in I, \qquad (4.5)\\
\sum_{k=0}^{u_i} z_{ik} &= 1, \forall\ i \in I, \qquad (4.6)\\
\sum_{k=1}^{u_i} z_{ik} &= x_i, \forall\ i \in I, \qquad (4.7)\\
\sum_{i\in S_{COV}(j)} y_{ij} &= 1, \forall\ j \in J,\\
x_i, y_{ij}, z_{ik} &\in \{0,1\}, \forall\ i \in I, \forall\ j \in J, \forall k = 1, ...u_i
\end{aligned}
$$

Where Equation (4.6) means a specific number of vehicles should be located at each station. $z_{i0} = 1$ means no vehicles are located at Station i, that is, Station i is not activated; Equation (4.7) means we can locate vehicles at a station only when it is activated.

The proof of the equivalence of Model $T-LSCPA$ and the original model $LSCPA$ can be found in Appendix G. Then z_i in Model $LSCPA$ is substituted by $\sum_{k=0}^{u_i} kz_{ik}$ from Model $T-LSCPA$, while y_{ij} and x_i take the same values as in Model $T-LSCPA$. We can see that the original non-linear mathematical model is transformed to a pure linear integer program, which can be solved by common solvers like $CPLEX$.

4.3. Heuristic

The linear integer model $T-LSCPA$ belongs to the class of NP-hard optimization problems. The computation time to solve this model will

exponentially increase with the number of decision variables, therefore it may require great computation efforts for solving large scale problems to achieve the global optimal. For example, the efficiency of $CPLEX$ in solving the proposed model may be quite unstable, which is dependent on whether or not the LP relaxation and cutting plane added at the nodes are tight. That is, generally it is difficult to predict the efficiency of solving the model in advance. Regarding the real case study in the next chapter, the optimality gap for some instances will remain to be larger than 4% after 1e+10 iterations.

On the other hand, some flexible constraints exist in the model $T-LSCPA$, like the availability constraint (4.5). The exact algorithm (like branch-and-cut algorithm used by CPLEX) needs to resolve the model even in case of only a little bit change to one such flexible constraint. This may be very time consuming and could be expensive for real applications. We are motivated to develop a heuristic which is not so sensitive to little changes in some parameters. Furthermore, our heuristic is decomposed into two parts which are implemented consecutively. So changes to some parameters may only lead to resolving one part instead of totally resolving the entire problem. With similar arguments, the heuristic is also helpful when the vehicles have to be located dynamically.

The heuristic is decomposed into two parts: Subroutine A is to decide which stations to be selected, and Subroutine B is to decide the fleet size at each station. Subroutine A is identical with the original LSC problem without the service constraint, that is, select the least number of stations in order that all demand nodes can be covered by at least one station nearby.

We adopt the algorithm provided by Toregas and Revelle (1973) to reduce the problem size. In Subroutine B, we develop a heuristic to decide the fleet size at each station.

Subroutine A: Decide which stations to be selected.

- Step $A.1$: Initialization. $J_r = J; I_s = \phi; I_c = I$
- Step $A.2$: Define the set of the essential stations as $I_e = \{i \in S_{COV}(j)\ \&\ |S_{COV}(j)| = 1, \forall j \in J_r\}$. Then

$$\begin{aligned} I_s &= I_s + I_e; \\ I_c &= I_c - I_e; \\ J_r &= J_r - \{j \in S_{PT}(i) \cap J_r | i \in I_e\}. \end{aligned}$$

- Step $A.3$: Define Node j_1 is dominated by j_2 if $\{S_{COV}(j_1) \cap I_c\} \supseteq \{S_{COV}(j_2) \cap I_c\}$. Update J_r as

$$J_r = J_r - \text{all dominated nodes in } J_r.$$

- Step $A.4$: Define Station i_1 is dominated by i_2 if $\{S_{PT}(i_1) \cap J_r\} \subseteq \{S_{PT}(i_2) \cap J_r\}$. Update I_c as

$$I_c = I_c - \text{all dominated stations in } I_c.$$

- Step $A.5$: After steps $A.1$-$A.4$, the remaining work is to decide the least number of stations in set I_c to cover all nodes in J_r. The branch and bound approach can be applied to solve the reduced problem. The stations selected in the solution are added to set I_s, which is the set of candidate stations we will select.

We can see that the time complexity for Step $A.2$ and Step $A.5$ is $O(|I| \log |J|)$, and the time complexities for Steps $A.3$ and $A.4$ are $O(|J|)$

and $O(|I|)$, respectively. Thus Subroutine A can be accomplished in $O(|I| \log |J| + |J|)$ time.

After Subroutine A, we develop Subroutine B to decide the number of vehicles required at each station.

Subroutine B: Decide the fleet size at each station. Subroutine B is implemented twice. For the first time, I_o is initialized as I_s; and for the second time, I_o is initialized as I. We compare the total costs required for these two results and select the one that leads to the lower cost.

- Step $B.1$: Compute the heaviest workload which can be assigned to each station, denoted as $\overline{\rho_i}$ by solving $1 - p_1 = FP^{AU}(\overline{\rho_i}, u_i)$, $\forall i \in I_o$, the detailed algorithm can be seen in Function "find the threshold workload" in Appendix F.

- Step $B.2$: Initialization for the second subroutine.

  Set an iteration index, $v = 1$; initialize the assigned workload for each station as $\rho_i^v = 0, \forall i \in I_o$; $J_v = J$; and the demand nodes assigned to each station as $S_{PT}(i) = \phi, \forall i \in I_o$;

- Step $B.3$: Decide the preferred station for the demand nodes.

  - Step $B.3.1$: For $j = \arg\max_{j \in J_v} \ \lambda_j$, set another iteration index, $w_j = 1$;

  - Step $B.3.2$: Find the preferred station for Node j, denoted as

    $$N_j^{w_j} = \arg \min_{\{i \in S_{COV}(j) \cap I_o \ \& \ i \notin \cup_{k=1}^{w_j - 1} N_j^k\}} dis_{ij}.$$

    where dis_{ij} denotes the distance from node i to node j, which indicates that the nearest activated station will be the preferred station for each node.

– Step $B.3.3$: If $\rho^{v}_{(N_j^{w_j})} + \frac{\lambda_j}{\mu_{(N_j^{w_j})j}} > \overline{\rho_{(N_j^{w_j})}}$, $w_j = w_j + 1$, then go to step $B.3.2$. Otherwise,

$$S_{PT}(N_j^{w_j}) = S_{PT}(N_j^{w_j}) + j; v = v + 1;$$

$$\rho^{v}_{(N_j^{w_j})} = \rho^{v-1}_{(N_j^{w_j})} + \frac{\lambda_j}{s_{(N_j^{w_j})j}}; J_v = J_{v-1} - j.$$

If $J_v \neq \phi$, then return to step $B.3.1$;

- Step $B.4$: For each station $\{i \in I_o | S_{PT}(i) \neq \phi\}$, decide the fleet size by solving $1 - p_1 = FP^{AU}(\sum_{j \in S_{PT}(i)} \frac{\lambda_j}{\mu_{ij}}, z_i)$

  Otherwise, for other stations $z_i = 0$;

- Step $B.5$: $z_i{}^* = \lceil z_i \rceil$ for all $i \in I$. Stop.

We can see that the time complexity for Step $B.1$, Step $B.4$ and Step $B.5$ is $O(|I|)$, and the time complexity for Steps $B.3$ from $B.3.1$ to $B.3.3$ is $O(|I||J| \log |J|)$. Thus Subroutine B can be accomplished in $O(|I||J| \log |J|)$ time. As a conclusion, the computation complexity for the entire heuristic including the two subroutines is $O(|I||J| \log |J|)$.

We can always implement Subroutine B twice with I_o initialized differently as indicated above to find the one that leads to the lower cost. We observe that when the value of c_1/c_2 is small, initializing I_o with I usually needs a lower cost. The reason lies in that, when the station cost is lower, more candidate stations may be selected and each demand may find a superior station with a higher service rate. As a result, the fleet size (and the vehicle cost) can be reduced. On the other hand, when the value of c_1/c_2 is large, initializing I_o with I_s usually provides a better solution.

We are interested in the optimality gap of the heuristic. Let Function $Fleet(w)$ be the required fleet size to cover the demand nodes if the total

workload is w, which can be obtained by solving

$$1 - p_1 = FP^{AU}(\sum_{j \in S_{PT}(i)} \frac{\lambda_j}{\mu_{ij}}, z_i)$$

as in Step $B.4$. We need some notations for our further analysis on the optimality gap of the heuristic.

k^H = number of stations activated from the heuristic results;

k^* = number of stations activated according to the exact optimal solutions;

I^H = the set of stations activated from the heuristic results ;

I^* = the set of stations activated according to the exact optimal solutions;

y_{ij}^H = the assignment variable determined by the heuristic results;

y_{ij}^* = the assignment variable determined by the exact optimal solutions;

c^H = the objective value from the heuristic results;

c^* = the exact optimal objective value.

Lemma 3. *The optimality gap of the proposed heuristic is only dependent on the distribution of service rates throughout all demand nodes. In particular, if all demand nodes share the same service rate, the optimality gap tends to 0.*

Proof.

$$\begin{aligned} c^H &= c_1 k^H + c_2 \sum_{i \in I^H} Fleet(\sum_{j \in S_{PT}(i)} \lambda_j y_{ij}^H) \\ &= c_1 k^H + c_2 [\sum_{i \in I^H \cap I^*} Fleet(\sum_{j \in S_{PT}(i)} \lambda_j y_{ij}^H) + \sum_{i \in I^H \setminus I^*} Fleet(\sum_{j \in S_{PT}(i)} \lambda_j y_{ij}^H)] \end{aligned}$$

$$\begin{aligned} c^* &= c_1 k^* + c_2 \sum_{i \in I^*} Fleet(\sum_{j \in S_{PT}(i)} \lambda_j y_{ij}^*) \\ &= c_1 k^* + c_2 [\sum_{i \in I^H \cap I^*} Fleet(\sum_{j \in S_{PT}(i)} \lambda_j y_{ij}^*) + \sum_{i \in I^* \setminus I^H} Fleet(\sum_{j \in S_{PT}(i)} \lambda_j y_{ij}^*)] \end{aligned}$$

As our heuristic always activate the fewest number of vehicles and assign each demand to the activated station which incurs least workload, so

$$\begin{aligned} c^H \leq\ & c_1 k^* + c_2 \sum_{i \in I^H \cap I^*} Fleet(\sum_{j \in S_{PT}(i)} \lambda_j y_{ij}^*) \\ +\ & k c_2 \sum_{i \in I^* \setminus I^H} Fleet(\sum_{j \in S_{PT}(i)} \lambda_j y_{ij}^*) \end{aligned}$$

where $k \geq 1$ is a characteristic parameter dependent on the distribution of service rate throughout all demand nodes. If all demand nodes share the same service rate, $k = 1$ means that the optimality gap tends to 0. This parameter also increases if the deviation of the service rates of demand is larger. □

However, we assume that the demand can only be covered by the stations in its coverage. The service rate generally does not deviate too much, which means that the optimality gap of the proposed heuristic is acceptable. The performance of the heuristic will also be tested by a series of experiments in the next chapter.

It can also be noticed that in our heuristic, the workload assigned to

each station $\sum_{j \in S_{PT}(i)} \lambda_j s_{ij} y_{ij}$ is pooled to determine the fleet size of each station, after that the ceiling number of the results is taken as the final solution. Therefore, the aggregation of $\lambda_j s_{ij}$ and the ceiling number of the results make the fleet size not so sensitive to both the availability requirement p_1 and $\lambda_j s_{ij}$. This insensitivity is also demonstrated by the numerical studies in the next chapter. On the other hand, we can see that Subroutine A is independent of such parameters. That is, even if we want to re-run the heuristic in case of changes to the parameters, we only need to resolve Subroutine B. The merit of this robustness is that if these parameters collected are not so accurate or these parameters change at some time, this is the usual case in many EMS systems, the system designed by the heuristic is still a good guidance for the managers.

4.4. Summary

This chapter focuses on developing the mathematical model for the demand assignment problem and proposing an approach to transform the model to be a pure integer program.

Section 4.1 proposes a mathematical model which uses FP^{AU} in the availability constraint. However, the introduction of a large number of assignment variables enlarges the model size. Furthermore, the non-linear model itself and the unknown decision variables inside the Poisson parameter should be addressed to make it solvable. Thus, Section 4.2 develops a tricky way to transform the original mathematical model to a pure integer program. Because the computational time to solve the model $T-LSCPA$ may still be quite long, we are motivated to propose a heuristic in Sec-

tion 4.3, and we also demonstrate its performances in terms of accuracy as well as insensitivity. The performance of the heuristic will be tested in the numerical studies in the next chapter.

Chapter 5 EXPERIMENTAL RESULTS AND DISCUSSIONS ON THE T-LSCPA

In this chapter, we further compare the results of our model (the T-LSCPA) with the BBKK1, QRLSCP and BRLSCP to study the pros and cons of the demand assignment, through the frequently used 55-node problem (see, Swain 1974, Church and ReVelle 1974, Daskin 1983, Church and Weaver 1986, Ball and Lin 1993). We also apply our model to a real case study for a known city S with 9,397 demand nodes and 47 candidate stations. The data of the 55-node problem are shown in Appendix H. Based on the results derived from the experimental results, we further provide a complete discussion on the advantages of the T-LSCPA over other models.

5.1. Case study 1: 55-node problem

We build up all the models, the BBKK1, T-LSCPA, the QRLSCP and the BRLSCP in the optimization software $AIMMS$3.14. $CPLEX$12.5 is selected to solve the Mixed Integer Program (MIP) part.

Following the parameter settings in Church and ReVelle (1974) and Ball and Lin (1993), we let the coverage radius (i.e., the radius of set $S_{PT}(i) : \forall i \in I$) to be 5, 10, and 15. The coverage radius corresponds to the required maximal response (travel) time. The service time is assumed to be the total travel time between the dispatching vehicle and the demand nodes, plus 30

Table 5.1.. The computation time for the 27 instances

computation time (s)	< 30	$30 \sim 70$	$70 \sim 200$	$200 \sim 1000$	$> 1,000$
No. of instances	11	9	3	2	2

Note: The computation time is calculated for the models running on the 3.47GHZ/24.0GB computer

minutes in preparation and processing. We also let the required availability p_1 to be 80%, 90%, and 95%, $c_1/c_2 = 10,\ 5,\ 0$. Note that $c_1/c_2 = 0$ implies that there is no cost for choosing new stations. In summary, we examine a total of 27 instances with 3 coverage radii, 3 availability levels, and 3 different values of c_1/c_2.

The computation time of solving the model T-LSCPA by $CPLEX12.5$ for these 27 instances is shown in Table 5.1. We can see that most instances (20 instances out of 27 instances) can be solved within 70s, i.e., the global optimal can be achieved. However, there are also 2 instances cannot be finished within 1,000s. In particular, the optimality gaps are 2.4% and 3.9% which cannot be reduced even after 1e+6 iterations. We keep the generated solutions (with gaps of 2.4% and 3.9%) as the final results for these two instances in our subsequent analysis. The unstable performance of $CPLEX$ is due to whether or not the LP relaxation and cutting plane added at the nodes are tight, generally it is difficult to predict the efficiency in advance. This is also the reason why we need to develop our heuristic.

As the performance of $CPLEX12.5$ when solving our model is not quite stable in regards of the computation time, we are motivated to develop our heuristic as in Section 4.3 especially for large scale problems. All these 27 instances can be completely solved by our heuristic within 10s with the

optimality gap less than 4.5% for those 25 instances whose optimal can be found by $CPLEX12.5$, and with the same optimality gaps for the two instances whose optimal cannot be obtained by $CPLEX12.5$ within 1e+6 iterations, which is acceptable in practice. We define TC_H^* and TC_L^* as the minimal total costs of the heuristic and the T-LSCPA, respectively, the optimality gap of the heuristic is defined as $g^H = \frac{TC_H^* - TC_L^*}{TC_L^*}$. Note that for the two instances where $CPLEX$ also cannot generate optimal solutions within acceptable computation time, we use the lower bounds provided by $CPLEX$ to calculate the optimality gaps.

Now we focus on the discussion on the benefits of demand assignment of our model. We consider the total cost obtained from the QRLSCP solution as a baseline, as QRLSCP usually results in the highest costs among all the models. Take the T-LSCPA as an example of comparison: We define TC_Q^* as the minimal total costs of the QRLSCP. The cost saving for the T-LSCPA, denoted as ϵ^L, is calculated by $\epsilon^L = \frac{TC_Q^* - TC_L^*}{TC_Q^*}$. Note that we use the solutions generated by $CPLEX$ to calculate the cost saving of Model T-LSCPA for all 27 instances, including the two where $CPLEX$ cannot generate optimal solutions within acceptable computation time. The cost saving for other models (the BBKK1, BRLSCP and the relaxed model to be discussed later) are defined in a similar way as well. Table 5.2 shows the cost saving of all the 27 instances. We observe that the T-LSCPA consistently leads to a significant saving, and the saving is consistently higher than the BBKK1. Furthermore, we have the following observations for the T-LSCPA:

1. The cost saving increases with the coverage radius. This observation

corresponds to our analysis in the previous sections. As mentioned, since a larger coverage area implies a higher overlap fraction, the T-LSCPA leads to higher cost saving. On the other hand, the cost saving for BBKK1 over the QRLSCP is quite stable over different radii as both have the same availability estimator. As indicated previously, the BBKK1 just further optimizes on the decision of station selection.

2. The cost saving of the T-LSCPA is consistently greater than that of the BBKK1, and such cost saving is also higher when the overlap is larger. The reason is similar with the comparison between the T-LSCPA and the QRLSCP.

3. The cost saving is larger when the availability requirement is high, which is consistent with the analysis in Section 3.1.1. The benefits of the demand assignment are more significant under a higher availability requirement.

4. The smaller the value of c_1/c_2, the larger the cost saving. The reason is that: When the station cost is lower, more stations may be selected to host ambulances. Each demand node may have a higher chance to be covered by more than one station. According to the discussion in Section 3.2, the advantage of the T-LSCPA over the QRLSCP or the BBKK1 increases as the overlap of the coverage areas increases.

Table 5.2.: Cost saving compared with the QRLSCP

$p_1 = 80\%; c_1/c_2 = 10$

	ϵ^{B1}	ϵ^{L}	ϵ^{R}	ϵ^{H}	ϵ^{B2}	g^{H}
radius=5	1.1%	3.2%	3.9%	1.8%	4.8%	1.4%
radius=10	1.8%	4.7%	5.2%	2.7%	4.4%	2.0%
radius=15	1.5%	7.2%	8.0%	5.2%	4.7%	2.1%

$p_1 = 80\%; c_1/c_2 = 5$

	ϵ^{B1}	ϵ^{L}	ϵ^{R}	ϵ^{H}	ϵ^{B2}	g^{H}
radius=5	1.6%	4.3%	5.2%	4.3%	6.1%	2.4%*
radius=10	1.5%	5.5%	6.7%	3.3%	5.7%	2.3%
radius=15	1.8%	7.9%	9.0%	5.7%	5.9%	2.4%

$p_1 = 80\%; c_1/c_2 = 0$

	ϵ^{B1}	ϵ^{L}	ϵ^{R}	ϵ^{H}	ϵ^{B2}	g^{H}
radius=5	1.3%	6.7%	8.1%	6.7%	8.5%	3.9%*
radius=10	1.7%	8.3%	10.4%	5.2%	8.6%	3.4%
radius=15	1.6%	9.6%	12.1%	7.3%	8.8%	2.5%

$p_1 = 90\%; c_1/c_2 = 10$

	ϵ^{B1}	ϵ^{L}	ϵ^{R}	ϵ^{H}	ϵ^{B2}	g^{H}
radius=5	1.5%	4.0%	4.6%	2.1%	5.2%	2.0%
radius=10	1.4%	5.2%	6.1%	2.9%	5.2%	2.4%
radius=15	1.2%	8.1%	8.9%	6.3%	5.7%	2.0%

$p_1 = 90\%; c_1/c_2 = 5$

	ϵ^{B1}	ϵ^{L}	ϵ^{R}	ϵ^{H}	ϵ^{B2}	g^{H}
radius=5	1.7%	4.6%	5.4%	3.4%	6.7%	1.3%
radius=10	1.8%	5.9%	6.9%	4.3%	6.9%	1.7%
radius=15	1.8%	8.3%	9.2%	6.0%	6.6%	2.5%

$p_1 = 90\%; c_1/c_2 = 0$

	ϵ^{B1}	ϵ^{L}	ϵ^{R}	ϵ^{H}	ϵ^{B2}	g^{H}
radius=5	1.6%	7.2%	8.9%	5.0%	10.2%	2.4%
radius=10	1.9%	8.9%	10.9%	6.5%	10.3%	2.6%
radius=15	1.4%	10.5%	13.0%	7.1%	10.4%	4.0%

$p_1 = 95\%; c_1/c_2 = 10$

	ϵ^{B1}	ϵ^{L}	ϵ^{R}	ϵ^{H}	ϵ^{B2}	g^{H}
radius=5	1.1%	5.1%	5.7%	3.2%	6.5%	2.0%
radius=10	1.5%	6.0%	6.9%	4.3%	6.6%	1.8%
radius=15	1.5%	9.2%	9.9%	7.6%	6.9%	1.8%

$p_1 = 95\%; c_1/c_2 = 5$

	ϵ^{B1}	ϵ^{L}	ϵ^{R}	ϵ^{H}	ϵ^{B2}	g^{H}
radius=5	1.7%	5.9%	6.6%	4.0%	7.7%	2.0%
radius=10	1.6%	6.9%	7.9%	5.3%	7.3%	1.7%
radius=15	1.4%	9.4%	10.8%	7.1%	7.8%	2.5%

$p_1 = 95\%; c_1/c_2 = 0$

	ϵ^{B1}	ϵ^{L}	ϵ^{R}	ϵ^{H}	ϵ^{B2}	g^{H}
radius=5	1.9%	8.4%	9.8%	5.9%	11.2%	2.7%
radius=10	1.3%	10.1%	11.9%	6.9%	11.3%	3.6%
radius=15	1.8%	11.7%	13.9%	7.7%	11.7%	4.5%

Note: ϵ^{B1}, ϵ^{L}, ϵ^{R}, ϵ^{H} and ϵ^{B2} denote the cost saving for the solutions from the BBKK1, T-LSCPA, the relaxed problem when y_{ij} is a real number, the heuristic and the BRLSCP, compared with the QRLSCP, respectively. The sign "$\square$" represents where the solution of the BRLSCP results in the actual ex-post availability below the requirement. The two numbers with "*" are the optimality gap for the two instances where the optimal solutions cannot be generated by $CPLEX$ within acceptable computation time, and the heuristic will generate the solutions with the same minimum objective value with the solver $CPLEX$.

It is interesting to study the cost saving ϵ^{R} for a *relaxed model*, where the variable y_{ij} in the T-LSCPA is a real number in $[0, 1]$ instead of a binary one, that is, the demand node j is assigned to Station i according to a pre-determined probability of y_{ij}, which is one of the decision variables. It can

be observed from Table 5.2 that, the additional cost saving for the relaxed problem over the original T-LSCPA is significantly affected by the the coverage radius and the value of c_1/c_2. The cost saving from the relaxation increases with the coverage radius, and becomes larger for a smaller value of c_1/c_2.

Moreover, we notice that the model BRLSCP may result in a lower cost than the relaxed model in 9 instances when the coverage radius is smaller (i.e., smaller overlap fraction). However, as discussed in Section 3.3, FP^B may undershoot the traditional failure probability when the overlap is small. To further examine whether the required availability is achieved, we undertake a simulation by $AutoMod$12.3.1 to find the ex-post availability for all cases in Table 5.2, including the solutions of all five models. In the simulation, 1e+8 emergency calls are generated for each model. Each call is served by the closest available ambulance if there is at least an idle one in the coverage radius, otherwise the demand is lost. In other words, the solutions from all the models are tested under the same ambulance dispatching rule (i.e., the closest available first). We denote the fraction of demand which cannot achieve the service requirement as the infeasibility, which is calculated by

$$\%Inf = \frac{\sum_{j \in S(Inf)} \lambda_j}{\sum_{j \in J} \lambda_j}$$

where $S(Inf)$ is the set of the demand node which cannot be covered with the service requirement.

We find that the infeasibility for the model BBKK1 only varies from 0.21% to 0.57% due to the randomness of demand arrivals. The model

BRLSCP will lead to up to 20% infeasibility. Regarding the solutions of our model, we will also lead to 2.31% to 6.22% infeasibility for these 27 instances. The reason is that, the pre-assigned station for each demand by our solutions is not always the closest activated one. The closest available vehicle, however, will be assigned to serve the demand in the ex-post simulation. Then we modify our solutions by re-assigning the closest activated stations to those demand which are pre-assigned to a farther one in advance. The detailed procedure of this modification can be found in Appendix I. After this modification, the infeasibility is also reduced to 0.38% to 0.78%. Although it is still a little larger than that of the model BBKK1, it is acceptable as the randomness of demand arrivals also exists in practice. Furthermore, the infeasibility for the solutions solved by our heuristic varies from 0.27% to 0.69%. That is, using our heuristic, we can get the solutions, which can be well applied to practice, efficiently.

Furthermore, we examine the performance of our heuristic. As the cost saving of the T-LSCPA and the heuristic over the benchmark model, the QRLSCP, is shown in Table 5.2, the optimality gap of the proposed heuristic, g^H, can be calculated with the following formula as shown in the last column of Table 5.2.

$$g^H = \frac{TC_H^* - TC_L^*}{TC_L^*} = \frac{\epsilon^L - \epsilon^H}{1 - \epsilon^L}$$

Note that for the two instances where $CPLEX$ cannot provide optimal solutions within acceptable computation time, TC_L^* is substituted by the lower bounds provided by $CPLEX$ to calculate the gaps. As in these two instances, the heuristic will generate the same objective value as the

solver, the optimality gaps will also be equal to those from the solver, i.e., 2.4% and 3.9%, respectively. Hence, the optimality gaps of the proposed heuristic are found to vary from 1.2% to 4.5% as mentioned above, which verifies the results of Lemma 3. We also compare the performance of the proposed heuristic to that of other models, and find that the solutions from the heuristic are always better than the QRLSCP, even better than the BBKK1, with the cost saving over the BBKK1 lying in a range of $0.7\% - 5.8\%$. The heuristic proposed also shows its advantage in addressing the computation time issue. It needs very little time (less than 10s which is negligible compared with that of QRLSCP and BBKK1) to provide the solutions.

Finally, we conduct sensitivity analysis for our heuristic. The results for the sensitivity analysis are explained as follows. Table 5.3 shows the basis solutions of one example out of the 1,000 valid results of the scenario where $p_1 = 90\%, c_1/c_2 = 5$. Table 5.4 shows the solutions of this example when the target availability requirement increases by 2% and 4%, where "$\square$" denotes the solutions which are changed due to the parameter fluctuation. Table 5.5 shows the solutions of the same example when the value of $\lambda_j s_{ij}$ fluctuates by 10%, where β is a fluctuation factor associated with the value of $\lambda_j s_{ij}$. Here, $\beta = 1$ means $\lambda_j s_{ij}$ takes the basic value, and for other values of β, $\lambda_j s_{ij}$ are replaced by $\beta\lambda_j s_{ij}$ respectively. As can be seen, the heuristic is robust to the change of the values of p_1 and $\lambda_j s_{ij}$. This robustness is worth mentioning because of three reasons. Firstly, the availability standard might need to be improved to a higher service level under some special conditions, and the system designed by our heuristic is

robust in this condition. Secondly, the system designed by our heuristic is stable when the parameter λ_j fluctuates due to the population change. For example, the population is growing and aging in Singapore, and the system designed by our heuristic will perform well for the changing demand. Finally, the fluctuation of s_{ij} is much more common due to traffic patterns, road and weather conditions.

Table 5.3.: The basis solutions ($p_1 = 90\%$, β=1)

Node	4	7	14	17	19	21	24	25	38	40	43	50	52	54
Fleet	9	7	2	4	6	4	3	3	3	3	0	2	0	0

Table 5.4.: Solution sensitivity analysis to p_1

p_1	4	7	14	17	19	21	24	25	38	40	43	50	52	54
90%	9	7	2	4	6	4	3	3	3	3	0	2	0	0
92%	9	7	2	4	6	4	3	4	3	3	0	2	0	0
94%	9	7	2	4	6	4	3	4	3	3	0	2	0	0

Table 5.5.: Solution sensitivity analysis to $\lambda_j s_{ij}$

β	4	7	14	17	19	21	24	25	38	40	43	50	52	54
0.9	9	6	2	4	6	4	2	3	3	3	0	2	0	0
1	9	7	2	4	6	4	3	3	3	3	0	2	0	0
1.1	9	8	2	4	7	4	3	4	3	3	0	2	0	0

5.2. Case study 2: real case study for a known city S

We also apply our model in the real case problem for a known city S. For the demand set, we use the demand data for the first half year of 2006 (from Jan, 2006 to Jun, 2006), and each demand node refers to a unique postal code which has generated demand during the period and its location is available in Google Map. As the demand may occur almost continuously from any location, we merge all the nodes closest to each postal code to one and get in total 9,397 demand nodes. We consider 47 fire stations or fire posts based on the data of May, 2015 as our candidate station set. For each station or fire post, historical travel time and distance for each travel (from Google Map) are used to calculate an average speed, which is then used to estimate the coverage radius, i.e., the coverage radius equals the product of the average speed and the response time (9 min criterion is used in our study). So totally we have a 9,397 $\times$ 47 assignment matrix. The average service time for each station equals to the response time plus one hour time (including cleaning time, time travel to hospital).

In this example, we only compare the solutions of our model with B-BKK1, as it will result in the best solutions among all the other models as discussed in the previous example. We show the solutions of our model and the model BBKK1 under different service requirements in Table 5.6. As shown in this table, the benefits of demand assignment will become quite large in practice, saving 47/50/56 vehicles under the service requirements of 85%, 90% and 95%. The reason why the benefits are larger is that the demand in the overlapped area is much more in practice than that in the small case study. Even after the modification, we still save 27/33/36 vehi-

Table 5.6.. Comparison under the real case study

	No. of required vehicles (%Inf)		
p_1	85%	90%	95%
BBKK1	97(0.33%)	108(0.43%)	123(0.79%)
T-LSCPA	50(6.23%)	58(6.61%)	67(7.13%)
after modification	70(0.71%)	75(0.97%)	87(0.68%)
heuristic	71(1.01%)	74(1.64%)	86(1.77%)

Note: We also adopt the modification indicated in Appendix I

cles compared with the results of BBKK1, while the infeasibility is reduced a lot compared with the solutions from the original model. Our heuristic can provide similar solutions with the results after modification from the original T-LSCPA solutions. With respect to the running time, the optimal solutions can be obtained by $CPLEX12.5$ in 78.94s, 1232.42s for the first two cases, while the optimality gap for the last instance remain to be 4.01% even after 1e+10 iterations respectively. On the other hand, the heuristic can reduce the computation time to no more than 25s.

It is also interesting to test the infeasibility of the solutions to our model (before and after modification) and the heuristic with $AutoMod12.3.1$. 10 years real emergency call data are studied and the results are also shown in Table 5.6. The modification procedure to address the infeasibility issue proposed in Section 5.1, which can be referred in Appendix I, also performs well in this real case problem. The infeasibility can be reduced to be insignificant after such modification. Furthermore, the infeasibility of the solutions from our heuristic is also similar to the results after modification.

5.3. Summary and discussions

In this chapter, we apply our model to the frequently used 55-node problem and a real case problem for a known city S, and obtain a series of interesting observations, which verify what we have analysed in chapters 3 and 4.

As a conclusion of the experimental study, the cost saving of our model consistently outperforms the results of BBKK1. In addition, such saving also increases with the coverage radius and the availability level target, but decreases with the ratio of c_1/c_2. Furthermore, when the overlap among the coverage areas is significant, or when the demand is covered by more stations, our model with assignment requires fewer vehicles than the BBKK1 and the QRLSCP. Compared to the BRLSCP, our model performs better when the overlap is large.

Furthermore, we also provide a modification procedure to make the solutions of our model to be applicable to practice. In particular, the solutions of our model are modified to ensure the availability almost achieved under practical closest available vehicle policy. From the studies from both these two cases, we find that this procedure performs well in terms of achieving the service requirement by not increasing many vehicles.

Moreover, the solutions of our heuristic are insensitive to the change of the inputs, such as p_1 and $\lambda_j s_{ij}$. Such robustness of our model supports its applications when the availability standard changes or the population or traffic conditions are not stable.

To summarize, our model (the T-LSCPA) proposed for EMS systems with demand assignment incorporated can be handy for the practitioners in the location field, particularly the decision makers for the urban areas.

Furthermore, the analysis for the benefits of the demand assignment is also beneficial for the researchers in other fields, such as the location-inventory modelling.

PART II

Application of the LSC models in the GrSC network design

Chapter 6 PROBLEM STATEMENT FOR THE GrSC NETWORK DESIGN

This chapter states the problem for the application of the LSC models in the GrSC network design in more details. In particular, we discuss the main issues this work mainly addresses, especially the relationship of modules and options for some product and a description of how carbon tariffs are imposed through the supply chain under our discussion.

6.1. Main issues in the GrSC network design

The supply chain network under consideration is a three-tier one which is denoted as $E\{N, G\}$, where N is the node set and G is the set of links. N consists of the set of potential suppliers S providing various options of modules, potential assembling facilities F and customers C, i.e., $N = S \cup F \cup C$.

Our study mainly copes with the decision making in the global supply chain network design including the following three main issues.

1. Module selection.

   Traditionally, supplier's capacity, purchasing price and transportation costs are the main concerns for the assembling facilities when purchasing modules. This study additionally takes the carbon issue into consideration. As indicated in Section 2.3, the cradle-to-

gate CO2 emission and the emission during future usage would bring about carbon tariffs if exported to member society. Furthermore, the emissions during transportation are also considered by restricting that these emissions contribute to the total emissions which cannot exceed a stated cap.

2. Production in assembling facilities.

   In traditional supply chain design, managers consider set up costs, production costs and transportation costs when making decisions for assembling facilities. This study also takes into account the CO2 emissions and carbon policies, which refer to carbon cap and cap tariffs in this thesis. In particular, our study tries to examine the impacts of carbon tariffs on the production quantity decision and technology selection. For example, does a higher carbon tariff rate motivate the facilities to invest in higher level technology with lower emissions (usually higher investment costs) as our intuition suggests? Or does the higher tariff rate just bring about a shift of the production from non-member facilities to member facilities? These questions will be answered in Chapter 8.

3. Transportation mode selection.

   The transportation cost and carbon emission would differ a lot for different transportation modes (see, Cristea et al. 2013). Thus the effects of different transportation modes should be taken into account when designing GrSC network. In this study, several transportation modes can be selected for both modules and final products.

4. Carbon tariff imposition.

Since the carbon tariff is usually generated when goods are exported from non-member countries to member countries, the allocation of products among facilities and customers becomes an interesting topic. This part will be further discussed through the results of the experimental study in Chapter 8.

The subsequent sections in this chapter will further clarify the above-mentioned issues in order to construct our model in Chapter 7.

6.2. Relationship of modules and options

We first further explain the relationship of *products, modules and options* in more details. As mentioned in Section 2.3, facilities purchase modules to assemble products and suppliers provide options for these modules. Each option has its corresponding unit purchasing price, cradle-to-gate emission and expected emission in future usage. With carbon issue considered, facilities would not only ask for low price options, they will struggle to balance the purchase costs and the CO2 emission since more CO2 emission from lower cost options may lead to a higher carbon tariff when exported to foreign markets.

The purchased modules are then used for assembling to final products in the facilities, which should determine the amount of products assembled from different combinations of module options. Each product is assumed to be assembled by several modules and there are several options to be selected for each module. A simple example is provided to illustrate the relationship of products, modules and options. As shown in Table 6.1, we consider a certain product which is composed from five modules

i_1, i_2, i_3, i_4, i_5, and for each module, there are four possible options, e.g., $\{j_{11}, j_{12}, j_{13}, j_{14}\}$ are the possible options for part i_1. Eventually one option is selected for each module of the product. Hence, the amount of total module options used for each product equals the amount of modules for the product. That is, in this example, each product has five modules and thus five total module options, where one possible combination of the module options is $\{j_{11}, j_{21}, j_{31}, j_{41}, j_{51}\}$. We also assume that products which are assembled from different combinations of options would be regarded as equivalent alternatives for the consumers, i.e., in this example, a total of 4^5 combinations will be treated equally by the customers.

Table 6.1.. Relationship of modules and options of a certain product

Mod	Opt_1	Opt_2	Opt_3	Opt_4
i_1	j_{11}	j_{12}	j_{13}	j_{14}
i_2	j_{21}	j_{22}	j_{23}	j_{24}
i_3	j_{31}	j_{32}	j_{33}	j_{34}
i_4	j_{41}	j_{42}	j_{43}	j_{44}
i_5	j_{51}	j_{52}	j_{53}	j_{54}

6.3. Technology selection and transportation mode selection

In terms of the technology selection, each factory which is operating is assumed to be able to invest on only one technology level. Different technology levels would have different investment costs, yet different production costs and different CO2 emissions when assembling products. Generally, higher investment costs on technology lead to lower production costs as

well as lower carbon emissions in the assembling process. Furthermore, as the demand in the market is assumed to be deterministic, inventory in the assembling facilities is not involved within our single period decision framework.

On the other hand, regarding the transportation for modules and final products, as mentioned previously, Cristea et al. (2013) indicate that a large variety of modes, e.g., ships, plane, trucks, rail, pipelines and so on, are adopted in global trade. Different transportation modes may have quite varied transportation costs and CO2 emission per quantity shipped. Thus several transportation modes can be selected in this study for both modules and final products.

6.4. Description of how carbon tariffs are imposed

This section highlights our presentation on the carbon tariff imposition mechanism across the supply chain in our study, which makes it possible for us to construct our model in the next chapter. This is also one of the main contributions of this thesis. To our best knowledge, this is a novel and easy-to-understand manner to quantify the mechanism to impose carbon tariffs.

When the final products are ready, they will be sent to the customers to satisfy the demand. As discussed previously, member countries may impose a border carbon tariff on imports from non-member countries. The carbon tariff varies a lot depending on where the facilities and markets are located. How to impose carbon tariffs is a big challenge in practice (see, Persson 2010). Here we simplify the mechanism, which is still reasonable

as mentioned in Section 2.4, to derive our model in the next section. A simple example is illustrated to show how the carbon tariffs are imposed from non-member countries to member countries in the three-tier supply chain network under our discussion.

As shown in Figure 6.1, a certain amount of product, say P is required at customer nodes C_1 which is located in a member country and C_2 which is located in a non-member country. To satisfy the demand, two facilities F_1 and F_2, which are based in a member country and a non-member country respectively, assemble the product from the modules provided by suppliers $S_i, i = 1, 2, 3$ or 4. The CO2 emissions during the assembling process are indicated in the figure, e.g., 60 units for F_1 and 100 units for F_2. To simplify our explanation, we assume one product is assembled from a single module, say, A. The suppliers provide modules with different cradle-to-gate CO2 emissions as well as the predicted CO2 emissions in future usage. Such data has been shown in the figure. For example, the module provided by supplier S_1 contributes 70 units of CO2 emissions from raw material acquisition, and the predicted CO2 emissions during usage are 80 units.

Carbon tariffs are only imposed when goods flow from non-member countries to member countries. For example, the flow of module A from S_2 to F_1, the flow of product P from F_2 to C_1, and so forth. Moreover, tariffs on the emissions during usage will be imposed only when the product is exported to a member country customer if the module is from a non-member country. For example, carbon tariffs will be imposed for the emissions during usage for the flow $S_2 \rightarrow F_1 \rightarrow C_1$. However, there is no tariff on the emissions during usage for the flow $S_2 \rightarrow F_1 \rightarrow C_2$. Such an assumption is

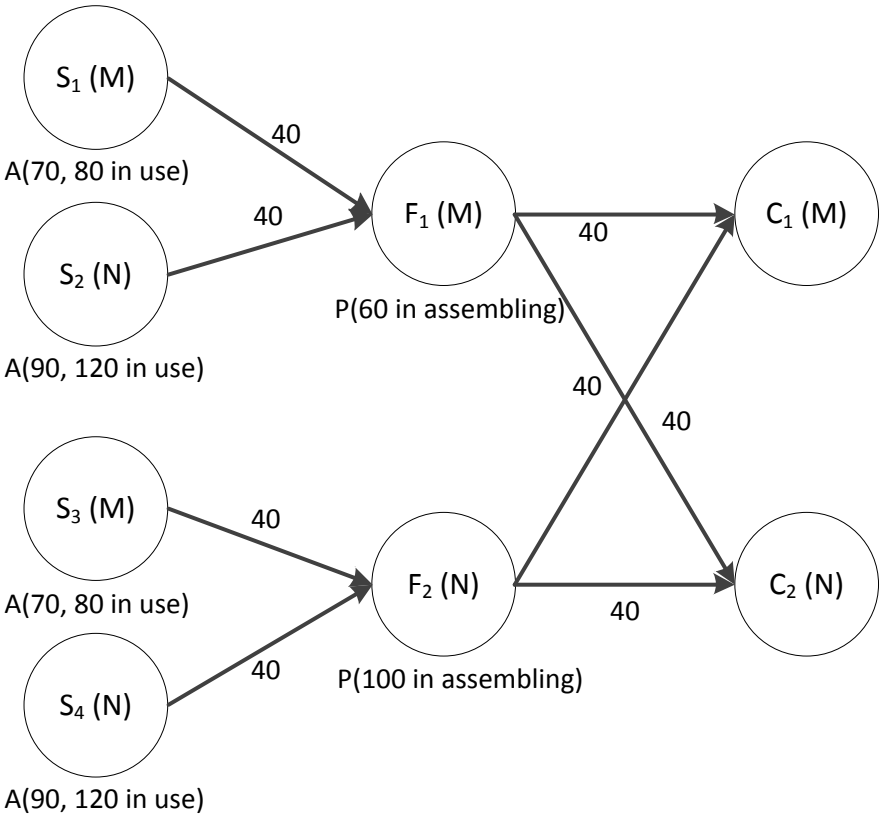

Figure 6.1.. The three-tier supply chain network

reasonable as the emissions during usage only impacts the customers.

Another assumption is that the CO2 emissions from the transportation from any supplier to any factory, or any factory to the customer are a constant 40. The CO2 emissions from transportation only increase the overall CO2 emissions, but will not bring about any carbon tariffs. As the transportation service can usually be provided by a third-party company, the carbon tariffs related to this part is beyond our current study.

According to the discussion above, we list the carbon tariffs that should be imposed for different routes as follows (the rates charged for each unit of carbon emission are assumed to be $\alpha_j, j = F_1, C_1$ in the member countries where F_1 and C_1 are located, respectively).

- Route (1): $S_1 \rightarrow F_1 \rightarrow C_1$

Because there are no goods flowing from non-member countries to member countries, the carbon tariffs charged are 0. The overall CO2 emissions are $70 + 80 + 40 + 60 + 40 = 290$.

- Route (2): $S_1 \rightarrow F_1 \rightarrow C_2$

  Similarly, as there are no goods flowing from non-member countries to member countries, the carbon tariffs charged are 0. The overall CO2 emissions are $70 + 80 + 40 + 60 + 40 = 290$.

- Route (3): $S_2 \rightarrow F_1 \rightarrow C_1$

  The carbon tariffs charged are $90\alpha_{F_1} + 120\alpha_{C_1}$. The overall CO2 emissions are $90 + 120 + 40 + 60 + 40 = 350$.

- Route (4): $S_2 \rightarrow F_1 \rightarrow C_2$

  The carbon tariffs charged are $90\alpha_{F_1}$. The difference from the previous flow results from the fact that customer C_2 is not a member country, and no tariffs will be imposed on the emissions during the usage. The overall CO2 emissions are $90 + 120 + 40 + 60 + 40 = 350$.

- Route (5): $S_3 \rightarrow F_2 \rightarrow C_1$

  The carbon tariffs charged are $100\alpha_{C_1}$. The overall CO2 emissions are $70 + 80 + 40 + 100 + 40 = 330$.

- Route (6): $S_3 \rightarrow F_2 \rightarrow C_2$

  Because there are no goods flowing from non-member countries to member countries, the carbon tariffs charged are 0. The overall CO2 emissions are $70 + 80 + 40 + 100 + 40 = 330$.

- Route (7): $S_4 \rightarrow F_2 \rightarrow C_1$

The carbon tariffs charged are $(90+120+100)\alpha_{C_1}$. The overall CO2 emissions are $90+120+40+100+40=390$.

- Route (8): $S_4 \rightarrow F_2 \rightarrow C_2$

  Because there are no goods flowing from non-member countries to member countries, the carbon tariffs charged are 0. The overall CO2 emissions are $90+120+40+100+40=390$.

In order to further explain how the carbon tariffs are imposed, especially for the emissions during usage, the following four-tier example (indicated in Figure 6.2) is used. In this example, S_3 is assumed to be located in a non-member country, provide the raw materials with 30 units of CO2 emissions, and the predicted emissions of such raw materials during final usage at node C will be 20. S_2 (in a member country) and S_1 (in a non-member country) both add some components with emissions of 25 and 20 units, and the predicted emissions during final usage for these added components are 30 and 40 units as well. The carbon tariffs imposed for this network would depend on whether or not the customer is based in a member country. If the customer C is in a member country, then the carbon tariffs imposed would be $30\alpha_{S_2}+(20+40)\alpha_C$, as all the emissions including the existent carbon and the carbon in future usage caused by a non-member company should be charged, where α_{S_2} and α_C are the rates in these two regions, respectively. On the other hand, if customer C is in a non-member country, the emissions in usage will not be charged. Then the total carbon tariffs charged should be $30\alpha_{S_2}$.

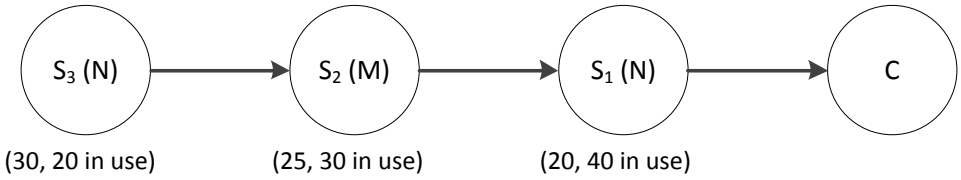

Figure 6.2.. The multi-tier supply chain network

6.5. Summary

This chapter states the problem for the GrSC network design, which serves as a preparation for the mathematical model construction in the next chapter.

Section 6.1 summarizes the four main issues for the LSC problem in designing the GrSC network: module selection; production decisions in assembling facilities; transportation mode selection and the carbon tariff mechanism. We assume that there are several options for each module of a product, with different purchasing prices, cradle-to-gate emissions and expected emissions in the future usage. In addition, the products assembled with different options will be treated equally by the customers. Managers can find a trade-off among those properties and decide which option should be selected for each module. Another critical issue is about how the carbon tariffs should be imposed. Actually, this is a difficult problem for operators so far. Section 6.4 clearly defines this mechanism based on some simplifying assumptions, which makes it possible to incorporate the carbon tariff mechanism into the mathematical model and to investigate the impacts of carbon tariff via applying the model in a real case study.

Chapter 7 MODEL CONSTRUCTION FOR THE GrSC NETWORK DESIGN

This chapter constructs a mathematical model in which CO2 emissions during usage and carbon tariff are incorporated into the GrSC network design. Firstly, we list the required notations which are used to formulate the model. Then the objective function as well as the constraints are presented.

7.1. List of notations

This section lists the necessary notations below to construct the models in the next section.

- General parameters:

i = the index of the module;

j = the index of the module option;

M_{ij} = the jth possible option for Module i of the product. $i \in I$, where I is the module set, $j \in J_i$, where J_i is the set of options for Module i. Module option M_{ij} is denoted as Module M_{ij} to simplify the subsequent presentation;

Cap = the CO2 emission cap for the global supply chain;

α_n = the carbon tariff rate in Node $n \in N$. e.g., $\alpha_n = 0$ if Node n is located in a non-member country;

$$Ind_n = \begin{cases} 1, & \text{if Node } n \text{ is located in a member country, where } n \in N; \\ 0, & \text{otherwise.} \end{cases}$$

- Parameters related to suppliers:

$C_s^1(M_{ij})$ = the purchasing price of Module M_{ij} from Supplier $s \in S$;

$E_s^1(M_{ij})$ = the cradle-to-gate CO2 emissions of Module M_{ij} provided by Supplier s;

$E_s^2(M_{ij})$ = the expected CO2 emissions during future usage of Module M_{ij} provided by Supplier s;

$u_s^S(M_{ij})$ = the maximum amount of Module M_{ij} provided by Supplier s;

- Parameters related to assembling facilities:

C_f^{Set} = set up cost for Factory $f \in F$;

L_f = the set of candidate technologies that can be invested by Factory f;

C_{fl}^{Inv} = the investment cost if Factory f invests on Technology $l \in L_f$;

C_{fl}^{Pro} = the production cost of assembling one unit of the product in Factory f if it invests on Technology $l \in L_f$;

E_{fl}^{Pro} = the average CO2 emissions of assembling one unit of the product in Factory f if it invests on Technology $l \in L_f$;

u_f^F = the maximum assembling capacity for the product in Factory f;

- Parameters related to customers:

d_c = the demand for the product at customer node $c \in C$ (note that the products which are made from different module options are regarded as equivalent alternatives from the perspective of the customers);

- Parameters related to transportation:

T = the set of possible transportation modes;

$C_t^{Tp}(M_{ij})$= the cost of transporting one unit of Module M_{ij} for 1,000 km if Mode $t \in T$ is selected;

$\overline{C_t^{Tp}}$ = the cost of transporting one unit of final product for 1,000 km if Mode $t \in T$ is selected;

$E_t^{Tp}(M_{ij})$= the average CO2 emissions of transporting one unit of Module M_{ij} for 1,000 km if Mode $t \in T$ is selected;

$\overline{E_t^{Tp}}$ = the average CO2 emissions of transporting one unit of final product for 1,000 km if Mode $t \in T$ is selected;

$dis_{tn_1n_2}$ = the distance for transportation Mode t between Node n_1 and Node n_2, where $n_1, n_2 \in N$. If Mode t between Node n_1 and Node n_2 is unavailable, we set $dis_{tn_1n_2}$ as a very big number.

- Decision variables:

$$y_f = \begin{cases} 1, & \text{if Factory } f \text{ is open;} \\ 0, & \text{otherwise.} \end{cases}$$

$$z_{fl} = \begin{cases} 1, & \text{if Factory } f \text{ invests on Technology } l \in L_f; \\ 0, & \text{otherwise.} \end{cases}$$

$x_{sfc}(M_{ij})$= the amount of Module M_{ij} purchased from Supplier s that finally will arrive in Customer c as final products after assembling in Factory f;

x_{flc} = the amount of final products which are assembled in Factory f under Technology l and finally arrive in Customer c;

$$g_{sf}^t(M_{ij}) = \begin{cases} 1, & \text{if transportation Mode } t \in T \text{ is selected for transporting Module } M_{ij} \text{ from Supplier } s \text{ to Factory } f; \\ 0, & \text{otherwise.} \end{cases}$$

$$\overline{g^t_{flc}} = \begin{cases} 1, & \text{if transportation Mode } t \in T \text{ is selected for transporting the final products from Factory } f \text{ under technology } l \text{ to Customer } c; \\ 0, & \text{otherwise.} \end{cases}$$

7.2. Mathematical model

With the notations in the previous section, we present our Mathematical Model for the Green Supply Chain problem (MMGSC) as follows. The objective of the MMGSC is to minimize the total costs which consist of the traditional logistics costs and the total carbon tariffs imposed. The components of the costs are listed below.

- Set up costs for assembling facilities (denoted as $LSUC$):

$$LSUC = \sum_{f \in F} C^{Set}_f y_f$$

- Investment costs on technologies for facilities (denoted as LIC):

$$LIC = \sum_{f \in F} \sum_{l \in L_f} C^{Inv}_{fl} z_{fl}$$

- Total purchasing costs for modules (denoted as $LPURC$):

$$LPURC = \sum_{s \in S} \sum_{i \in I} \sum_{j \in J_i} C^1_s(M_{ij}) \sum_{c \in C} \sum_{f \in F} x_{sfc}(M_{ij})$$

- Total production costs for assembling products (denoted as $LPROC$):

$$LPROC = \sum_{f \in F} \sum_{l \in L_f} C_{fl}^{Pro} \sum_{c \in C} x_{flc}$$

- Total transportation costs (denoted as $LTPC$):

$$\begin{aligned} LTPC &= \sum_{t \in T} \sum_{i \in I} \sum_{j \in J_i} C_t^{Tp}(M_{ij}) \sum_{c \in C} \sum_{f \in F} \sum_{s \in S} x_{sfc}(M_{ij}) g_{sf}^t(M_{ij}) dis_{tsf} \\ &+ \sum_{t \in T} \overline{C_t^{Tp}} \sum_{c \in C} \sum_{f \in F} \sum_{l \in L_f} x_{flc} \overline{g_{flc}^t} dis_{tfc}, \end{aligned} \tag{7.1}$$

  where the first term is the costs for transporting modules from suppliers to facilities, and the second term is the costs for sending final products to the customers.

- Carbon tariffs (denoted as TAR) as the monetary form imposed on the global supply chain:

$$\begin{aligned} TAR &= \sum_{s \in S} \sum_{i \in I} \sum_{j \in J_i} E_s^1(M_{ij}) \sum_{f \in F} \sum_{c \in C} x_{sfc}(M_{ij}) \alpha_f Ind_f (1 - Ind_s) \\ &+ \sum_{f \in F} \sum_{l \in L_f} E_{fl}^{Pro} \sum_{c \in C} x_{flc} \alpha_c Ind_c (1 - Ind_f) \\ &+ \sum_{s \in S} \sum_{i \in I} \sum_{j \in J_i} E_s^2(M_{ij}) \sum_{f \in F} \sum_{c \in C} x_{sfc}(M_{ij}) \alpha_c Ind_c (1 - Ind_s) \\ &+ \sum_{s \in S} \sum_{i \in I} \sum_{j \in J_i} E_s^1(M_{ij}) \sum_{f \in F} \sum_{c \in C} x_{sfc}(M_{ij}) \alpha_c Ind_c (1 - Ind_s) \\ &\quad \times (1 - Ind_f) \end{aligned}$$

As mentioned above, the objective of the MMGSC is to minimize the total costs which consist of the above-mentioned components, i.e, the ob-

jective of Model MMGSC is

$$\min c = LSUC + LIC + LPURC + LPROC + LTPC + TAR$$

Next, we discuss the constraints of Model MMGSC. The first constraint is that the overall CO2 emissions in the global supply chain network cannot exceed the predetermined cap, i.e., Cap. To derive this constraint, we have to find the overall CO2 emissions across the global supply chain network first. The overall CO2 emissions include the following three parts:

1. CO2 emissions from modules (denoted as $EMOD$):

$$\begin{aligned} EMOD &= \sum_{s\in S}\sum_{i\in I}\sum_{j\in J_i} E_s^1(M_{ij}) \sum_{f\in F}\sum_{c\in C} x_{sfc}(M_{ij}) \\ &+ \sum_{s\in S}\sum_{i\in I}\sum_{j\in J_i} E_s^2(M_{ij}) \sum_{f\in F}\sum_{c\in C} x_{sfc}(M_{ij}), \end{aligned}$$

   where the first term refers to the CO2 emissions of the modules from raw material acquisition, i.e., cradle-to-gate emissions, and the second term refers to the CO2 emissions from the future usage of the modules.

2. CO2 emissions during production process (denoted as $EPRO$):

$$EPRO = \sum_{f\in F}\sum_{l\in L_f} E_{fl}^{Pro} \sum_{c\in C} x_{flc}$$

3. CO2 emissions from transportation (denoted as $ETRP$):

$$\begin{aligned} ETRP &= \sum_{t\in T}\sum_{i\in I}\sum_{j\in J_i} E_t^{Tp}(M_{ij}) \sum_{c\in C}\sum_{f\in F}\sum_{s\in S} x_{sfc}(M_{ij}) g_{sf}^t(M_{ij}) dis_{tsf} \\ &+ \sum_{t\in T} \overline{E_t^{Tp}} \sum_{c\in C}\sum_{f\in F}\sum_{l\in L_f} x_{flc} \overline{g_{flc}^t} dis_{tfc}, \end{aligned} \tag{7.2}$$

where the first term is the CO2 emissions for transporting modules from suppliers to facilities, and the second term is the CO2 emissions for transporting final products from facilities to end customers.

Then we obtain our first constraint below.

- CO2 emission cap restriction:

$$EMOD + EPRO + ETRP \leq Cap \tag{7.3}$$

As discussed above, the left hand side denotes the overall emissions of the global supply chain network, which should not exceed the cap predetermined.

The other constraints of Model MMGSC are as follows.

- supplier's capacity restriction:

$$\sum_{c\in C}\sum_{f\in F} x_{sfc}(M_{ij}) \leq u_s^S(M_{ij}), \qquad \forall s \in S,\ \forall i \in I,\ \forall j \in J_i$$

- assembling factory's investment on technology:

$$\sum_{l\in L_f} z_{fl} = y_f, \qquad \forall f \in F$$

This constraint means that, if a factory operates, it can invest on a certain level of technology; otherwise, no technology is selected.

- assembling factory's capacity:

$$\sum_{c \in C} x_{flc} \leq u_f^F z_{fl}, \qquad \forall f \in F,\ \forall l \in L_f$$

- balance between the modules and products assembled:

$$\sum_{s \in S} \sum_{j \in J_i} x_{sfc}(M_{ij}) = \sum_{l \in Lf} x_{flc}, \qquad \forall i \in I,\ \forall c \in C,\ \forall f \in F$$

As the product is assembled from several modules, this constraint means that the amount of each module from all kinds of options should be equal, and this number should equal the amount of final product as well.

- end demand satisfaction:

$$\sum_{f \in F} \sum_{l \in Lf} x_{flc} = d_c,\ \forall c \in C \tag{7.4}$$

Note that in this constraint, usually we can relax the "=" by "$\geq$" and get an unchanged optimal solution. However, we will show some counterexamples via the numerical studies in Chapter 8 where the relaxed constraint may bring changes to the original optimal solutions.

- one transportation mode should be selected for each module or each

final product:

$$\sum_{t\in T} g^t_{sf}(M_{ij}) = 1, \qquad \forall s \in S,\ \forall f \in F,\ \forall i \in I,\ \forall j \in J_i$$

$$\sum_{t\in T} g^t_{flc} = 1, \qquad \forall f \in F,\ \forall l \in L_f,\ \forall c \in C$$

- binary variables:

$$y_f \in \{0,1\}, \qquad \forall f \in F$$

$$z_{fl} \in \{0,1\}, \qquad \forall f \in F,\ \forall l \in L_f$$

$$g^t_{sf}(M_{ij}) \in \{0,1\}, \qquad \forall s \in S, \forall f \in F, \forall i \in I, \forall j \in J_i, \forall t \in T$$

$$\overline{g^t_{flc}} \in \{0,1\}, \qquad \forall f \in F,\ \forall l \in L_f,\ \forall c \in C, \forall t \in T$$

- nonnegative variables:

$$x_{sfc}(M_{ij}) \geq 0, \qquad \forall s \in S,\ \forall f \in F,\ \forall c \in C,\ \forall i \in I,\ \forall j \in J_i$$

$$x_{flc} \geq 0, \qquad \forall f \in F,\ \forall l \in L_f,\ \forall c \in C$$

7.3. Model analysis

In Model MMGSC, there are some non-linear terms in the objective function and a constraint. In particular, they are:

1. The term $LTPC$ of the objective function which is specified in Equation (7.1);

2. The term $ETRP$ of Constraint (7.3) which is specified in Equation (7.2).

That is, MMGSC is a mixed integer non-linear program. In this section we introduce an approach to transform our model with those non-linear terms to a pure linear one without changing the solution space. Moreover, we examine how the value of carbon tariff rate affects decision making on technology selection, total costs and emissions. The insights derived in this section will be verified by a real case study in the next chapter.

7.3.1. Transformation to a pure linear model

In order to re-write the first term of Equation (7.1), we redefine another variable $x^t_{sfc}(M_{ij})$ to denote the modules transported from Supplier s to Factory f by Mode t; and will finally be sent to Customer c as final products.

$$\sum_{t\in T}\sum_{i\in I}\sum_{j\in J_i} C^{Tp}_t(M_{ij}) \sum_{c\in C}\sum_{f\in F}\sum_{s\in S} x_{sfc}(M_{ij}) g^t_{sf}(M_{ij}) dis_{tsf}$$

can then be re-written as a linear form:

$$\sum_{t\in T}\sum_{i\in I}\sum_{j\in J_i} C^{Tp}_t(M_{ij}) \sum_{c\in C}\sum_{f\in F}\sum_{s\in S} x^t_{sfc}(M_{ij}) dis_{tsf}$$

by adding the following two constraints:

$$\begin{aligned} \sum_{t\in T} x^t_{sfc}(M_{ij}) &= x_{sfc}(M_{ij}), \forall s, f, c, i, j \\ x^t_{sfc}(M_{ij}) &\leq Q \times g^t_{sf}(M_{ij}), \forall t, s, f, c, i, j \end{aligned}$$

where Q is a big enough number; $\max\{u^F_f, \forall f \in F\}$ is a viable choice for the value of Q.

With the above-mentioned method, we can re-write the linear form

of equations (7.1) and (7.2) by introducing two more decision variables: $x^t_{sfc}(M^p_{ij})$ and x^t_{flc},

Equation (7.1) is re-written as:

$$\begin{aligned} LTPC &= \sum_{t\in T}\sum_{i\in I}\sum_{j\in J_i} C^{Tp}_t(M_{ij}) \sum_{c\in C}\sum_{f\in F}\sum_{s\in S} x^t_{sfc}(M_{ij}) dis_{tsf} \\ &+ \sum_{t\in T} \overline{C^{Tp}_t} \sum_{c\in C}\sum_{f\in F}\sum_{l\in L_f} x^t_{flc} dis_{tfc} \end{aligned}$$

Equation 7.2 is re-written as:

$$\begin{aligned} ETRP &= \sum_{t\in T}\sum_{i\in I}\sum_{j\in J_i} E^{Tp}_t(M_{ij}) \sum_{c\in C}\sum_{f\in F}\sum_{s\in S} x^t_{sfc}(M_{ij}) dis_{tsf} \\ &+ \sum_{t\in T} \overline{E^{Tp}_t} \sum_{c\in C}\sum_{f\in F}\sum_{l\in L_f} x^t_{flc} dis_{tfc} \end{aligned}$$

For the above transformation, we have to add four more constraints, as follows.

$$\begin{aligned} \sum_{t\in T} x^t_{sfc}(M_{ij}) &= x_{sfc}(M_{ij}), \forall s, f, c, i, j \\ \sum_{t\in T} x^t_{flc} &= x_{flc}, \forall f, l \in L_f, c \\ x^t_{sfc}(M_{ij}) &\leq Q \times g^t_{sf}(M_{ij}), \forall t, s, f, c, i, j \\ x^t_{flc} &\leq Q \times \overline{g^t_{flc}}, \forall t, f, l \in L_f, c \end{aligned}$$

Through the above transformation, the MMGSC is modelled as a mixed integer linear program thus far. As mentioned previously, the carbon emissions during future usage as well as the carbon tariff are both incorporated into the model. As a mixed integer linear programming model, it is demonstrated that this transformed model needs quite little computation-

al burden when applied in a real case study in the next chapter, i.e., the computational time for each experiment instance can be solved within less than two seconds of computer time.

7.3.2. The impacts of carbon tariffs on technology selection

We are interested in the impacts of carbon tariffs on the technology selection. Assume a customer C^M from a member country has a demand d for some product, which can be satisfied by either a member factory F^M or a non-member factory F^N. Both factories have enough capacity to satisfy the demand of C^M. The non-member factory F^N is currently using a low-level technology, with unit production cost of c_l and unit production emission of e_l. It is to be decided whether F^N should invest on higher-level technologies. Without loss of generality, we assume that if the factory wants to invests on a higher level technology, the investment costs consist of a constant part (denoted as Inv) and a variable part (denoted as t). The corresponding unit cost saving $f(t)$ and unit emission saving $g(t)$ are only dependent on the variable part. (Note: In reality, there may be several technology levels which should be defined as discrete. Here we adopt a continuous function for analysis. When we derive the optimal t to be invested, we just select the near-optimal technology level as our real decision.) We also assume that both $f(t)$ and $g(t)$ follow typical shapes of marginal benefit curves, i.e., both functions are increasing with t, while the marginal increments are declining, that is, both are concave functions of t. Thus, the unit production cost when F^N invests t on technology is $c_l - f(t)$, and the unit production emission is $e_l - g(t)$.

Generally, Customer C^M will import from F^N without carbon tariff imposition, due to the lower production costs. When the carbon tariffs are imposed, does F^N have motivation to update to a higher level technology, and to which specific level if so? We compare the total production costs and emission costs resulting from carbon tariff imposition in order to answer this question.

The costs at low-level technology are: $C(0) = d(c_l + \alpha e_l)$, where α is the carbon tariff rate. If Factory F^N invests t on technology, the costs become $C(t) = Inv + t + d[(c_l - f(t)) + \alpha(e_l - g(t))]$. We define the *Net Benefit of Technology* t ($NB(t)$) as:

$$NB(t) = C(0) - C(t) = -Inv - t + d[f(t) + \alpha g(t)] \tag{7.5}$$

As $f(t)$ and $g(t)$ are both concave functions of t, $NB(t)$ is also a concave function of t. We denote the optimal point as t_0 which satisfies

$$\frac{\partial NB(t)}{\partial t}|t_0 = -1 + d\frac{\partial f(t_0)}{\partial t_0} + d\alpha\frac{\partial g(t_0)}{\partial t_0} = 0$$

The marginal increment will decline with t within $[0, t_0]$. Under different tariff rates ($\alpha_1 < \alpha_2 < \alpha_3$), the graphs of $NB(t)$ are as shown in Figure 7.1.

Hence, the optimal decisions on technology selection for Factory F^N under different carbon tariff rates are:

$\alpha < \alpha_2 (e.g.,\ \alpha = \alpha_1)$: keep the low-level technology, as $\max NB(t)|_{\alpha_1} = NB(t_1)|_{\alpha_1} < 0$;

$\alpha = \alpha_2$: is a critical value to invest on higher technology, as $\max NB(t)|_{\alpha_2} =$

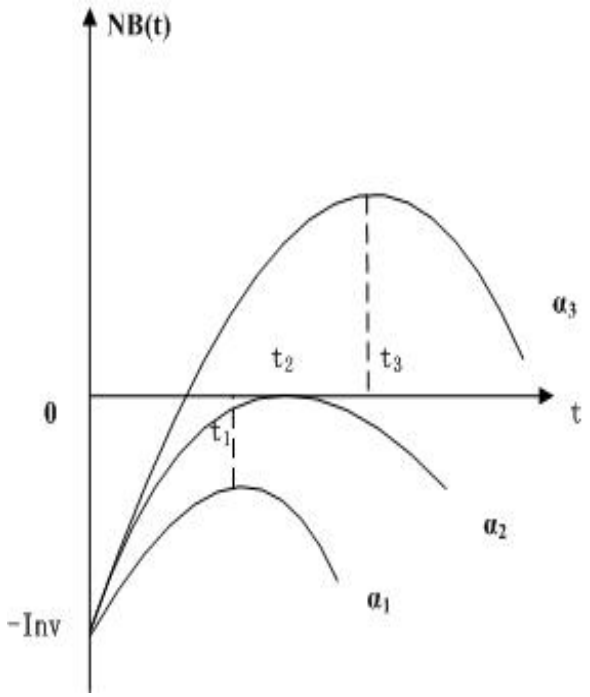

Figure 7.1.. The graphs of $NB(t)$

$NB(t_2)|_{\alpha_2} = 0$;

$\alpha > \alpha_2 (e.g.,\ \alpha = \alpha_3)$: invest t_3 on technology, as $\max NB(t)|_{\alpha_3} = NB(t_3)|_{\alpha_3} > 0$.

The general case for the above optimal decision selection is quite difficult to prove as the specific forms of $f(t)$ and $g(t)$ are unknown. Then the relationship between the $\max NB(t)$ and α is difficult to find. However, We can provide the formal proof if $f(t)$ and $g(t)$ are assumed as some specific forms as follows. Because they are both increasing and concave functions of t as discussed previously in this section, we assume $f(t) = g(t) = \ln t$. Generally $t > 1$, $f(t)$ and $g(t)$ are then non-negative. $\max NB(t)$ can be obtained when $t = t_0$ which makes the following equation hold.

$$\frac{\partial NB(t_0)}{\partial t_0} = -1 + d\frac{\partial f(t_0)}{\partial t_0} + d\alpha\frac{\partial g(t_0)}{\partial t_0} = -1 + d\frac{1}{t_0} + d\alpha\frac{1}{t_0} = 0.$$

That is, $t_0 = d(\alpha + 1)$. Then,

$$\max NB(t) = NB(t_0) = -Inv - t + d(\alpha + 1) \ln [d(\alpha + 1)].$$

The derivative of $\max NB(t)$ is

$$\frac{\partial \max NB(t)}{\partial \alpha} = d(\alpha + 1).$$

Generally, $d(\alpha + 1) > 0$ indicates that the derivative of $\max NB(t)$ is an increasing function of α. If $\max NB(t) < 0$, there exists a critical point α_2 which makes $\max NB(t) < 0|_{\alpha=\alpha_2} = 0$. Then,

when $\alpha < \alpha_2$, $\max NB(t) < 0$, i.e., the factory should keep the low-level technology;

when $\alpha > \alpha_2$, $\max NB(t) > 0$, i.e., the factory should invest on higher technology.

We can see that factories tend to use higher level technologies with a higher tariff rate, and the net benefits of using higher technologies are also higher.

On the other hand, assume that if Customer C^M imports products from the member-factory F^M, the corresponding costs (production costs and emission costs) are $C(Mr) = dc_{Mr}$, where c_{Mr} is the unit production cost for Factory F^M. As mentioned, $c_{Mr} > c_l$. When the carbon tariff rate α is too high, in particular, if

$$C(0) - \max NB(t)|_\alpha > C(Mr)$$

Customer C^M will directly import products from Factory F^M to save the

tariff costs, even though the production costs are higher. The insights from this section will be further verified by the numerical study in Section 8.1.1.

7.3.3. The impacts of carbon tariffs on total emissions

Following the discussion above, however, if the capacity of F^M, denoted as u_{F^M}, is not enough for the demand of Customer C^M, i.e., $u_{F^M} < d$, then C^M must import at least $d - u_{F^M}$ amount of product from F^N. By Formula (7.5), $NB(t)$ in Factory F^N is increasing with $d - u_{F^M}$. Hence, when $d - u_{F^M}$ is too low, F^N will have no motivation to transfer to higher level technology. At this stage, the total emissions will increase with $d - u_{F^M}$ as the production in F^N brings about heavier emissions than that of F^M.

When $d - u_{F^M}$ continues increasing and the maximum net benefits are positive, i.e., $\max NB(t)|_{d-u_{F^M}} > 0$, Factory F^N will take actions (higher level technology) to reduce the production costs as well as the tariff cost resulting from carbon emissions. Then the total emissions will go down.

When $d - u_{F^M}$ is too large and Factory F^N reaches its highest possible technology, the increase of product flow from F^N to C^M may lead to higher emissions, if the transportation load from F^N to C^M is heavier than that of F^M. The insights from this section will be further verified by the numerical study in Section 8.1.2.

7.4. Summary

In this chapter, we derive our mathematical model MMGSC based on the analysis in Chapter 6. The model MMGSC is constructed in such a way that it takes into account carbon emissions during usage and carbon tariff

when designing the GrSC network. The original model is not a linear model, i.e., there are some non-linear terms in the objective function and a constraint, and we propose a way to transform it to be linear. Moreover, we examine how the value of carbon tariff rate affects decision making on technology selection, total costs and emissions. The insights derived in this section will be verified by a real case study in the next chapter.

Chapter 8 EXPERIMENTAL RESULTS AND DISCUSSIONS ON THE MMGSC

In Chapter 7, a mixed integer programming model (MMGSC) is proposed that considers emissions during usage as well as carbon tariffs imposed for imported goods in a supply chain network. This chapter applies it in a real case study, in order to explore the managerial insights for the company in our real case study from the proposed model. As Company G is a leading international company in computer industry and the impacts in our study are significant, we believe this study is representative, and the insights from this study can be generalized to similar companies without much modification. In fact, the obtained results have also been verified with the Manufacturing Business Unit's General Manager of Company G.

8.1. Numerical experimental study and results

The model is built up in optimization software $AIMMS3.14$, and each problem is solved by CPLEX 12.5. All experiments are implemented on the 3.40GHZ/8GB computer. Note that the computation time of each run is less than 2s, and so we believe our model can be similarly applied in larger scale problems.

The real case is motivated by the global supply chain network design of Company G, a major computer company in Taiwan. The Company

provides a comprehensive product line and its products have been sold all over the world. For example, the current annual worldwide demand for one type of notebook, say Type I notebook, from Company G is around 100,000 units. The sales distribution (quantities and %) of Type I notebook around the world is as shown in Table 8.1.

Table 8.1.. Sales distribution of Type I notebook

Euro	AUS	China	Korea	Russia	Taiwan	UK	USA
12,376	2,723	6,064	32,104	13,144	14,505	3,837	15,248
12.38%	2.72%	6.06%	32.10%	13.14%	14.50%	3.84%	15.25%

Company G is now a pioneer in carbon reduction programs and thus is interested in its network design under carbon reduction policies. In this case, we are required to provide some managerial guidance on the green supply chain network design if carbon tariffs are involved.

We then build our optimization model to investigate the notebook supply chain network for this case. For simplicity, each Type I notebook is assumed to be assembled by five modules (motherboard, display, battery, chassis and other parts), and we provide two options for each module (lower emission yet higher cost option and higher emission yet lower cost option). The network under our analysis consists of six potential facilities and five potential suppliers, as shown in Table 8.2. This study is motivated by Company G's plan of outsourcing some production to local facilities as long as the costs can be reduced. Following Böhringer et al. (2013), we label China, Taiwan and Russia as non-member countries, and other countries as

member countries. Moreover, each supplier node or customer node is also assumed to represent the aggregative supply or demand in a region.

Table 8.2.. The network under analysis

	Euro	AUS	China	Korea	Russia	Taiwan	UK	USA
Member?	✓	✓		✓			✓	✓
Factory?	✓		✓	✓		✓	✓	✓
u_f^F	10k		25k	10k		25k	20k	30k
Supplier?	✓		✓			✓	✓	✓

The subsequent sections aim to answer the questions of concern to the managers in Company G, as mentioned in Section 1.2.

8.1.1. The impacts of carbon tariff on production quantity decision and technology selection

This section tries to explore the impacts of carbon tariffs on the technology selection for factories of Company G. (Note: The technology selection means whether we update the current technology to an upper status. This update needs much lower cost than if we change the equipment. The latter one may not be feasible for the company when the budget is not enough. Also, our time horizon is considered to be one calendar year in this study.) We start with the network where the capacities of the potential facilities are shown as Table 8.2, and each supplier's capacity is assumed to be 1.2 $\times$ the corresponding factory's capacity. We then study the impact on the network design for varying carbon tariff rates α, that is, the impacts on

the network design are studied by varying carbon tariff rate α. Note that in this study the tariff rates are assumed to be identical for all countries or regions in the network. The results of the product flow indicate that when no carbon tariff is imposed, i.e., $\alpha = 0$, Factories in China and Taiwan would reach full scale production to meet the demand which cannot be domestically satisfied in Russia, Euro, Australia and Korea, due to the lower production costs in these two non-member regions. Table 8.3 also shows that when $\alpha = 0$, China would select the middle level technology and Taiwan would select low level (high emission) technology. We also find that the facilities only use the low cost (high emission) module options since no tariff will be imposed for the emissions. The total costs are around $ 40,764K and the total emissions are around 2,103,542K kgs.

Table 8.3.. Technology level selection under different carbon tariff rates

technology	α=0	α=0.3	α=0.6
Euro	M	M	M
China	M	M	L
Korea	M	M	M
Taiwan	L	L	L
UK	M	M	H
USA	H	H	H

Note: L: Low level (high emission) technology; M: Middle level (medium emission) technology; H: High level (low emission) technology

If the carbon tariff rate increases, say, α = \$0.3 per kg and \$0.6 per kg, the production amount in China and Taiwan decreases. Besides the domestic demand, China only produces enough to satisfy the demand in Russia, as Russia is a non-member country and no tariff will be imposed; Taiwan only produces enough to cover Australia's demand due to the near

distance. Euro and Korea, however, instead tend to transfer their imports from UK and USA to avoid the tariff. Especially, when $\alpha = \$0.6$, the production amount in China is only 19,208, which is exactly equal to the total demand in China and Russia. Furthermore, Taiwan only produces 17,228 units for the domestic demand and Australia's demand. That is, with the increase of the carbon tariff rate, non-member facilities tend to produce less and export products to non-member customers to avoid the tariff expense. On the other hand, the demand in member customers is preferred to be met by member facilities. In terms of the module selection, the factory in Korea would import a larger quantity of low emission (high cost) modules from China to reduce the carbon tariff when α increases. The important interpretation is that the imposition of carbon tariffs seems to be like setting up a barrier between member society and non-member society, as opponents of the carbon tariff argue in Section 2.4. However, carbon tariff imposition indeed improves the competitiveness of products from member factories in member customers as the proponents propagandise.

According to Table 8.3, we also observe that the factory in China will take on low level technology when α increases to \$0.6. Taiwan will always take the low level technology when α reaches \$0.3 or \$0.6 as well. Moreover, the corresponding total costs and emissions when α is \$0.3 (\$0.6) are $42,152K$ $(43,796K)$ and $2,102,977K$ $(2,102,463K)$ kg. It is therefore worth noting that, as long as the demand in member customers can already be satisfied by member factories, the imposition of carbon tariffs alone cannot force the non-member factories to take low emission; besides, the mechanism alone has little effect in curbing the total emissions.

Since only imposing carbon tariffs sometimes has little effect as mentioned, it would therefore be interesting to study how to motivate the non-member facilities to invest in low emission technologies. One possible method is to lower the total emission cap restriction. In the example above, China will invest on the low emission technology if the cap is reduced to 76.2% of the primary amount; the corresponding number for the Taiwan factory is 68.5%. However, this is the impact of the carbon cap instead of the carbon tariff, which is also discussed by Wang et al. (2011), and thus it is not repeated in this study.

The reason why carbon tariffs have little impact in forcing non-member facilities to take low emission technologies is because the member facilities can already satisfy the demand in member customers. We are interested in examining the condition when the maximum production capacity of the member facilities is restricted, so that member customers must import some products from the non-member facilities. This is a reasonable assumption as member countries usually curb the scale of facilities, and meet the domestic demand by importing from foreign facilities. Furthermore, many non-member facilities usually produce more than their own demand to earn profits.

We assume a scenario where there is no demand in non-member customers, that is, the demand in non-member customers (China, Russia and Taiwan) is set to be 0. Since the non-member facilities can produce excessive goods besides the amount needed to satisfy the demand in a non-member society; thus, the demand in the non-member customers is not considered any more. The capacities of some facilities are adjusted as:

$u^F_{Korea} = 0$, $u^F_{UK} = 3,837$ and $u^F_{USA} = 25,000$, and the rest of the facilities maintain their original capacities, in order to ensure that member customers must import products from non-member facilities. The reason to create this case is as follows. In the scenario created here, Korea factory's capacity is assumed to be 0, then the total demand, i.e., 32,104 as shown in Table 8.1, must be satisfied by foreign factories; we let UK factory's capacity equals to its demand 3,837, then the market in UK can be taken out of our consideration in this case, making the results easier to understand. Regarding the US factory in the original case, it produces to meet the demand from Euro and domestic market. We re-set the capacity to 25k to make some of the demand from these two markets (total 27,624 including 12,376 from Euro and 15,248 from domestic market) should import from other factories. To summarize, this scenario is studied to examine the impacts of carbon tariff if we break the barrier artificially.

In this scenario, the products in both China and Taiwan will be exported to member customers. The results show that both facilities would be forced to take on high technology even when the tariff rate reaches \$0.2 per kg, as they want to reduce the carbon tariffs imposed when their products are exported to member customers. The product flow in the optimal network design ($\alpha = \$0.1$) is shown in Figure 8.1. Furthermore, both facilities will choose the low emission modules to assemble products to reduce the tariffs, too.

Another interesting issue is to study the case where a threshold minimum production rate is imposed. Facilities usually have a threshold minimum production rate to operate. In other words, if a factory operates, it cannot

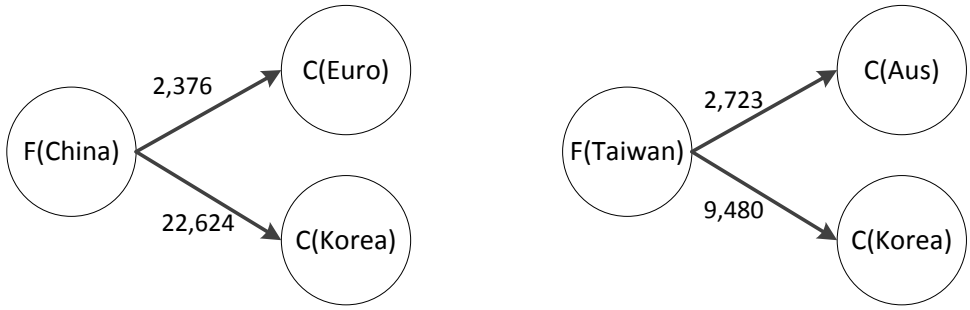

Figure 8.1.. The flow of products in the network

produce with a rate below that threshold. If the demand has to be just satisfied as Constraint (7.4) indicates, some customers may have to import from other factories under this minimum production rate restriction. Take the factory in Taiwan for example; when the threshold production rate exceeds 12,203, while the actual demand (from Australia and Korea) is only 12,203, it has no other choice but to ask the UK and the USA to cover this amount of demand. However, if Constraint (7.4) is relaxed by replacing "=" by "≥", the results are as shown in Figure 8.2. For example, under the primary carbon cap with the carbon tariff rate equals to \$0.1 per kg, the Taiwan factory would operate even if the minimum production rate was 17,200. In particular, it produces 4,997 redundant products. That is, factories may prefer to produce redundant products to satisfy the minimum production rate restriction to avoid other higher costs resulting from operating other factories. When the tariff goes up, this minimum production rate to motivate the factory to operate will decrease, as the tariffs will also increase the costs from the Taiwan factory. Moreover, for those 12,203 products which will be delivered to member customers, the low emission modules are selected; while for the 4,997 redundant products, the factory will choose the low cost yet high emission module options, as this part brings in no tariff at all. On the other hand, when the carbon cap

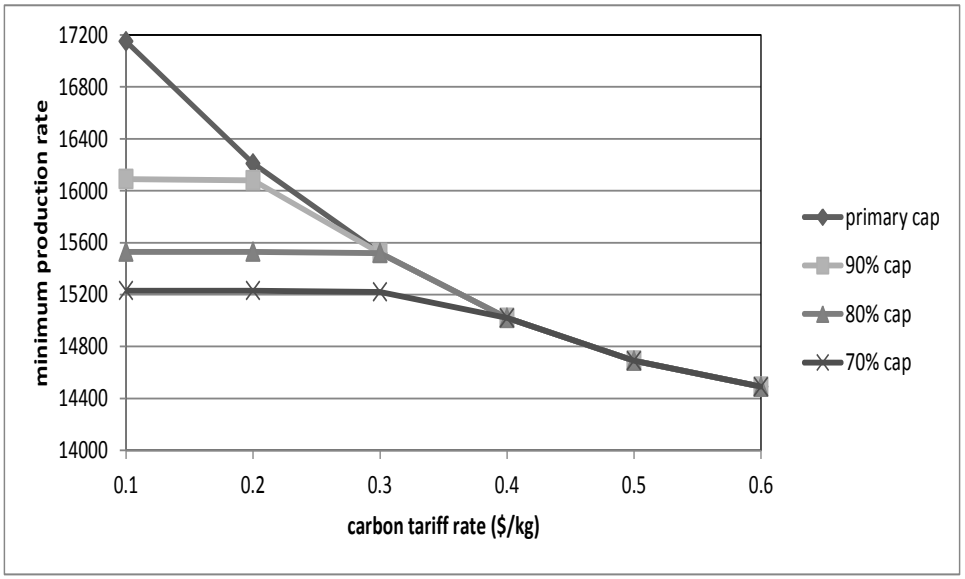

Figure 8.2.. Minimum production rate to operate a factory

is decreased by 10%, 20% and 30%, we observe that the threshold minimum production rate also decreases because producing so much redundant products may not be allowed with a stricter carbon cap.

8.1.2. The impacts on total costs and emissions

As discussed previously, the product flow exported from non-member facilities to member customers can motivate the non-member facilities to take some actions, e.g., adoption of low emission technology, usage of low emission modules and joining the member society. This section conducts another analysis to discover the impact on the total costs and emissions of such flow. The scenario under study is assumed as: total capacities of all facilities equal the total demand. The total demand of the non-member customers, i.e., China, Russia and Taiwan, is 33,173, and the total demand of the rest of member customers is 66,287. We reset the capacity of non-member facilities to be 33173+γ (capacity of China: capacity of Taiwan =5:4), and the capacity of member facilities to be 66,287−γ (capacities of

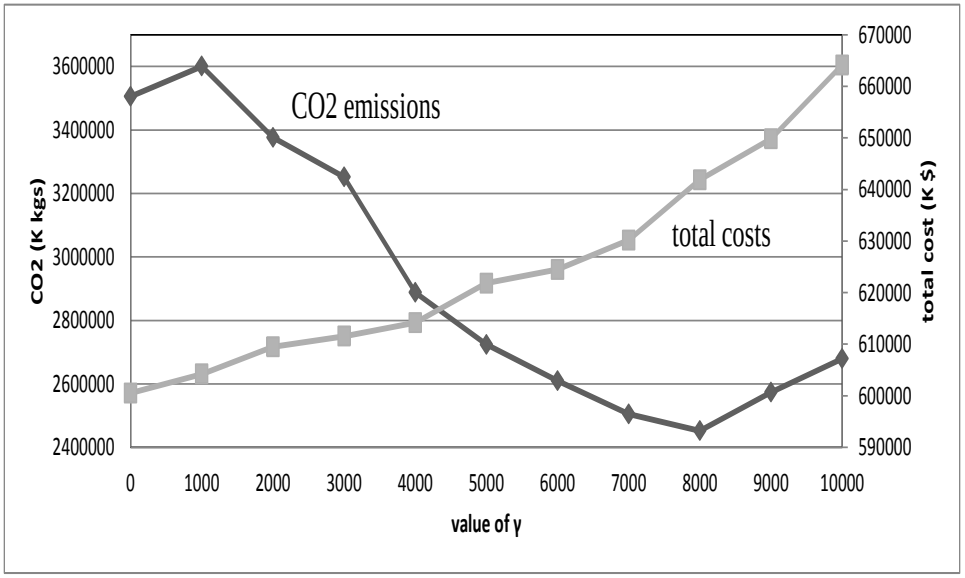

Figure 8.3.. Total costs and emissions for different values of γ

Europe: Korea: UK: USA = 1:1:3:2). The total costs and emissions for different values of γ are shown in Figure 8.3. The total emissions first increase and then decrease sharply and then increase again. The reason is that, at the first stage, the flow from non-member factories to member customers is too low to motivate the non-member factories to take high level technologies. With the flow continuing to increase, the non-member factories will update to higher technology levels to reduce the emissions and the tariff costs. At the final stage (say, γ exceeds 8,000), the factories have already reached the high level technology levels, and the emissions increase from transportation exceed the emission reduction from low emission module selection.

On the other hand, the total costs consistently increase with γ. This is intuitive because the investment cost is higher for high level technologies, lower emission modules and the transportation costs. However, it is interesting to find that the total costs curve is relatively flat within the range [0, 6,000] (increases by 3.9%), where the overall emission decreases a lot,

by about 25.6%. That is, we can reduce the CO2 emissions a lot by not increasing the costs much.

8.1.3. The impacts of supply range on the total costs with and without carbon tariffs

This section intends to examine the impacts of supply range on the costs whether or not the carbon tariff is considered. The suppliers' capacity is set as what Table 8.4 shows, where $\beta > 1$ is the range parameter, which means the supply is enlarged to $\beta \times$ the primary capacity. For different values of β, the cost reduction is shown in Figure 8.4. We can see that when the suppliers' capacity is enlarged, the cost decreases, which coincides with the results in Wang et al. (2011). Furthermore, our results indicate that the cost reduction is larger when the carbon tariff is considered because when the suppliers' capacities are larger, the network is more flexible to decrease the flow transferring modules from non-member suppliers to member facilities or member customers, and mitigate the carbon tariffs in the end. Another finding is that the cost reduction is greater when the carbon cap is looser.

Table 8.4.. The suppliers' capacities

	Euro	China	Taiwan	UK	USA
supplier's capacity	10K$\times\beta$	25K$\times\beta$	25K$\times\beta$	20K$\times\beta$	20K$\times\beta$

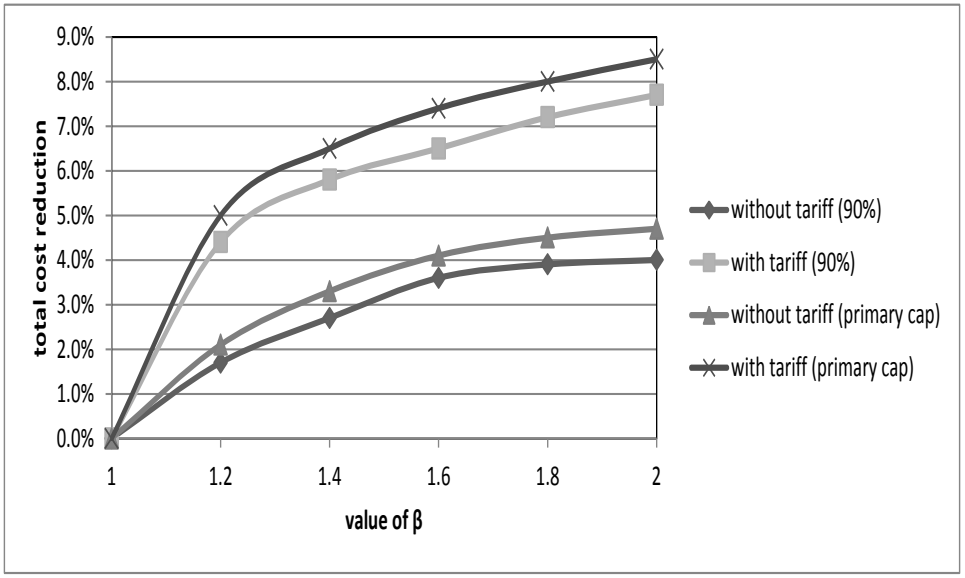

Figure 8.4.. Cost reduction for different supply ranges

8.1.4. Factory's willingness to join the member society

Finally, we are interested in exploring the willingness of non-member facilities to join the member society. The scenario above is studied to derive the results. Figure 8.5 shows the threshold cost that China's factory would like to pay for joining the member society. To explain the results, we denote the following notations:

p = the marginal profit of one product (this number is assumed to be \$100 to derive the results in Figure 8.5);

x_1 = the optimal production quantity of China's factory before joining the member society;

x_2 = the optimal production quantity of China's factory after joining the member society;

C_{mbr} = the part of membership fee allocated to Company G's factory in China if China joins the member society;

TAR_{CN} = the carbon tariffs imposed if China does not join the member society.

Company G's factory in China is willing to join the member society if

$$x_1 p - TAR_{CN} \leq x_2 p - C_{mbr}$$

In other words, if the cost allocated to Company G is less than $(x_2 - x_1)p + TAR_{CN}$, Company G prefers that China join the member society. For example, by solving the model, when the tariff rate is \$0.1, the factory in China will produce 25,000 units whether or not China belongs to the member society, and the tariff imposed when China is not a member country is \$500. Thus, it can be concluded that Company G's factory in China is willing to join the member society as long as the expense itself is less than \$500. According to Figure 8.5, we further find that non-member factory is willing to spend more to join the member society when the tariff rate increases or the total carbon cap increases. The reason why the non-member factory would like to spend more to join in the member society with a higher carbon cap is that, the higher the carbon cap, the more products the non-member factory will assemble due to the cost advantages, then the tariff imposed on such products will also be higher, therefore the factory will have more incentives to join in the member society.

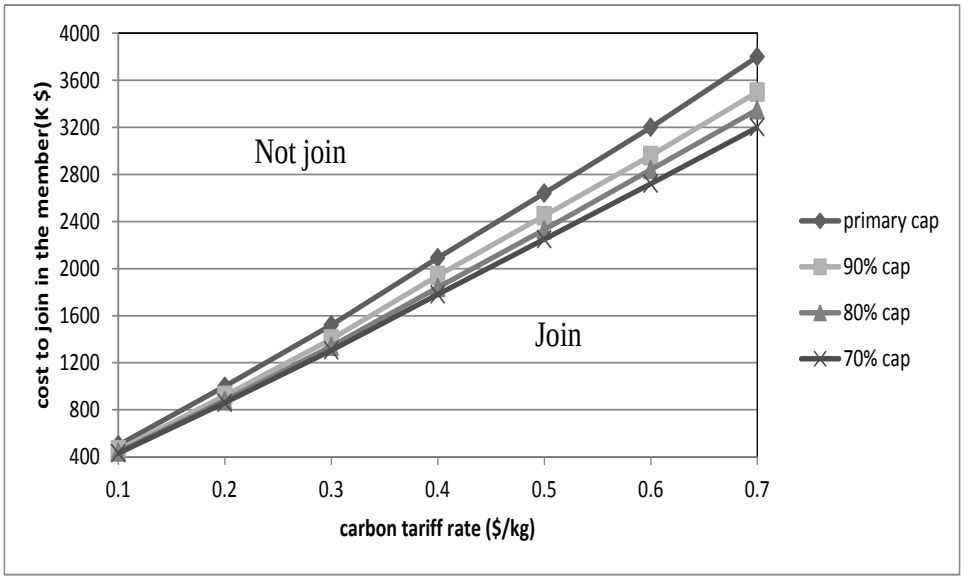

Figure 8.5.. Threshold cost that Company G's factory in China prefers to join the member society

8.2. Summary and discussions

In this chapter, we apply the model MMGSC to a real case study, i.e., Company G's case, to derive the managerial insights for its top level managers.

To summarize, only imposing carbon tariffs sometimes has little effect to force the non-member facilities to take actions to lower the carbon emissions. The member facilities will produce to satisfy the demand in the member society; and the demand in the non-member society would be satisfied by the non-member facilities themselves. In other words, the imposition of carbon tariffs seems to be like building up a barrier between member society and non-member society. We also find that facilities may produce more than the actual demand if carbon tariffs are imposed.

Another finding is that, non-member facilities are willing to spend more to join the member society when the tariff rate is higher or the total carbon cap is larger.

Through examining the relationship of total costs and emissions, we find that the CO2 emissions can be reduced a lot while the overall costs only increase a little. Moreover, when the supply range increases, the cost reduction is larger when carbon tariffs are considered than the condition where carbon tariffs are not considered. This is because when the suppliers' capacities are larger, the supply chain is more flexible to decrease the flow transferring modules from non-member suppliers to member facilities or member customers, and mitigate the carbon tariffs and total costs in the end.

Chapter 9 CONCLUSIONS AND FUTURE RESEARCH

This thesis studies two applications of the LSC models: the application in the EMS systems and the application in the GrSC network design. For the former one, our work contributes to an understanding of the benefits of demand assignment in the LSC models with service constraints. For the latter one, our work helps to shed light on studying the effects of the carbon tariff through a mathematical model and conducting a complete experimental study. This chapter summarizes and discusses the main research results in this thesis as explained in previous chapters. Potential further research directions are also presented.

9.1. Summary of research results

Part I of this thesis contributes to an understanding of the benefits of demand assignment in the LSC models with service constraints. Our model (i.e., the LSCPA) proposed for EMS systems with demand assignment incorporated can be handy for the practitioners in the location field, particularly the decision makers for the urban areas. Furthermore, the analysis for the benefits of the demand assignment is also beneficial for researchers in other fields, such as the location-inventory modelling.

Chapter 3 introduces the failure probability under demand assignment

in the LSC models in EMS systems. A failure probability FP^A is defined to derive our model with assignment. It is demonstrated that FP^A is always no larger than the failure probability estimate used in the QRLSCP and the BBKK1. Hence, using FP^A in the model always results in fewer vehicles required. In addition, we derive an upper bound for FP^A, denoted as FP^{AU}, and demonstrate that this bound asymptotically approaches to FP^A whether the workload is quite large or extremely small. The model with this upper bound also shows its advantages (results in fewer vehicles required) than the existing models when the overlap among the coverage areas is significant, or when the demand is covered by more stations. Finally, we demonstrate that FP^{AU} almost always overshoots the traditional failure probability, with the only exception that in the two station case and only one vehicle at each station (underestimation is no more than 1.06% though). This important finding verifies the viability of using FP^{AU} in the availability constraint.

Chapter 4 develops the mathematical model for the demand assignment problem which uses FP^{AU} in the availability constraint. The difficulties of using FP^{AU} in the availability constraint lie in two aspects. Firstly, the model size is enlarged because the introduction of a large number of assignment variables enlarges the model size. On the other hand, the model itself is non-linear. We then develop a tricky approach to transform the model to a linear one, $T-LSCPA$, which can be solved by current solvers like $CPLEX$. However, the linear integer model $T-LSCPA$ still belongs to the class of NP-hard optimization problems. The computation time to solve this model will exponentially increase with the number of decision

variables, therefore it may require great computation efforts for solving large scale problems to achieve the global optimal. For example, the efficiency of $CPLEX$ in solving the proposed model may be quite unstable, which is dependent on whether or not the LP relaxation and cutting plane added at the nodes are tight. That is, generally it is difficult to predict the efficiency of solving the model in advance. On the other hand, some flexible constraints exist in the model $T - LSCPA$, like the availability constraint (4.5). The exact algorithm (like branch-and-cut algorithm used by CPLEX) needs to resolve the model even in case of only a little bit change to one such flexible constraint. This may be very time consuming and could be expensive for real applications. We are motivated to develop a heuristic which is not so sensitive to little changes in some parameters. To address the computation time issue, we develop a heuristic which is decomposed into two parts and has only polynomial computation time and good performance as well. As a result, changes to some parameters may only lead to resolving one part instead of totally resolving the entire problem. With similar arguments, the heuristic is also helpful when the vehicles have to be located dynamically.

Chapter 5 runs many numerical examples and obtains a series of interesting observations, which in turn verify the analytical results in chapters 3 and 4. We observe that the cost saving of our model consistently outperforms the results of the BBKK1. In addition, such saving also increases with the coverage radius and the availability level target, but decreases with the ratio of C_1/C_2. In more details, when the overlap among the coverage areas is significant, or when the demand is covered by more stations,

our model with assignment requires fewer vehicles than the BBKK1 and the QRLSCP. Compared to the BRLSCP, our model performs better when the overlap is large. We also propose a modification procedure to make the solutions of our model to be applicable to practice. In particular, the solutions of our model are modified to ensure the availability almost achieved under practical closest available vehicle policy. Finally, the solutions of our heuristic are insensitive to the change of the inputs. Such insensitivity further supports the application of the heuristic under the conditions where the availability standard changes, the population or traffic conditions are not stable.

Part II of this thesis studies the effects of the carbon tariff via constructing a mathematical model and then applying this model in a real case study.

Chapter 6 is about the problem statement for the GrSC network design. In this chapter, we highlight two main topics for the LSC problem in designing the GrSC network: module selection and the carbon tariff mechanism. Regarding the first one, it is assumed that several options with different properties can be selected for each module of a product. Moreover, these products assembled with different options are treated equally by the consumers. Managers therefore can find a trade-off among those properties and decide which option should be selected for each module. The other critical issue is about how the carbon tariffs should be imposed. Section 6.4 clearly defines this mechanism based on some simplifying assumptions, which makes it possible to incorporate the carbon tariff mechanism into a mathematical model and to investigate the impacts of such a mechanism

via applying the model in a real case study.

Chapter 7 derives the mathematical model MMGSC based on the analysis in Chapter 6, which takes into consideration carbon emissions during usage as well as the carbon tariffs when designing the GrSC network. As the original model is non-linear, we propose a tricky way to transform it to be a mixed integer linear programming model.

Chapter 8 applies the model MMGSC to a real case study, i.e., Company G's case, to derive managerial insights for its managers. It is concluded that only imposing carbon tariffs may not motivate non-member facilities to adopt low emission technologies due to the barrier set up by the tariff between member society and non-member society. However, this motivation can be realized if the capacity of member facilities is restricted in the meantime. Furthermore, we also study the willingness of non-member facilities to join the member society, i.e. the maximum expense the facilities would like to pay for joining the member society. Finally, we find that enlarging the suppliers' capacity will bring about a higher cost reduction when carbon tariffs are imposed than when carbon tariffs are not considered, due to the flexibility on the supply chain brought by the increase of the supply range.

9.2. Limitations of this study and future research

This section lists the limitations or assumptions of this study, and indicates some potential future research directions accordingly. Part of I the thesis:

1. In the application of LSC models in EMS systems, we discussed the conditions where our estimated failure probability, i.e., failure prob-

ability with assignment, possibly undershoots the traditional failure probability (under closest available vehicle policy) which is the common performance target in practice. Moreover, with our demand assignment, not always the closest available vehicles would respond the emergency calls. However, this is not allowed in EMS systems sometimes. A possible direction is to add another constraint which restricts the assigned station to each demand node to be the closest one among all activated stations. This would inevitably increase the total costs due to the stricter constraint set. However, it would be interesting to investigate under what conditions the solutions with this additional constraint will also lead to better performance than the existing models. Besides, whether the new model with this additional constraint always overshoots the traditional failure probability or not is also of interest.

2. The emergency calls share the same urgency degree in our study. We can extend our model to the situation where different urgency degrees of the emergency calls are taken into consideration. We then assign ambulances according to the urgency degrees (usually on a one-to-four scale) of the calls. Note that the coverage radii for the calls of different urgency degrees may be different.

3. The current heuristic can be regarded as the inner algorithm. We can develop an outer algorithm in order to modify the set of stations to be activated, targeting at finding global optimal or local "good enough" solutions. This inevitably would lead to the computation time increase, so how to achieve a trade-off between the computation

time and the accuracy of heuristic is an interesting topic.

Part II of this thesis:

1. In our proposed model, the carbon tariff is linear related to the emissions, which may not be the case sometimes. We can extend our work to explore those cases where the relationship is not a simple linear one, as long as the carbon tariff mechanism can be clearly stated, as was the case in Section 6.4.

2. This study assumes that the carbon tariff rate in different countries is equal. A possible extension of this study is how to allocate customers to factories if the tariff rates in different countries are not identical.

3. We consider deterministic demand and one period design in this study. Multi-period supply chain design and stochastic demand requirement are interesting topics to investigate where inventory and lead time have to be considered.

4. We do not consider batch size in our study. The results in Section 8.1.1 may be different when this is considered.

Bibliography

Abdallah, T., Diabat, A. and Rigter, J. (2013) Investigating the option of installing small scale pvs on facility rooftops in a green supply chain. *International Journal of Production Economics*, **146**(2), 465–477.

Alsalloum, O. and Rand, G. (2006) Extensions to emergency vehicle location models. *Computers and Operations Research*, **33**(9), 2725–2743.

Aly, A. and White, J. (1978) Probabilistic formulation of the emergency service location problem. *The Journal of the Operational Research Society*, **29**(12), 1167–1179.

Australian Government, . (2013) *How Australia's carbon price is working one year on.*, Available from *http* : *//creativecommons.org/licenses/by/*3.0/*au*. Accessed July 20 2011.

Ball, M. and Lin, F. (1993) A reliability model applied to emergency service vehicle location. *Operations Research*, **41**(1), 18–36.

Baron, O., Berman, O., Kim, S. and Krass, D. (2009) Ensuring feasibility in location problems with stochastic demands and congestion. *IIE Transactions*, **41**(5), 467–481.

Beamon, B.M. (1999) Designing the green supply chain. *Logistics Information Management*, **12**(4), 332–342.

Benjaafar, S., Li, Y. and Daskin, M. (2013) Carbon footprint and the management of supply chains: Insights from simple models. *IEEE Transactions on Automation Science and Engineering*, **10**(1), 99–116.

Beraldi, P. and Bruni, M. (2009) A probabilistic model applied to emergency service vehicle location. *European Journal of Operational Research*, **196**(1), 323–331.

Berman, O., Drezner, Z. and Krass, D. (2009) Cooperative cover location problems: the planar case. *IIE Transactions*, **42**(3), 232–246.

Berman, O. and Krass, D. (2002) The generalized maximal covering location problem. *Computers and Operations Research*, **29**(6), 563–581.

Böhringer, C., Carbone, J.C. and Rutherford, T.F. (2013) The strategic value of carbon tariffs. *CESifo Working paper*, .

Borrás, F. and Pastor, J. (2002) The ex-post evaluation of the minimum local reliability level: an enhanced probabilistic location set covering model. *Annals of Operations Research*, **111**(1), 51–74.

British Standards Institute, . (2008) *PAS 2050: 2008, Specification for the Assessment of the Life Cycle Greenhouse Gas Emissions of Goods and Services.*, .

Broder, J. (2009) Obama opposes trade sanctions in climate bill. Available from New York Times at *http* : *//www.nytimes.com*/2009/06/29/*us*/*politics*/29*climate.html*. Accessed June 2 2013, .

Brotcorne, L., Laporte, G. and Semet, F. (2003) Ambulance location and relocation models. *European Journal of Operational Research*, **147**(3), 451–463.

Budge, S., Ingolfsson, A. and Erkut, E. (2009) Approximating vehicle dispatch probabilities for emergency service systems with location-specific service times and multiple units per location. *Operations Research*, **57**(1), 251–255.

Chaabane, A., Ramudhin, A. and Paquet, M. (2012) Design of sustainable supply chains under the emission trading scheme. *International Journal of Production Economics*, **135**(1), 37–49.

Chapman, S. and White, J. (1974) Probabilistic formulations of emergency service facility location problems. *ORSA/TIMS Paper. San Juan, Puerto Rico*, .

Chen, X., Benjaafar, S. and Elomri, A. (2013) The carbon-constrained eoq. *Operations Research Letters*, **41**(2), 172–179.

Chen, X. and Hao, G. (2014) Sustainable pricing and production policies for two competing firms with carbon emissions tax. *International Journal of Production Research*, (ahead-of-print), 1–13.

Church, R. and ReVelle, C. (1974) The maximal covering location problem. *Papers in Regional Science*, **32**(1), 101–118.

Church, R. and Weaver, J. (1986) Theoretical links between median and coverage location problems. *Annals of Operations Research*, **6**(1), 1–19.

Corcoran, T. (2008) Carbon Tariff Trade War?. Available from National Post at $http://www.nationalpost.com/opinion/columnists/story.html?id=6dcb1379-42e1-4200-a656-cef828629182\&k=92193$. Accessed May 7 2013.

Cosbey, A. (2008) Border carbon adjustment. In: *IISD Background Paper for*

the Trade and Climate Change Seminar, June, pp. 18–20.

Cristea, A., Hummels, D., Puzzello, L. and Avetisyan, M. (2013) Trade and the greenhouse gas emissions from international freight transport. *Journal of Environmental Economics and Management*, **65**(1), 153–173.

Daskin, M. (1983) A maximum expected covering location model: formulation, properties and heuristic solution. *Transportation Science*, **17**(1), 48–70.

Daskin, M. (1995) *Network and discrete location: models, algorithms, and applications.*, John Wiley & Sons.

Daskin, M. and Dean, L. (2005) Location of health care facilities. *Operations Research and Health Care. Springer US*, pp. 43–76.

De Maio, V., Stiell, I., Wells, G. and Spaite, D. (2003) Optimal defibrillation response intervals for maximum out-of-hospital cardiac arrest survival rates. *Annals of Emergency Medicine*, **42**(2), 242–250.

Dekker, R., Bloemhof, J. and Mallidis, I. (2012) Operations research for green logistics–an overview of aspects, issues, contributions and challenges. *European Journal of Operational Research*, **219**(3), 671–679.

Demailly, D., Quirion, P. et al. (2005) *Leakage from climate policies and border tax adjustment: lessons from a geographic model of the cement industry*, CESifo Venice Summer Institute, July.

Diabat, A. and Al-Salem, M. (2015) An integrated supply chain problem with environmental considerations. *International Journal of Production Economics*, **164**(1), 330–338.

Dong, Y., Ishikawa, M. and Hagiwara, T. (2015) Economic and environmental

impact analysis of carbon tariffs on chinese exports. *Energy Economics*, **50**(3), 80–95.

Drake, D.F. (2011) Carbon tariffs: Impacts on technology choice, regional competitiveness, and global emissions. Tech. rep., Harvard Business School, working paper, available from *http* : *//hbswk.hbs.edu/item/*6755.*html*. Accessed May 11 2013.

Enkvist, P., Nauclér, T. and Rosander, J. (2007) A cost curve for greenhouse gas reduction. *McKinsey Quarterly*, **1**(1), 35–45.

Erkut, E., Ingolfsson, A. and Erdoğan, G. (2008) Ambulance location for maximum survival. *Naval Research Logistics (NRL)*, **55**(1), 42–58.

Erlang, A. (1917) Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges. *Elektrotkeknikeren*, **13**(1), 138–155.

Farahani, R.Z., Hassani, A., Mousavi, S.M. and Baygi, M.B. (2014) A hybrid artificial bee colony for disruption in a hierarchical maximal covering location problem. *Computers & Industrial Engineering*, **75**(1), 129–141.

Fitch, J. (2005) Response times: myths, measurement and management. *Journal of Emergency Medical Services*, **30**(9), 46–56.

Fleischmann, M., Beullens, P., Bloemhof-Ruwaard, J.M. and Wassenhove, L.N. (2001) The impact of product recovery on logistics network design. *Production and Operations Management*, **10**(2), 156–173.

Gielen, D. and Moriguchi, Y. (2002) Co2 in the iron and steel industry: an analysis of japanese emission reduction potentials. *Energy Policy*, **30**(10), 849–863.

Gleason, J. (1975) A set covering approach to bus stop location. *Omega*, **3**(5), 605–608.

Goetschalcks, M. and Fleischmann, B. (2008) Strategic network design. In: *Supply chain management and advanced planning*, pp. 117–132. Springer.

Goetschalckx, M., Vidal, C.J. and Dogan, K. (2002) Modeling and design of global logistics systems: A review of integrated strategic and tactical models and design algorithms. *European Journal of Operational Research*, **143**(1), 1–18.

Gros, D., Egenhofer, C., Fujiwara, N., Guerin, S. and Georgiev, A. (2010) *Climate Change and Trade: Taxing carbon at the border?*, CEPS Paperbacks. Available from: $http://www.ceps.eu/book/climate-change-and-trade-taxing-carbon-border$. Accessed May 8 2013.

Houser, T., Bradley, R., Staley, B.C., Werksman, J. and Heilmayr, R. (2008) *Leveling the carbon playing field: international competition and US climate policy design*, Peterson Institute for International Economics. Available from $http://www.wri.org/publication/leveling-the-carbon-playing-field$. Accessed May 9 2013.

Hua, G., Cheng, T. and Wang, S. (2011) Managing carbon footprints in inventory management. *International Journal of Production Economics*, **132**(2), 178–185.

Hwang, H. (2004) A stochastic set-covering location model for both ameliorating and deteriorating items. *Computers and Industrial Engineering*, **46**(2), 313–319.

Institute of Medicine of the National Academies, . (2006) *Emergency Medical*

Services: At the Crossroads, National Academy of Sciences, Washington, D.C.

Institute of Public Affairs, . (2009) *Costly, ineffectual and protectionist carbon tariffs.*, Available from :*http* : *//sustainabledev.org/wp − content/uploads/*2009/*Carbontariffs.pdf*. Accessed May 10 2013.

Jaber, M.Y., Glock, C.H. and El Saadany, A.M. (2013) Supply chain coordination with emissions reduction incentives. *International Journal of Production Research*, **51**(1), 69–82.

Kuik, O. and Hofkes, M. (2010) Border adjustment for european emissions trading: Competitiveness and carbon leakage. *Energy Policy*, **38**(4), 1741–1748.

Kuo, T.C., Huang, S.H. and Zhang, H.C. (2001) Design for manufacture and design for x: concepts, applications, and perspectives. *Computers & Industrial Engineering*, **41**(3), 241–260.

Lai, H. (2012) *Ziqitza Health Care introduces ambulance outsourcing service.*, Available from SME news at *http* : *//news.indiamart.com/story/ziqitza − health − care − introduces − ambulance − outsourcing − service −* 161007.*html*. Accessed June 8 2012.

Marianov, V. and Serra, D. (1998) Probabilistic, maximal covering location allocation models for congested systems. *Journal of Regional Science*, **38**(3), 401–424.

Marianov, V. and Serra, D. (2002) Location–allocation of multiple-server service centers with constrained queues or waiting times. *Annals of Operations Research*, **111**(1-4), 35–50.

Martí, J.M.C., Tancrez, J.S. and Seifert, R.W. (2015) Carbon footprint and re-

sponsiveness trade-offs in supply chain network design. *International Journal of Production Economics*, **166**(1), 129–142.

Mayorga, M.E., Bandara, D. and McLay, L.A. (2013) Districting and dispatching policies for emergency medical service systems to improve patient survival. *IIE Transactions on Healthcare Systems Engineering*, **3**(1), 39–56.

McLay, L.A. and Mayorga, M.E. (2013) A dispatching model for server-to-customer systems that balances efficiency and equity. *Manufacturing & Service Operations Management*, **15**(2), 205–220.

Monjon, S. and Quirion, P. (2010) How to design a border adjustment for the european union emissions trading system? *Energy Policy*, **38**(9), 5199–5207.

Monjon, S. and Quirion, P. (2011) A border adjustment for the eu ets: Reconciling wto rules and capacity to tackle carbon leakage. *Climate Policy*, **11**(5), 1212–1225.

Moore, M.O. (2011) Implementing carbon tariffs: A fool's errand? *The World Economy*, **34**(10), 1679–1702.

O'Connell, S. and Stutz, M. (2010) Product carbon footprint (pcf) assessment of dell laptop-results and recommendations. In: *Sustainable Systems and Technology (ISSST), 2010 IEEE International Symposium on*, pp. 1–6. IEEE.

Ong, M., Ho, K., Tan, T., Koh, S., Almuthar, Z., Overton, J. and Lim, S. (2009) Using demand analysis and system status management for predicting ED attendances and rostering. *American Journal of Emergency Medicine*, **27**(1), 16–22.

Paterson, C., Kiesmüller, G., Teunter, R. and Glazebrook, K. (2011) Inventory models with lateral transshipments: A review. *European Journal of Operational Research*, **210**(2), 125–136.

Pereira, M.A., Coelho, L.C., Lorena, L.A. and De Souza, L.C. (2015) A hybrid method for the probabilistic maximal covering location–allocation problem. *Computers & Operations Research*, **57**(1), 51–59.

Persson, S. (2010) Practical aspects of border carbon adjustment measures–using a trade facilitation perspective to assess trade costs. *ICTSD Global Platform on Climate Change, Trade Policies and Sustainable Energy, Issue Paper No 13. International Centre for Trade and Sustainable Development, Geneva & National Board of Trade, Stockholm.*, (13).

Plane, D. and Hendrick, T. (1977) Mathematical programming and the location of fire companies for the denver fire department. *Operations Research*, **25**(4), 563–578.

Rajagopalan, H., Saydam, C. and Xiao, J. (2008) A multiperiod set covering location model for dynamic redeployment of ambulances. *Computers and Operations Research*, **35**(3), 814–826.

Ramudhin, A., Chaabane, A. and Paquet, M. (2010) Carbon market sensitive sustainable supply chain network design. *International Journal of Management Science and Engineering Management*, **5**(1), 30–38.

Reinaud, J. (2008) Issues behind competitiveness and carbon leakage. *Focus on Heavy Industry. Paris: IEA. IEA Information Paper*, **2**.

Restrepo, M., Henderson, S. and Topaloglu, H. (2009) Erlang loss models for the static deployment of ambulances. *Health Care Management Science*, **12**(1),

67–79.

ReVelle, C. and Hogan, K. (1988) A reliability-constrained siting model with local estimates of busy fractions. *Environment and Planning B: Planning and Design*, **15**(3), 143–152.

Revelle, C. and Hogan, K. (1989) The maximum reliability location problem and α-reliablep-center problem: Derivatives of the probabilistic location set covering problem. *Annals of Operations Research*, **18**(1), 155–173.

ReVelle, C. and Hogan, K. (1989) The maximum availability location problem. *Transportation Science*, **23**(3), 192.

Savas, E. (1969) Simulation and cost-effectiveness analysis of New York's emergency ambulance service. *Management Science*, **15**(12), 608–627.

Savaskan, R.C., Bhattacharya, S. and Van Wassenhove, L.N. (2004) Closed-loop supply chain models with product remanufacturing. *Management Science*, **50**(2), 239–252.

Schilling, D., Jayaraman, V. and Barkhi, R. (1993) A review of covering problems in facility location. *Location Science*, **1**(1), 25–55.

Schreuder, J. (1981) Application of a location model to fire stations in rotterdam. *European Journal of Operational Research*, **6**(2), 212–219.

Shen, Z., Coullard, C. and Daskin, M. (2003) A joint location-inventory model. *Transportation Science*, **37**(1), 40–55.

Sindico, F. (2008) The eu and carbon leakage: how to reconcile border adjustments with the wto? *European Energy and Environmental Law Review*, **17**(6), 328–340.

Snyder, L.V. (2006) Facility location under uncertainty: a review. *IIE Transactions*, **38**(7), 547–564.

Solomon, S. (2007) *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC*, vol. 4, Cambridge University Press.

Sorensen, P. and Church, R. (2010) Integrating expected coverage and local reliability for emergency medical services location problems. *Socio-Economic Planning Sciences*, **44**(1), 8–18.

Sošić, G. (2006) Transshipment of inventories among retailers: Myopic vs. farsighted stability. *Management Science*, **52**(10), 1493–1508.

Srivastava, S.K. (2007) Green supply-chain management: a state-of-the-art literature review. *International Journal of Management Reviews*, **9**(1), 53–80.

Swain, R. (1974) A parametric decomposition approach for the solution of uncapacitated location problems. *Management Science*, **21**(2), 189–198.

Tognetti, A., Grosse-Ruyken, P.T. and Wagner, S.M. (2015) Green supply chain network optimization and the trade-off between environmental and economic objectives. *International Journal of Production Economics, in press*, .

Toregas, C. and Revelle, C. (1973) Binary logic solutions to a class of location problem*. *Geographical Analysis*, **5**(2), 145–155.

Toregas, C., Swain, R., ReVelle, C. and Bergman, L. (1971) The location of emergency service facilities. *Operations Research*, **19**(6), 1363–1373.

Toro-Díaz, H., Mayorga, M.E., Chanta, S. and McLay, L.A. (2013) Joint location

and dispatching decisions for emergency medical services. *Computers & Industrial Engineering*, **64**(1), 917–928.

Van Asselt, H. and Brewer, T. (2010) Addressing competitiveness and leakage concerns in climate policy: An analysis of border adjustment measures in the us and the eu. *Energy Policy*, **38**(1), 42–51.

Walker, W. (1974) Using the set-covering problem to assign fire companies to fire houses. *Operations Research*, **22**(2), 275–277.

Wang, F., Lai, X. and Shi, N. (2011) A multi-objective optimization for green supply chain network design. *Decision Support Systems*, **51**(2), 262–269.

Waxman, R.H.A. and Markey, R.E.J. (2009) American clean energy and security act of 2009. *Washington: US House of Representatives*, .

Wee, K. and Dada, M. (2005) Optimal policies for transshipping inventory in a retail network. *Management Science*, **51**(10), 1519–1533.

Wolff, R. (1982) Poisson arrivals see time averages. *Operations Research*, **30**(2), 223–231.

Yang, J. and Qin, Z. (2007) Capacitated production control with virtual lateral transshipments. *Operations Research-baltimore then Linthicum*, **55**(6), 1104.

Ye, L., Ye, C. and Chuang, Y. (2011) Location set covering for waste resource recycling centers in taiwan. *Resources, Conservation and Recycling*, **55**(11), 979–985.

Zhou, X., Kojima, S. and Yano, T. (2010) *Addressing carbon leakage by border adjustment measures review of current studies*, Institute for Global Environmental Studies, Hayama, Japan.

APPENDICES

A. Proof of Lemma 1

Proof. Due to the Poisson arrivals, the arrival rate at station i is $\lambda_i = \sum_{j \in S_{PT}(i)} \lambda_j y_{ij}$ with assignment.

We derive the average service time at station i as follows. We notice that the service time at station i is s_{ij} which are assumed to be continuous random variables following any general distribution when the demand originates from node j with a probability of $\frac{\lambda_j y_{ij}}{\sum_{j \in S_{PT}(i)} \lambda_j y_{ij}}$ for all possible nodes $\{j \in S_{PT}(i) | y_{ij} = 1\}$. That means the service time at station i is a mixture distribution of the service times for all possible nodes assigned to station i. According to the properties of the mixture distribution, the service time at station i is a convex combination of the service times for all possible nodes assigned to station i.

On the other hand, the arrivals originate from node j with a probability of $\frac{\lambda_j y_{ij}}{\sum_{j \in S_{PT}(i)} \lambda_j y_{ij}}$. Correspondingly, the average service time is $\frac{1}{\mu_{ij}}$ with a probability of $\frac{\lambda_j y_{ij}}{\sum_{j \in S_{PT}(i)} \lambda_j y_{ij}}$. Hence, the average service time at station i is also a convex combination of $\frac{1}{\mu_{ij}}$, i.e., $\frac{\sum_{j \in S_{PT}(i)} \frac{\lambda_j}{\mu_{ij}} y_{ij}}{\sum_{j \in S_{PT}(i)} \lambda_j y_{ij}}$. Equivalently, the average service rate at station i, denoted as μ_i, is $\frac{\sum_{j \in S_{PT}(i)} \lambda_j y_{ij}}{\sum_{j \in S_{PT}(i)} \frac{\lambda_j}{\mu_{ij}} y_{ij}}$. As a

result,

$$\rho_i^A = \lambda_i/\mu_i = \sum_{j \in S_{PT}(i)} \frac{\lambda_j}{\mu_{ij}} y_{ij}.$$

Hence, the lemma holds. □

B. Proof of Proposition 1

Proof. Let $P_X(\rho_i^A, n)$ be the probability density function (PDF) of X, and $F_X(\rho_i^A, n)$ be the cumulative density function (CDF) of X, where $X \sim POISSON(\rho_i^A)$ and n is the non-negative integer parameter. The failure probability with assignment FP_i^A and right hand side of inequalities (3.2) FP_i^{AU} are reformulated as follows.

$$FP_i^A = \frac{(\rho_i^A)^{z_i}/z_i!}{\sum_{s=0}^{z_i}(\rho_i^A)^s/s!} = \frac{e^{-\rho_i^A}(\rho_i^A)^{z_i}/z_i!}{\sum_{s=0}^{z_i} e^{-\rho_i^A}(\rho_i^A)^s/s!} = \frac{P_X(\rho_i^A, z_i)}{F_X(\rho_i^A, z_i)}$$

$$FP_i^{AU} = Pr(\sum_{j \in S_{PT}(i)} D_j(1/\mu_{ij})y_{ij} \geq z_i) = Pr(\widehat{D}_i \geq z_i) = 1 - F_X(\rho_i^A, z_i - 1)$$

Then, we have

$$
\begin{aligned}
& P_X(\rho_i^A, z_i) \leq F_X(\rho_i^A, z_i) \\
\Longrightarrow \quad & P_X(\rho_i^A, z_i)[1 - F_X(\rho_i^A, z_i)] \leq F_X(\rho_i^A, z_i)[1 - F_X(\rho_i^A, z_i)] \\
\Longrightarrow \quad & P_X(\rho_i^A, z_i) \leq F_X(\rho_i^A, z_i) - F_X^2(\rho_i^A, z_i) + P_X(\rho_i^A, z_i) F_X(\rho_i^A, z_i) \\
\Longrightarrow \quad & \frac{P_X(\rho_i^A, z_i)}{F_X(\rho_i^A, z_i)} \leq 1 - F_X(\rho_i^A, z_i) + P_X(\rho_i^A, z_i) \\
\Longrightarrow \quad & \frac{P_X(\rho_i^A, z_i)}{F_X(\rho_i^A, z_i)} \leq 1 - F_X(\rho_i^A, z_i - 1) \\
\Longrightarrow \quad & FP_i^A \leq FP_i^{AU}
\end{aligned}
$$

Hence, the lemma holds. □

C. Proof of Proposition 2

Firstly, we show that $\forall i \in I$,

$$
\begin{aligned}
\frac{\partial P_X(\rho^A, z)}{\partial \rho^A} &= \frac{\partial}{\partial \rho^A} \frac{e^{-\rho^A}(\rho^A)^z}{z!} \\
&= -\frac{e^{-\rho^A}(\rho^A)^z}{z!} + \frac{e^{-\rho^A}(\rho^A)^{z-1}}{(z-1)!} \\
&= P_X(\rho^A, z-1) - P_X(\rho^A, z) \\
&= (\frac{z}{\rho^A} - 1) P_X(\rho^A, z).
\end{aligned}
$$

When the same logic is applied for each term of $F_X(\rho^A, z)$, we get

$$\frac{\partial F_X(\rho^A, z)}{\partial \rho^A} = F_X(\rho^A, z-1) - F_X(\rho^A, z) = -P_X(\rho^A, z).$$

When $\rho^A \to 0$, $FP^{AU} - FP^A$ approaches to zero because both FP^{AU} and FP^A approach to zero, which indicates our bound FP^{AU} asymptotically converges to the failure probability with assignment FP^A.

And,

$$\begin{aligned}
\lim_{\rho^A \to 0} \frac{FP^{AU}}{FP^A} &= \lim_{\rho^A \to 0} \frac{F_X(\rho^A, z)[1 - F_X(\rho^A, z-1)]}{P_X(\rho^A, z)} \\
&= \lim_{\rho^A \to 0} \frac{-P_X(\rho^A, z)[1 - F_X(\rho^A, z-1)] + \frac{z}{\rho^A} F_X(\rho, z)}{(\frac{z}{\rho^A} - 1) P_X(\rho^A, z)} \\
&= \lim_{\rho^A \to 0} \frac{\frac{z}{\rho^A} - 1}{\frac{z}{\rho^A}} \\
&= 1. \qquad \text{(C.1)}
\end{aligned}$$

On the other hand, when $\rho^A \to \infty$,

$$\begin{aligned}
&\lim_{\rho^A \to \infty} FP^{AU} - FP^A \\
= &\lim_{\rho^A \to \infty} 1 - F_X(\rho^A, z-1) - \frac{P_X(\rho^A, z)}{F_X(\rho^A, z)} \\
= &\lim_{\rho^A \to \infty} \frac{F_X(\rho^A, z) - F_X(\rho^A, z-1) F_X(\rho^A, z) - P_X(\rho^A, z)}{F_X(\rho^A, z)} \\
= &\lim_{\rho^A \to \infty} \frac{-P_X(\rho^A, z)[1 + F_X(\rho^A, z-1) + (\frac{z}{\rho^A} - 1)] + P_X(\rho^A, z-1) F_X(\rho^A, z)}{-P_X(\rho^A, z)} \\
= &\lim_{\rho^A \to \infty} \frac{-[1 + F_X(\rho^A, z-1) + (\frac{z}{\rho^A} - 1)] + (\frac{z}{\rho^A}) F_X(\rho^A, z)}{-1} \\
= &\ 0. \qquad \text{(C.2)}
\end{aligned}$$

And,

$$\begin{aligned}
\lim_{\rho^A \to \infty} \frac{FP^{AU}}{FP^A} &= \lim_{\rho^A \to \infty} \frac{F_X(\rho^A, z)[1 - F_X(\rho^A, z-1)]}{P_X(\rho^A, z)} \\
&= \lim_{\rho^A \to \infty} \frac{-P_X(\rho^A, z)[1 - F_X(\rho^A, z-1)] + P_X(\rho^A, z-1)F_X(\rho^A, z)}{(\frac{z}{\rho^A} - 1)P_X(\rho^A, z)} \\
&= \lim_{\rho^A \to \infty} \frac{-[1 - F_X(\rho^A, z-1)] + \frac{z}{\rho^A}F_X(\rho^A, z)}{(\frac{z}{\rho^A} - 1)} \\
&= 1, \qquad \text{(C.3)}
\end{aligned}$$

where L'Hopital's Rule is applied in $(C.1)$, $(C.2)$, and $(C.3)$.

D. Transition Rate Matrix for the two station case: $z = 2$, shown in Table D.1

E. Transition Rate Matrix for the three station case

E.1. 3 station case, $z = 1$, $o = 0$, shown in Table E.1

E.2. 3 station case, $z = 1$, $0 < o \leq 0.346$, shown in Table E.2

E.3. 3 station case, $z = 1$, $0.346 < o \leq 0.973$, shown in Table E.3

Table D.1.. The transition rate matrix (2 stations, $z = 2$)

$z = 2$	00	01	02	10	11	12	20	21	22
00	$-\lambda(2-o)$	$\lambda(1-0.5o)$		$\lambda(1-0.5o)$					
01	1	$-\lambda(2-o)-1$	$\lambda(1-0.5o)$		$\lambda(1-0.5o)$				
02		2	$-(2+\lambda)$			λ			
10	1			$-\lambda(2-o)-1$	$\lambda(1-0.5o)$		$\lambda(1-0.5o)$		
11		1		1	$-\lambda(2-o)-2$	$\lambda(1-0.5o)$		$\lambda(1-0.5o)$	
12			1		2	$-(3+\lambda)$			λ
20				2			$-(2+\lambda)$	λ	
21					2		1	$-(3+\lambda)$	λ
22						2		2	-4

Table E.1.. The transition rate matrix (3 stations, $z = 1$, $o = 0$)

$S1$	000	001	010	011	100	101	110	111
000	-3λ	λ	λ	0	λ	0	0	0
001	1	$-(2\lambda+1)$	0	λ	0	λ	0	0
010	1	0	$-(2\lambda+1)$	λ	0	0	λ	0
011	0	1	1	$-(2+\lambda)$	0	0	0	λ
100	1	0	0	0	$-(2\lambda+1)$	λ	λ	0
101	0	1	0	0	1	$-(2+\lambda)$	0	λ
110	0	0	1	0	1	0	$-(2+\lambda)$	λ
111	0	0	0	1	0	1	1	-3

Table E.2.. The transition rate matrix (3 stations, $z = 1$, $0 < o \leq 0.346$)

$S2$	000	001	010	011	100	101	110	111
000	$-(3\lambda - 1.5o)$	$\lambda(1 - 0.5o)$	$\lambda(1 - 0.5o)$	0	$\lambda(1 - 0.5o)$	0	0	0
001	1	$-(2\lambda - 0.5o + 1)$	0	$\lambda(1 - 0.25o)$	0	$\lambda(1 - 0.25o)$	0	0
010	1	0	$-(2\lambda - 0.5o + 1)$	$\lambda(1 - 0.25o)$	0	0	$\lambda(1 - 0.25o)$	0
011	0	1	1	$-(2 + \lambda)$	0	0	0	λ
100	1	0	0	0	$-(2\lambda - 0.5o + 1)$	$\lambda(1 - 0.25o)$	$\lambda(1 - 0.25o)$	0
101	0	1	0	0	1	$-(2 + \lambda)$	0	λ
110	0	0	1	0	1	0	$-(2 + \lambda)$	λ
111	0	0	0	1	0	1	1	-3

Table E.3.. The transition rate matrix (3 stations, $z = 1$, $0.346 < o \leq 0.973$

$S3$	000	001	010	011	100	101	110	111
000	$-(6b+3c+a)$	$2b+c+a/3$	$2b+c+a/3$	0	$2b+c+a/3$	0	0	0
001	1	$-(6b+2c+a+1)$	0	$3b+c+a/2$	0	$3b+c+a/2$	0	0
010	1	0	$-(6b+2c+a+1)$	$3b+c+a/2$	0	0	$3b+c+a/2$	0
011	0	1	1	$-(2+a+4b+c)$	0	0	0	$a+4b+c$
100	1	0	0	0	$-(6b+2c+a+1)$	$3b+c+a/2$	$3b+c+a/2$	0
101	0	1	0	0	1	$-(2+a+4b+c)$	0	$a+4b+c$
110	0	0	1	0	1	0	$-(2+a+4b+c)$	$a+4b+c$
111	0	0	0	1	0	1	1	-3

F. Function "find the threshold workload"

```
for (j in Nodes2) do
        loop_index2:=0;
        for(k in k_number_of_vehicles) do
                loop_index3 := 0.01;    !estimated lamda_k_alpha
                loop_index2+=1;    ! z
                RHS :=0;    !RHS is loss probability
            while (loop_index3 < (service_rate(j) *loop_index2)) do
                loop_index4:=0; !k
                bottom:=0;
                while (loop_index4 < loop_index2) do    !to get the denominator of erlang B
                    bottom+= (((loop_index3/service_rate(j))^loop_index4) /Factorial (loop_
                    index4));
                    loop_index4 +=1;
                endwhile;

                RHS:= 1-power(2.71828183, -(loop_index3/service_rate(j))) * bottom ;

                if((1-RHS) > Required_a )    then
                    loop_index3+=0.01;
                else
                    lamda_k_alpha(k,j):= loop_index3-0.01;
                    loop_index3:=service_rate(j)*loop_index2+1;
                endif;
            endwhile;
        endfor;
endfor;
```

G. Proof of the equivalence of Model T-LSCPA and LSCPA

Proof. First, we show that the integer variable z_i in Model $LSCPA$ can be rewritten as $\sum_{k=0}^{u_i} kz_{ik}$ in Model $T-LSCPA$ for any station i, where z_{ik} is a binary variable representing whether or not k vehicles are located at Station i.

We consider two cases. In the first case, Station i is activated, that is, $x_i = 1$. Because one and only one number of vehicles are located at Station i,

$$\sum_{k=1}^{u_i} z_{ik} = x_i = 1$$

Suppose that k^S vehicles are located at Station i, it is easy to verify that

$$z_i = k^S = k^S z_{ik^S} = \sum_{k=1}^{u_i} kz_{ik}$$

In the other case, Station i is close, that is, $x_i = 0$. It is followed that $z_{ik} = 0$ for all $k > 0$, and $z_{i0} = 1$. Hence,

$$z_i = 0 = 0 \times z_{i0} + \sum_{k=1}^{u_i} kz_{ik} = \sum_{k=0}^{u_i} kz_{ik}$$

As a summary of these two cases, we can see that the integer variable z_i in Model $LSCPA$ is equivalent to $\sum_{k=0}^{u_i} kz_{ik}$ in Model $T-LSCPA$ for any station i.

Variables x_i and y_{ij} in Model $T-LSCPA$ are kept unchanged as in Model $LSCPA$.

The parameter, w_{ik}, which denotes the maximum workload which can be served by Station i if k vehicles are located there, is predetermined as in Appendix F. This is to ensure that Constraint (4.5) in Model $T - LSCPA$ is equivalent to Constraint (4.2) in Model $LSCPA$. Constraints (4.6) and (4.7), as mentioned above, are proposed to make $\sum_{k=0}^{u_i} kz_{ik}$ in Model $T - LSCPA$ equivalent to z_i in Model $LSCPA$, and also ensure that Constraint (4.3) holds as well. The constraint about variable y_{ij} keeps unchanged.

Therefore, Model $T - LSCPA$ is a linear equivalence of Model $LSCPA$ with the original variable z_i substituted by $\sum_{k=0}^{u_i} kz_{ik}$. □

H. The 55-node problem, shown in Table H.1

Table H.1.: The detailed information of the 55-node problem

Node	x-coordinate	y-coordinate	arrival rate
1	32	31	7.1
2	29	32	6.2
3	27	36	5.6
4	29	29	3.9
5	32	29	3.5
6	26	25	2.1
7	24	33	2
8	30	35	1.9
9	29	27	1.7
10	29	21	1.7
11	33	28	1.6
12	17	53	1.5
13	34	30	1.4
14	25	60	1.2
15	21	28	1.2
16	30	51	1.1
17	19	47	1
18	17	33	1

19	22	40	0.9
20	25	14	0.9
21	29	12	0.9
22	24	48	0.8
23	17	42	0.8
24	6	26	0.8
25	19	21	0.8
26	10	32	0.7
27	34	56	0.6
28	12	47	0.6
29	19	38	0.6
30	27	41	0.6
31	21	35	0.6
32	32	45	0.5
33	27	45	0.5
34	32	38	0.5
35	8	22	0.5
36	15	25	0.5
37	35	16	0.5
38	36	47	0.4
39	46	51	0.4
40	50	40	0.4
41	23	22	0.4

42	27	30	0.4
43	38	39	0.4
44	36	32	0.4
45	32	41	0.3
46	42	36	0.3
47	36	26	0.3
48	15	19	0.3
49	19	14	0.3
50	45	19	0.3
51	27	5	0.3
52	52	24	0.2
53	40	22	0.2
54	40	52	0.2
55	42	42	0.2

I. Modification to address infeasibility

```
First Part
Update Y1(i,j) if assigned station is not closest
Min_index is the index of station which 1. Has more
than 0 ambulance & 2. Is closest to i

Y1(i,j) := Y(i,j);
for (i in Nodes) do
     min_dis:= 500000; // set a large value
     loop_index2:=0;
     min_index :=0;
     for (j in Nodes2) do
          loop_index2+=1;
          if sum[k, Z(j,k)] = 1 then // >0 ambulances
               if Shortest_dist(i,j) < min_dis then
                    min_dis := Shortest_dist(i,j);
                    min_index := loop_index2;
               endif;
          endif;
     endfor;
     loop_index3:=0;
     if min_index > 0 then
          for (j in Nodes2) do
               loop_index3+=1;
               Y1(i,j) := 0;
               if loop_index3 = min_index then
                    Y1(i,j) := 1;
               endif;
          endfor;
     endif;
endfor;


Second part
Copy from BBK1 model
Only changes is that station only receive assigned
demand and empty station's demand is 0
num_of_server_at_j_to_meet_alpha(j) is the final
solution

for (j in Nodes2) do
     lamda_N_j_new(j) := 0;
     for( i in Nodes) do
          if(Y1(i,j) = 1) then
               lamda_N_j_new(j) += arrival_rate(i);
          endif;
```

```
     endfor;
endfor;

// below are all same as BBK1, except demand are new
updated demand (lamda_N_j_new(j))
for(j in Nodes2) do
     loop_index2:=1;
     for(k in k_number_of_vehicles) do
          loop_index:=0;
          first_term_in_Po(j,k):=0;
          while(loop_index<= (loop_index2-1)) do
               first_term_in_Po(j,k) +=
((lamda_N_j_new(j)/service_rate(j))^loop_index/Factori
al(loop_index));
               loop_index+=1;
          endwhile;
     loop_index2+=1;
     endfor;
endfor;


for(j in Nodes2) do
     loop_index:=1;
     for(k in k_number_of_vehicles) do
     Po(j,k) :=0;

          if(lamda_N_j_new(j) >= (loop_index *
service_rate(j))) then
          Po(j,k) := 0;
          loop_index+=1;
          else
          Po(j,k) := ( 1/ ((first_term_in_Po(j,k)) +
(((lamda_N_j_new(j)/service_rate(j))^loop_index)
*(1/Factorial(loop_index))
*((loop_index*service_rate(j))/((loop_index*service_ra
te(j))-lamda_N_j_new(j)))))) ;
          loop_index +=1;
          endif;
     endfor;

endfor;

for(j in Nodes2) do
     loop_index:=1;
     for(k in k_number_of_vehicles) do
```

```
     Prob_matrix(j,k):= 0;
          if(lamda_N_j_new(j) >= (loop_index *
service_rate(j))) then
          Prob_matrix(j,k) := 0;
          loop_index+=1;
          else
     Prob_matrix(j,k):=
((((lamda_N_j_new(j)/service_rate(j))^loop_index)*Po(j
,k))/((Factorial(loop_index))*(1-
(lamda_N_j_new(j)/(loop_index*service_rate(j))))));
     loop_index +=1;
     endif;
     endfor;
endfor;

for(j in Nodes2) do

     for(k in k_number_of_vehicles) do
          P_at_least_one_free(j,k) :=0;

          if(Prob_matrix(j,k) > 0)then


               P_at_least_one_free(j,k) :=(1-
Prob_matrix(j,k) );

          endif;
     endfor;

endfor;

for(j in Nodes2)do
     loop_index:=1;
     for(k in k_number_of_vehicles) do

          num_of_server_at_j_to_meet_alpha(j) :=0;
          if(P_at_least_one_free(j,k) >= Required_a)
then
               num_of_server_at_j_to_meet_alpha(j) :=
loop_index;
          else
          loop_index+=1;
          endif;
     endfor;
endfor;
```