



Optimization of Container Line Networks with Flexible Demands

Plum, Christian Edinger Munk

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Plum, C. E. M. (2013). *Optimization of Container Line Networks with Flexible Demands*. DTU Management Engineering.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Ph.D. thesis

Optimization of Container Line Networks with Flexible Demands

Christian Edinger Munk Plum

April 17, 2013

Dansk titel:

Optimering af container netværk med fleksibel fragtliste

Type: Ph.D.-afhandling

Forfatter: Christian Edinger Munk Plum

ISBN-nr: 978-87-92706-06-5

DTU Management

Department of Management Engineering

Technical University of Denmark

Produktionstorvet, Building 426,

DK-2800 Kgs. Lyngby, Denmark

Phone: +45 45 25 48 00, Fax: +45 45 25 48 05

phd@man.dtu.dk

Maersk Line

Operations - Network Strategy

Esplanaden 50,

DK-1098 Kbh K., Denmark

Phone: +45 33 63 33 63

Grafik: Anders Bülow

Tryk: Schultz Grafisk A/S

April, 2013

Abstract

Liner shipping is at the core of the world's supply chains, with an estimated 36 % of the value of global merchandise trade being shipped in containers. The containers, carried on thousands of container vessels in intricate networks operated by global liner shipping carriers, constitute a very important part of the world economy. Although maritime transport is an environmental friendly transport mode, it is also an industry which emits millions of tonnes of CO₂ yearly. Highlighting the need to control the environmental impact. Container carriers operate in a highly competitive market, where the assets must be deployed in the best way possible to create a healthy business.

To better manage the assets invested in container shipping and to control the use of fossil fuels used by the liner shipping industry, optimization methods for liner shipping is studied in this thesis. The domain is investigated and models supporting decision making is developed, with the aim of reducing costs, bunker consumption and increasing the revenue and service levels of a liner shipping company.

These problems are complex, dealing with millions of containers traveling on hundreds of vessels calling hundreds of ports, all over the world. From an industry view, the complexity of managing these networks has grown tremendously in the last decades, with yearly two-digit growth rates. Developing from a few trades with 10 - 20 vessels deployed, which could be considered independently, to a massively interconnected network spanning the globe.

From a mathematical point of view these problems are among the very hardest, in the class of *NP*-hard problems. And when considering network design problems, these are among the most difficult of *NP*-hard problems, which rarely can be solved to optimality for medium or large instances.

The problems are studied with operations research techniques, which successfully have been used, for planning and logistical problems in other transportation industries, resulting in great improvements.

The thesis follows two main research directions, the first focuses on liner shipping network design, the second on other decision problems faced in liner shipping, that can be studied with operations research methods.

Research in liner shipping network design has been relatively scarce until recently, and the research that has been done, have been distributed, lacked focus and agreement on which aspects where important and relevant to include. To alleviate this a thorough description of the domain of liner shipping is given, explaining the industry in the words of an operations researcher. At the same time a set of benchmark instances, LINER-LIB 2012 is introduced. It is the hope that this description together with the benchmark instances will provide a common ground for research in liner shipping network design, enabling comparisons of methods and easing entry barriers for new researchers in this important field.

This thesis presents three different approaches for liner shipping network design. The first presents a model that allows for the creation of services (loops of vessels, following the same route) connecting to form a liner shipping network. This model is decomposed in a novel manner, where the partial flow of demand and construction of services is done in the same subproblem. This work also introduces an interesting aggregation of demands, which can greatly reduce the problem sizes. A second approach to liner shipping network design models how demand can flow on services, as opposed to flowing directly between ports. This allows for the creation of more complicated networks than previously seen. By solving an important issue faced by previous liner shipping network design methods, allowing a service to call the same ports, multiple times. This model is implemented and run on the LINER-LIB 2012 instances providing results for these. Lastly a model focusing on the design of a single service is considered. The method optimizes the profit, while considering operational and commercial aspects of liner shipping as capacity of vessels and transit time of demand, which is a very important factor for designing actual container services.

The second part of the thesis considers two decision problems faced in liner shipping. These are important as the fierce competition in liner shipping, gives very small margins of profit. Therefore a successful carrier, will need to control the details, and consistently manage the operational

challenges well. Two examples of this is studied: how to purchase bunker fuel considering contracts and how to manage disruptions in the sailing schedule of a vessel.

Bunker fuel is a huge expense for a liner shipping company, and at current market rates it constitutes up to 30 % of a networks operational cost, equalling in billions of dollars for large container carriers. To manage this, carriers will often use contracts for delivery of bunker to ensure supply and achieve a small discount on volume. As these contracts are shared between vessels, it constitutes a shared resource, which must be distributed optimally. A model is formulated, which is decomposed, implemented and run on real world problem instances of 500+ vessels and 500+ contracts. The method allows a global liner carrier to efficiently plan bunker purchases for their vessels, using a large number of bunker contracts to lower costs.

Container vessels operate on tight schedules to meet the customers transit time requirements, reach their port berth slots and catch connections to other vessels. Often disruptions occur to the schedules due to adverse weather, mechanical failures and port delays. These disruptions have a great impact on the service provided to customers and the cost for recovering from them are high. A mixed integer programming model is developed, which can suggest an optimal mitigation for a given disruption. The model considers common disruption scenarios and is run on four real cases, finding optimal solutions in less than 5 seconds. The cases show up to 58 % savings in recovery costs compared to manually realized recovery costs.

This thesis aims at opening up research in the important area of liner shipping network design in a number of ways. It gives a thorough introduction to the domain, presents a number of benchmark instances and proposes several models for liner shipping network design, which highlights important and not previously studied aspects of the problem. This allows further research in the area to use some of this work as building blocks for new methods. Two operational problems faced in liner shipping are considered with good results, showing the breadth of research areas existing in liner shipping.

From an industry point of view, models assisting with design of a single service or small networks have been presented and both the operational models shows promise for implementation in an actual decision support system. These can help overcome some of the complex problems faced in liner shipping, showing that operations research techniques can be applied to real liner shipping problems.

Resumé (Summary in Danish)

Containerlinjefart er grundpillen i verdens forsyningskæder, hvor 36 % af værdien af den globale samhandel anslås afskibet i containere. Containerne, der transporterer gods på tusindvis af containerskibe i komplekse netværk, udgør derfor en vigtig del af verdensøkonomien. Selvom søtransport er en miljøvenlig transportform, er det også en industri, der udleder millioner af tons CO₂ årligt. Derfor kan der være store miljømæssige fordele af selv relativt små reduktioner i containertransportindustriens CO₂-omsætning. Containerrederier opererer samtidig på et stærkt konkurrencepræget marked, hvor skibene skal anvendes på den bedst mulige måde for at skabe en sund forretning. Den praktiske betydning af at optimere på containerskibenes drift er således ganske betragtelig.

Denne afhandling undersøger optimeringsmetoder til linjerederiernes containerdrift med henblik på at forvalte de midler, der investeres i containersøfart bedst muligt. Og samtidigt betragte brugen af fossile brændstoffer, der anvendes af linjeskibsfarten. I afhandlingen undersøges domænet for linjefart, og der udvikles modeller, som understøtter branchens beslutningsprocesser.

Der er tale om ganske komplekse problemer, som vedrører millioner af containeres rejser på hundredvis af skibe, som anløber hundredvis af havne over hele verden. Fra en branchebetragtning er kompleksiteten i styringen af dette netværk vokset voldsomt i de seneste årtier, da der har været årlige tocifrede vækstrater i fragtmængden. Udviklingen siden 1970'erne, er gået fra få ruter med 10 - 20 skibe på hver selvstændig rute, til et massivt sammenhængende netværk, der spænder hele kloden og stiller krav til en kompliceret rekonceptualisering af rutetænkningen, og dermed til planlægningen af containerskibenes anvendelse, med henblik på at reducere omkostningerne og bunkerforbruget, samt at øge linjerederiernes indtægter og serviceniveau.

Fra et matematisk synspunkt er de studerede problemer blandt de hårdeste, tilhørende klassen af *NP*-hårde problemer. Heraf er netværksdesignproblemerne, blandt de absolut vanskeligste *NP*-hårde problemer, og selv små problemer kan vanskeligt løses til optimalitet.

Med henblik på at undersøge problemerne er der i afhandlingen anvendt operationsanalytiske teknikker, som er blevet anvendt til effektivt at løse planlægnings- og logistiske problemer i andre transportindustrier, med store forbedringer i driften som resultat.

Afhandlingen følger to primære forskningsretninger. Den første fokuserer på design af linjefart-netværk. Den anden på andre beslutningsproblemer i linjefart, som kan studeres med operationsanalyse.

Forskning inden for design af linjefartsnetværk har indtil nu været begrænset, og den forskning der er blevet udført, har været spredt med manglede fokus og enighed om, hvilke aspekter der har været vigtige at inddrage. For at afhjælpe dette, indeholder denne afhandling en grundig beskrivelse af domænet for linjefart, og beskriver branchen i operationsanalytiske termer. Samtidig er der udviklet et sæt benchmarkinstanser, LINER-LIB 2012. Det er håbet, at denne domænebeskrivelse sammenholdt med benchmarkinstanserne, vil give et fælles grundlag for fremtidig forskning i design af linjefartsnetværk, der gør det muligt at sammenligne metoder og reducere adgangsbarrierer for nye forskere på dette vigtige område.

I forhold til de forskningsretninger, der fokuserer på linjefartsnetværk undersøges tre forskellige tilgange til design af linjefartsnetværk. Den første model giver mulighed for etablering af *linier* (En gruppe skibe med samme og forudbestemt rute), der mødes i havne, og derved danner et linjefartsnetværk. Denne model bliver dekomponeret på en nyskabende måde, hvor en delvis rutning af containerne og konstruktion af linjer foregår i samme delproblem. Dette arbejde introducerer også en interessant aggregering af container korridorer (containere fra og til samme havn), som i høj grad kan reducere problemets størrelse. Den anden model for design af linjerederiers netværk, modellerer hvordan efterspørgslen kan rutes på linier, i modsætning til at rutes direkte mellem havne. Dette giver mulighed for dannelse af mere komplicerede netværk end set i tidligere modeller. Tidligere netværksdesign modeller har ikke tilladt en linje at kalde de samme havne flere gange. Dette tillades af denne nye model. Modellen er implementeret og testet på LINER-LIB 2012 instanserne. Den tredje model fokuserer på udformningen af en enkelt linie. Metoden optimerer profit, og samtidig tages overvejelser om operationelle og kommercielle aspekter af linjefart, som skibenes kapacitet og containernes transittid i betragtning, hvilket er vigtige faktorer for at

designe container linier.

Der undersøges to operationelle beslutningsproblemer indenfor linjefart. Det første problem vedrører beslutninger om kontrakter ved indkøb af bunkerolie. Det andet problem handler om hvordan forstyrrelser i et skibs sejlplan håndteres bedst muligt. Disse er vigtige, da den hårde konkurrence mellem rederierne giver meget små profitmargener. Derfor er det en forudsætning for overlevelse, at rederier kontrollerer detaljerne og konsekvent styrer de operationelle udfordringer effektivt.

Optimal beslutningstagning om kontrakter ved indkøb af bunkerolie er afgørende idet bunkerbrændstof er en stor udgift for et rederi. Udgiften udgør op mod 30 % af netværkets driftsomkostninger svarende til milliarder af dollars årligt for store containerrederier. For at styre dette, vil rederierne ofte bruge kontrakter om levering af bunkerolie til at sikre forsyningen og opnå rabat på volumen. Da disse kontrakter kan deles mellem skibene, skal fordelingen styres optimalt mellem de forskellige skibe. I afhandlingen formuleres en model for problemet, som dekomponeres, implementeres og testes på virkelige instanser med 500+ skibe og 500+ kontrakter. Metoden giver mulighed for at et globalt rederi effektivt kan planlægge bunkerindkøb for deres skibe, ved at bruge et stort antal bunkerkontrakter for at opnå lavere omkostninger.

Optimal håndtering af forstyrrelser i skibenes sejlplaner er afgørende af flere grunde. For det første fordi en vigtig konkurrenceparameter er imødekomme af kundernes transitkrav. For det andet fordi det er vigtigt at skibene når deres kajtider og forbindelser til andre skibe, idet der opereres med ganske stramme tidsplaner. Der opstår ofte forstyrrelser i tidsplaner på grund af dårligt vejr, mekaniske fejl og havne forsinkelser. Disse forstyrrelser har en stor betydning for den service kunderne oplever, og samtidig er omkostningerne ved at genoprette sejlplanen høje.

Der udvikles en heltalsprogrammeringsmodel, som kan foreslå en optimal genopretningsplan for en given forstyrrelse af sejlplanen. Modellen er kørt på almindeligt forekommende forstyrrelser, og testet på fire faktiske instanser. Optimale løsninger findes på mindre end 5 sekunder. Instanserne viser, at der kan spares op til 58 % i genoprettelsesomkostninger, for at indhente sejlplanen, i forhold til den manuelt valgte genopretnings plan.

Samlet set sigter denne afhandling mod, på en række forskellige måder, at åbne forskningsfeltet indenfor design af linjefartsnetværk. Der gives en grundig introduktion til domænet, præsenteres en række benchmarkinstanser og fremlægges flere modeller for design af linjefartsnetværk, som belyser vigtige aspekter af problemfeltet, som ikke tidligere er studeret. Fremtidig forskning på området kan derfor benytte dette arbejde som byggesten til nye metoder. To operationelle problemer i linjefart er blevet undersøgt med gode resultater, hvilket viser bredden af interessante forskningsområder der findes i linjefart.

Fra et branchesynspunkt, er der blevet præsenteret interessante modeller, som bistår med design af en enkelt linie eller små netværk. Begge de operationelle modeller viser lovende resultater, der åbner for implementering i egentlige beslutningsstøttesystemer i praksis. Disse kan hjælpe med at overvinde nogle af de meget komplekse udfordringer i linjefart, og viser, at operationsanalytiske teknikker kan anvendes til håndtering af reelle problemer, der opstår hos et linjefarts rederi.

Preface

The work presented in this dissertation constitute in part the fulfillment of the requirements for acquiring the degree of *Philosophiae Doctor* at Technical University of Denmark.

The project has been carried out for Maersk Line in cooperation with the Technical University of Denmark (DTU) in a Industrial Ph.D. project co-funded by the Danish Ministry of Science, Innovation and Higher Education. The Ph.D. study was performed at DTU Management Engineering, from January 2010 to April 2013. The *Industrial Ph.D. programme* aims to perform research which is interesting not only academically but also to the industry. Aiming to generate research which can have a practical relevance to industry.

Professor, Ph.D. David Pisinger, DTU Management Engineering, supervised the study, Analytics Manager, Ph.D. Mikkel M. Sigurd, Maersk Line co-supervised the project on behalf of Maersk Line.

The thesis deals with different aspects of optimization arising within liner shipping. The thesis consists of an introduction in Chapter 1, six research papers in Chapters 2, 3, 4, 5, 6 and 7 and a conclusion in Chapter 8. Two of the six papers are published, one appears in conference proceedings and three are currently submitted to peer reviewed journals within the field. All research papers are self contained within their domain, and contain separate biographies. All papers are co-authored.

The Ph.D. thesis consists of four parts. The first introduces the subject and explains the domain of liner shipping in two chapters. The second part presents different approaches to liner shipping network design in three chapters. The third part presents other uses of operations research in liner shipping, with one chapter on disruption management and one on bunker purchasing. The last chapter concludes the dissertation. A conference contribution on a feeder network design problem is given in Appendix, Chapter 10.

Copenhagen, Denmark, April 2013

Christian Edinger Munk Plum

Acknowledgements

This thesis would not have been without the help and support of many talented and spirited people.

First I would like to thank my supervisors David Pisinger and Mikkel Mühlhoff Sigurd. Without David's never ending support and continuous optimism, (*you have infinitely much time left*, until one month before deadline) this work would not have the ambition I hope you will see in it. The project would not have been without Mikkel's support and he has kept the drive to work on projects which could have a real world impact in Maersk Line alive, when things looked impossible. Mikkel's help opening the doors in Maersk Line has been invaluable.

A great thanks to Juan-José Salazar-González for welcoming me to beautiful La Laguna and great guidance on the single string design project. Thanks to Jorge Riera-Ledesma for making me feel at home at ULL and good lunch discussions.

I have been lucky to coauthor articles with other researchers, without their help this project would have been impossible: Jose Fernando Alvarez for his untiring work with our many iterations of the benchmark paper, Peter Neergaard Jensen for his invaluable work on our bunker projects, Guy Desaulniers and Mads Kehlet Jepsen for guidance and help on the path based project, Bo Vaaben and Jakob Dirksen for great discussions and work on our disruption project. Thanks to Line Blander Reinhardt for cooperation on the West Africa project, I am looking forward to see its future! A special thanks to Berit Dangaard Brouer for cooperation in many projects, great discussions, driving ambitions and lifting mood. It has been very encouraging to have a good companion throughout this project.

Thanks to Charlotte Vilhelmsen and Christian Vad Carsten for revisions of papers. Many others at DTU Management have helped me in little and big ways, in model discussions, life and everything in between, making coffee, cake and all other things making this thesis. Thanks to all of you.

A great many people have helped me at Maersk Line, and the actuality and relevance, in the project I own to them. Firstly thanks to Jørgen Harling who had the foresight to sponsor the project. Appreciations to Rolf Nielsen for his support and help keeping focus. Thanks to Niels Rasmussen for his vast knowledge and help keeping the project on the right track. And the many people who have listened to ideas and clarified my vague notions of what shipping is about: Niels Madsen, Jakob Skov, Claus Anker, Steffen Conradsen, Søren Toft and many others. Thanks to my great colleagues Thomas Gerken, Joël Raucq, David Magid Rousøe, Manoj Marne and Wouter Groeneweg for support, encouragements, discussions and help.

Thanks to my family for great support and encouragement, and to Aslan for being super and making sure I got up early. Lastly a great thanks to my wonderful girlfriend Eva Secher Mathiasen for her encouragements, patience and the time to make this possible.

Contents

I	Liner Shipping - Importance and Domain	1
1	Introduction	3
1.1	Impact	5
1.2	Thesis overview	7
2	A base integer programming model and benchmark suite for liner shipping network design	13
2.1	A benchmark suite for research on liner shipping network design problems	14
2.2	Literature review	15
2.3	Contribution	18
2.4	Domain of the liner shipping network design problem	19
2.5	The benchmark suite for LSNDP	27
2.6	Reference model for the LSNDP	33
2.7	Computational results	43
2.8	Conclusion and future work	51
2.9	Appendix - Figures	52
II	Liner Shipping Network Design	59
3	The Single Service Design Problem in Liner Shipping	61
3.1	Introduction	61
3.2	Mathematical Formulation	64
3.3	Computational Results	66
3.4	Conclusion and Further work	69
4	A Path Based Model for a Green Liner Shipping Network Design Problem	73
4.1	Introduction	74
4.2	Service Based Model	78
4.3	Preliminary computational Results	84
4.4	Conclusion	85
5	A Service Flow Model for the Liner Shipping Network Design Problem	89
5.1	Introduction	90
5.2	Literature	91
5.3	Mathematical Formulation	93
5.4	Computational results	99
5.5	Conclusion	101

III	Operational Liner Shipping Problems	105
6	The Vessel Schedule Recovery Problem (VSRP) - a MIP model for handling disruptions in liner shipping	107
6.1	Introduction	108
6.2	The Liner Shipping Business	110
6.3	Literature review	112
6.4	The Vessel Schedule Recovery Problem - (VSRP)	114
6.5	Computational results	120
6.6	Conclusion and future work	124
7	Bunker Purchasing with Contracts	131
7.1	Introduction	131
7.2	Bunker Purchasing with Contracts	134
7.3	Algorithm	137
7.4	Computational Results	138
7.5	Conclusion and Further Work	139
IV	Conclusion	147
8	Conclusion	149
8.1	Contribution	150
8.2	Trends in Liner Shipping Optimization Research	151
8.3	Final Remarks	153
V	Appendix	155
9	Bachelor and Master Thesis Projects	157
10	Models for Feeder Vessel Scheduling with Split Deliveries and Time Windows	159
10.1	Mathematical Model Of FVSP	160
10.2	Outlook	161

Part I

Liner Shipping - Importance and Domain

Chapter 1

Introduction

As you read this, thousands of container ships are carrying millions of containers to and from every corner of the earth. These gigantic networks carry nearly any imaginable commodity and product engineered by mankind, from waste paper enroute for recycling, to high tech electronic equipment and everything in between. Huge vessels which enable a core activity of our society, *trade*, by allowing for large scale, low unit cost transport between any part of the world.

Operating such networks are very expensive, with the largest carriers having yearly operating costs of two digit billions of dollars. Also in environmental terms the maritime industry is expensive, and has a huge impact with 2.7 % of world CO2 emissions, according to Psaraftis and Kontovas [20].

The container shipping industry has grown tremendously in the last couple of decades, with yearly growth rates of 8-12 %. This growth has mainly been supported by ever growing vessels, where in 1975 state of the art container vessels as the *Adrian Maersk*-Class had a capacity of 1.200 twenty foot containers (TEU), today the *EEE*-class currently in production will have a capacity of 18.000 TEU.

A liner shipping network, consists of a number of *services* which are cyclic roundtrips calling a fixed number of ports, sailed by a number of vessels. These are much like common bus routes and are served at a fixed frequency, usually weekly, by a number of similar vessels. At ports, the vessels can tranship containers between each other, allowing containers to reach any part of the world, served by the network.

The growing scale of the networks, have made them increasingly hard to manage and control, as the number of interactions between vessels, ports and containers have grown manifold making it difficult to foresee ripple effects of changes to networks and routings, and to optimize on these.

This is the scope of this thesis: how can decision support tools be used to optimize the tactical planning and execution of a liner shipping network on a planning and operational level.

Solution approach Operations Research (*OR*) techniques will be used to study these problems. OR has been used successfully for many types of logistical, planning and decision problems, over the past 70 years and with the great increase in computational power, it has been used and studied for many applications areas.

The structures governing liner shipping are well known, Stopford [22]. I.e what the charter cost of a vessel per day is, what the container capacity of a vessel is, how the fuel consumption relate to the speed, etc. All these are based on technical or economical attributes, which for a studied time period can be defined fairly well. We know what the ships, containers and ports are capable off, and we can describe their interactions in detail. The problem is how we deploy these assets optimally, with the vast number of configurations they allow. A container shipping network consisting of hundreds of vessels, hundreds of ports and millions of containers, can be combined in exponentially many ways. This is where operations research techniques are a good tool, for solving a problem with known underlying structure, which allows many configurations. To face this combinatorial challenge OR can find good solutions in an enormous solution space.

In many logistical industries as airline [24], trucking [23], and rail [10] OR techniques are used to achieve great results in improvements of service levels and effectiveness. But the great breakthrough has yet to materialize for liner shipping, as Christiansen et al. [4] writes in a 2004 review on the field:

We notice that there has been little research on liner shipping and strategic shipping (e.g., fleet size) problems during the last decade. This may be surprising, especially when we consider the increase in container traffic and the large number of mergers in the container shipping industry.

This has partly been due to shipping being a conservative industry, keeping their cards tight, but also because proposed optimization solutions [1, 21] have been unable to solve the scales of the problems seen in real world settings and an inability to incorporate all relevant business rules.

These impediments are showing signs of change. Lately Maersk Line has started to use optimization to increase effectiveness in several of their businesses, in various settings such as bunker purchase optimization [14] and optimization of container flows. Other container lines have also started supporting research in liner shipping or are using optimization methods such as APL, OOCL and CSAV [6]. The process of designing and maintaining a liner shipping network remains a very manual task, relying on the knowledge of experienced planners. However, as the word of the good results of the before mentioned projects spreads, the knowledge and interest in optimization techniques increases - as does employees willingness to invest time and data to develop them.

This facilitates that the time is mature for a larger breakthrough of optimization in container shipping. The increased transparency of the industry allows us to investigate the posed problems in detail and analyze, which business rules are critical for success. This should be used in combination with the best optimization techniques known in order to develop specialized methods, that can solve the container shipping problems in practice.

Since the review by Christiansen et al. [4] the research interest in liner shipping problems has picked up, and a recent review of Meng et al. [12] lists 18 articles published on *containership routing and scheduling in liner shipping* until 2006 and 52 articles published from 2007 and forward, which illustrates the great increase in interest in this research field. Or as stated in the recent review article by Christiansen et al. [5]:

Research on ship routing and scheduling has blossomed during the last decade. Comparing to the former decade its volume has more than doubled, and the same is true for the variety of research outlets. The research seems to be catching up with the increasing world fleet and trade.

Liner Shipping Network Design The main research focus of this project is developing methods to design liner shipping networks. Additionally other problems occurring in liner shipping have been studied, where it has seemed obvious that OR methods would be able to add value.

As mentioned, a global container shipping network is costly to operate. Therefore, even a small improvement of the network's utilization, costs, service levels, etc. would have very large impact. The cost structure of the network can be very volatile and by developing models that quickly can investigate an increased cost or reduced demand, the network can quickly be modified to adapt for these changed market conditions. Hence if we can develop methods that can be used to create liner shipping networks, which are more efficient, cheaper to run and emits less CO₂, a liner company can respond to market changes faster. The focus of choosing the right demand for the right network, will give a company a competitive advantage by having methods to create networks that caters for the demand according to its revenue.

Uncertainty As the studied problems are real world problems, most of the data are to some degree uncertain. For instance the consumption of bunker will depend on vessel draft, last hull scraping, weather etc. Some of these factors could be modeled at the cost of great complexity with little added accuracy, while others would be impossible to forecast. Furthermore, models considering data uncertainty are usually harder to solve than deterministic ones. Hence the developed models, in this thesis, are all deterministic. For the network design models it is already hard to scale to real world instance sizes, so there is little reason endeavoring to construct a stochastic

model. For the other methods it makes sense to construct a deterministic model before starting on a stochastic version.

Industrial Ph.D. The role of an Industrial Ph.D is to bridge two worlds, the theoretical world of academia and the practical of the industry, and in the spanning of these construct some new invention, which can add value to the industry by leveraging on the state of the art methods used in academia. I have found this role very challenging, and very giving. The unique position of having full access to the industry, through data and access to encouraging industry stakeholders, with detailed knowledge of operational problems and their dynamics have given a unique opportunity to study liner shipping problems. Using actual data and considering the triggers considered decisive and difficult by the problem owners in details. The challenge of making the research applied has been omnipresent in this project. Discussing problems faced in the company with the problems owners and evaluating whether they were suitable for consideration with OR methods. Once a problem was selected the continuous challenge was: what is *necessary* to capture in the model to correctly represent its complexity versus what is *possible* to solve using OR methods. This trade off is not always simple to find, as a stakeholder may mention constraints, which very seldom occur in practice, and which can be dealt with at a small cost, and thus can be relaxed. On the other hand, a stakeholder may not mention other characteristics, which is core to understanding and dealing with the problem, due to him/her seeing it as out of problem scope.

1.1 Impact

The contribution of this thesis can be seen as the individual contributions in each research paper, which is discussed in detail in the corresponding papers. Highlights are:

- The liner shipping network design benchmark data set of LINER-LIB 2012. This detailed real life data set allows OR researchers to investigate the LSNDP and compare developed methods. A detailed domain description scopes liner shipping network design in terms of OR.
- A basic model for the LSNDP, which captures its core structures. An algorithm has been constructed and implemented, reporting the first results for the LINER-LIB 2012 instances.
- A path based model for a liner shipping problem is presented together with a novel aggregation scheme for the demands, which greatly reduces the number of demands that must be considered.
- A model and algorithm is developed for constructing a single liner shipping service, which can solve instances of real world size, of up to 25 ports. While considering vessel capacity and demand path duration limits. The path durations limits are a key commercial factor in designing services to guarantee customer service levels.
- A novel service flow formulation for a liner shipping network design problem is presented, which allows for multiple interconnected services with any number of recurrent port calls to a port, a problem not solved by other liner shipping models. The implementation finds solutions for two of the LINER-LIB 2012 instances.
- A model and experimental results are presented for the *Vessel Schedule Recovery Problem*, which can be used for disruption management in liner shipping. Savings of up to 58 % are reported compared with actual chosen recoveries.
- A model, algorithm and implementation considering the problem of purchasing bunkers for a fleet of scheduled vessels, while considering a number of bunker contracts, is presented. Managing this huge cost effectively seen over a fleet of vessels, can have significant economical impact.

On a broader level this Ph.D. project has aimed at opening the area of liner shipping for research by operational research techniques and reversely developing methods that are realistic enough to be implementable by a liner shipping company. From this point of view a number of promising results have been reached. The LINER-LIB 2012 benchmark suite makes it possible for new researchers with skills in OR, to enter the domain of liner shipping, as it provides a thorough description of the field, its drivers and constraints. At the same time the LINER-LIB 2012 data provides a detailed data set, allowing a researcher to start directly with this. This data can also be used for other research projects in liner shipping, as it gives, costs, distances and demands, which are relevant in other settings. These data have already been used by researchers from a number of countries around the world, for network design and other liner shipping projects. The developed methods for the Liner Shipping Network Design Problem (LSNDP) still do not scale well and lack too many commercial and operational constraints, to be of value for full scale network design. However, with the initialization of the *Competitive Liner Shipping Network Design* project ¹ and the use of LINER-LIB 2012, in projects around the world, it has laid an important foundation for something, which can be used by the world's liner shipping companies in the future to optimize their networks. On the smaller scale the Single String Design project of Chapter 3 scales well and could be used to design real world services with some adaptation.

Two other projects related to operational problems met in liner shipping has been studied, one focusing on recovery of a disrupted vessel schedule and another on bunker purchasing with contracts. Papers have been written on both of these works, the first published and the second submitted, and thus have helped mature the OR field with respect to liner shipping. Both have been run experimentally on real problem instances, scaling well and showing good results. They are being evaluated by Maersk Line / Maersk Oil Trading, for their value in a full implementation and as such have shown promise to be successful in adding value, both for the research community but also for the business. As stated in Pisinger et al. [13] by Director of Maersk Line Situation Room, Steffen Conradsen:

The disruption management study has identified how complex the contingency handling process is when something unforeseen happens in the Maersk Line network. Today we are dependant on manual evaluation based on experience and time available to take the right decision. With a disruption tool several options can be identified, downstream consequences assessed and the most efficient solution be presented. This will have a large potential upside and also improve our reaction time to changes which inevitably will happen in our global network. Participating in the project has made it evident to me that we need to identify and develop the right tools for disruption management to reduce our operational expenses.

In effect one of the papers of this thesis has given a foundation for further research in LSNDP problems, and a further three have developed methods on which sound business cases can be built.

Problem difficulty When investigating liner shipping network design problems, it is natural to investigate related and well-studied routing problems as the Traveling Salesman Problem (TSP) and especially the Vehicle Routing Problem (VRP), which both can be solved efficiently for large problems instances. A conclusion of this thesis project, is that the liner shipping network design problem is considerably harder to solve than VRP type problems, and that we should not hope to construct optimal algorithms, in the foreseeable future, that can solve instance sizes as seen in VRP.

This can be motivated by highlighting a number of structural differences, which adds to the complexity:

- Non simple cycles are allowed in services for LSNDP. As seen in Chapter 5, this requires a considerable number of additional integer variables to model, due to the increased complexity.

¹A recently started research project at DTU Management, in cooperation with Maersk Line and funded by the Danish Maritime Fund.

- There is no depot in LSNDP. Knowing where we must start and end a route or service, greatly reduces complexity and allows for efficient algorithms.
- Transshipment of demand, delinks vessel/vehicle path and demand/commodity path, requiring these to be solved separately, while coordinating their paths.
- No precedence relation, e.g. without a depot, a service under construction is allowed to unload a demand, before it is loaded, as the path is cyclic.
- LSNDP allows for split pickup and delivery, increasing the solution space.
- Liner shipping networks are large, with full instances having 10000+ commodities and 100+ services.

As these factors give the LSNDP a very unconstrained solution space compared to VRP and due to the increased complexity solving medium or large instances to optimality seems unlikely on a shorter horizon.

LSNDP formulations A problem of the research in liner shipping network design problems, as stated in Brouer et al. [2], is that as many models exists for the problem, as there are published articles. Unfortunately this thesis adds to that tradition with different models presented in Chapter 2, Chapter 3, Chapter 4 and Chapter 5. There are several reasons for this: Firstly, the presented formulations are still not catching the core aspects in a compact formulation, e.g. the model of Chapter 2 / Brouer et al. [2] having an exponential number of rows and columns. Secondly, it is a very complex real world problem, and even though the models presented in this thesis, catch much of the complexity, there is still a long way to go before relevant constraints as path duration limits, draft constraints, empty repositing, etc. is all represented in one grand model. Lastly, it has been a goal to construct models, which allow design of algorithms, that can scale to larger instance sizes. Finding the right tradeoff between complexity and scaling has been a recurrent struggle in the project, which shows for developing new models, which allows for more efficient algorithms. Still the model of Brouer et al. [2] catch the absolute core of the problem and a solution to the LINER-LIB 2012 instances should have a close resemblance of what could be realistic in practice.

1.2 Thesis overview

This thesis is divided in three major parts. Part I motivates for the relevance of the problem and introduces the domain. The second part, Part II investigates how the Liner Shipping Network Design Problem can be modeled and the third part, Part III studies other operational problems met in Liner Shipping. Afterwards the thesis is concluded, and finally an appendix is given with other relevant material.

Part I: Liner Shipping - Importance and Domain consists of the introduction to the thesis, which is fairly compact, as the introduction and explanation of the domain and main problem, follows in Chapter 2 *A base integer programming model and benchmark suite for liner shipping network design*. This paper explains operational, commercial and other rules governing liner shipping, with the aim of modeling problems in this domain as OR problems. A basic model for LSNDP is formulated describing core problem characteristics as cost structure and service roundtrips. A number of benchmark instances (LINER-LIB 2012) in different sizes are presented, that are based on real world data. A heuristic column generation based algorithm is proposed and implemented, and the first results are reported for the benchmark instances. We hope that the domain description and benchmark instances LINER-LIB 2012 will facilitate and encourage much more research in LSNDP. The work has been presented as follows:

- Published in DTU Mananagement Technical Report number 19 2011 (Løfstedt, Alvarez, Plum, Pisinger, and Sigurd [11]).

- Accepted for publication in Transportation Science (Brouer, Alvarez, Plum, Pisinger, and Sigurd [2]).

Part II: Liner Shipping Network Design presents three papers all concerning liner shipping network design problems. The first considers the design of a single service, while the two latter deals with the design of several services.

The paper in Chapter 3, *The Single Service Design Problem in Liner Shipping* considers how to optimally construct a single liner shipping service. A number of ports that must be called and demands, with a maximal path duration limit, which may be carried on a service with a fixed capacity is given. Arc and Path flow models are presented and a *Branch-and-Cut-and-Price* algorithm is devised and implemented. Instances of up to 25 ports can be solved which is very promising as real world services seldom call more than 20 ports. The work has been presented as follows:

- Extended Abstract accepted for Proceedings of the International Multi Conference of Engineers and Computer Scientists, 2012, as *The Multi-commodity One-to-one Pickup-and-delivery Traveling Salesman Problem with Path Duration Limits* (Plum, Pisinger, Salazar-González, and Sigurd [15]).
- Submitted for Computers and Operations Research (Plum, Pisinger, Salazar-González, and Sigurd [18]).

Chapter 4, *A Path Based Model for a Green Liner Shipping Network Design Problem*. Presents a novel model for a liner shipping network design problem. An aim of the work is to alleviate a problem met in other formulations of LSNDP: that cost and revenue decomposes to master, respectively sub-problem, giving slow convergence. Cost and revenue are decomposed into the same subproblem, generating partial routings and services, which are combined in the master. The paper also presents a novel aggregation of demands, which greatly reduces the number of commodities to be considered. The work has been presented as follows:

- Extended abstract accepted for Proceedings of the International Multi Conference of Engineers and Computer Scientists, 2011 (Jepsen, Løfstedt, Plum, Pisinger, and Sigurd [8]).
- Published in DTU Management Technical Report number 10 2012 (Jepsen, D., Plum, Desaulniers, Pisinger, and Sigurd [9]).

Chapter 5, *A Service Flow Model for the Liner Shipping Network Design Problem*. Presents a novel modeling approach for a Liner Shipping Network Design problem. Instead of flowing the commodities using variables following the arcs between ports, the commodities are flowed on arcs to and from the services. This gives some nice properties of the model, in particular the formulation is able to model any number of portcalls to the same ports, which is often encountered in liner shipping services, for instance *butterfly* loops. The model has been implemented and results are presented for two of the instances of LINER-LIB 2012. The work has been presented as follows:

- Submitted for special issue of European Journal of Operational Research on *Maritime Logistics* (Plum, Pisinger, and Sigurd [16]).

Part III: Operational Liner Shipping Problems considers other problems encountered in liner shipping, that can be considered with OR methods. The paper in Chapter 6, *The Vessel Schedule Recovery Problem (VSRP)-a MIP model for handling disruptions in liner shipping* presents a model, that aims to mitigate the effects of a vessel encountering a disruption on its schedule. A model is presented, implemented and tested on four real world cases. The objective considers several factors as fuel costs, number of containers who misses a connections, etc. These are weighted differently, in different runs, to generate solutions with varying characteristics. In all cases the model finds similar or better solutions and is able to solve the model in seconds, allowing for use in real operational settings. The work has been presented as follows:

- Published in European Journal of Operational Research (Brouer, Dirksen, Pisinger, Plum, and Vaaben [3]).

Chapter 7, *Bunker Purchasing with Contracts*. The paper considers how a liner shipping company can purchase fuel (bunker) for a fleet of scheduled vessels, considering operational constraints while respecting a number of contracts for bunker across the world. These contracts are a shared resource between the vessels, which must be used optimally. A model is presented and due to the size of real world instances it is decomposed, giving a subproblem for each vessel, which generates columns for a master problem considering the contract limits. This is implemented in a column generation algorithm. The algorithm is run on very large real world data and results are presented, for instances which can not be solved by a MIP model. The work has been presented as follows:

- Extended abstract accepted for the 3rd International Conference on Computational Logistics, ICCL'12 (Farina, Jensen, Plum, and Pisinger [7]).
- Submitted for Maritime Economics and Logistics (Plum, Jensen, and Pisinger [17]).

Part V: Appendix Chapter 10, *Feeder vessel scheduling with split deliveries and time windows* considers the design of a feeder vessel network originating in a single hub port, which caters for a number of feeder ports. The model considers special conditions met in West Africa, requiring the consideration of draft limits when calling a port. Ports serve vessels *First come - First Serve*, as opposed to ports with fixed berth slots. No implementation is given. The work has been presented as follows:

- Extended abstract accepted for the conference TRISTAN 8, 2013 (Plum, Reinhardt, Pisinger, and Sigurd [19]).

Bibliography

- [1] Richa Agarwal and Ozlem Ergun. Ship scheduling and network design for cargo routing in liner shipping. *Transportation Science*, 42(2):175–196, 2008.
- [2] Berit D. Brouer, J. Fernando Alvarez, Christian E. M. Plum, David Pisinger, and Mikkel M. Sigurd. A base integer programming model and benchmark suite for liner shipping network design. *Transportation Science*, 2013. Accepted by Transportation Science.
- [3] Berit D. Brouer, Jakob Dirksen, David Pisinger, Christian E.M. Plum, and Bo Vaaben. The vessel schedule recovery problem (vsrp) - a mip model for handling disruptions in liner shipping. *European Journal of Operational Research*, 224(2):362 – 374, 2013.
- [4] Marielle Christiansen, Kjetil Fagerholt, and David Ronen. Ship routing and scheduling: status and perspectives. *Transportation Science*, 38(1):1–18, 2004.
- [5] Marielle Christiansen, Kjetil Fagerholt, Bjørn Nygreen, and David Ronen. Ship routing and scheduling in the new millennium. *European Journal of Operational Research*, 2012.
- [6] Rafael Epstein, Andres Neely, Andres Weintraub, Fernando Valenzuela, Sergio Hurtado, Guillermo Gonzalez, Alex Beiza, Mauricio Naveas, Florencio Infante, Fernando Alarcon, Gustavo Angulo, Cristian Berner, Jaime Catalan, Cristian Gonzalez, and Daniel Yung. A strategic empty container logistics optimization in a major shipping company. *Interfaces*, 42:5–16, 2012.
- [7] Federico Farina, Peter N. Jensen, Christian E. M. Plum, and David Pisinger. Bunker purchasing with contracts. 2012. Presented at <http://iwi.econ.uni-hamburg.de/iccl2012/index.htm>.
- [8] Mads K. Jepsen, Berit Løfstedt, Christian E. M. Plum, David Pisinger, and Mikkel M. Sigurd. A path based model for a green liner shipping network design problem. In *Proceedings of the IAENG, ICOR 2011*, 2011.
- [9] Mads K. Jepsen, Berit D., Christian E. M. Plum, Guy Desaulniers, David Pisinger, and Mikkel M. Sigurd. A path based model for a green liner shipping network design. Technical Report 10.2012, Technical University of Denmark, Management Engineering, November 2012.
- [10] Leo Kroon et al. The new dutch timetable: The or revolution. *Interfaces*, 39:6–17, 2009.
- [11] Berit Løfstedt, J. Fernando Alvarez, Christian E. M. Plum, David Pisinger, and Mikkel M. Sigurd. An integer programming model and benchmark suite for liner shipping network design. Technical Report 19, DTU, Technical University of Denmark, 2010.
- [12] Qiang Meng, Shuaian Wang, Henrik Andersson, and Kristian Thun. Containership routing and scheduling in liner shipping: Overview and future research directions. *Transportation Science*, 2013. To appear in.
- [13] David Pisinger, Rune M. Jensen, and Mikkel M. Sigurd. Enerplan - research achievements. <http://www.maersk.com/Innovation/LeadingthroughInnovation/ENERPLAN/Documents/success.pdf>, April 2012.
- [14] Christian E. M. Plum and Peter Neergaard Jensen. Minimization of bunker costs. Master’s thesis, 2007.
- [15] Christian E. M. Plum, David Pisinger, Juan-José Salazar-González, and Mikkel M. Sigurd. The multi-commodity one-to-one pickup-and-delivery traveling salesman problem with path duration limits. In *Proceedings of the International MultiConference of Engineers and Scientists 2012 Vol II*, 2012.

-
- [16] Christian E. M. Plum, David Pisinger, and Mikkel M. Sigurd. A service flow model for the liner shipping network design problem. Submitted for special issue of European Journal of Operational Research on *Maritime Logistics*, 2012.
- [17] Christian E. M. Plum, Peter N. Jensen, and David Pisinger. Bunker purchasing with contracts. Submitted for *Maritime Economics and Logistics*, 2013.
- [18] Christian E. M. Plum, David Pisinger, Juan-José Salazar-González, and Mikkel M. Sigurd. The single service design problem in liner shipping. Submitted for *Transportation Research Part B: Methodological*, 2013.
- [19] Christian E. M. Plum, Line B. Reinhardt, David Pisinger, and Mikkel M. Sigurd. Feeder vessel scheduling with split deliveries and time windows. Accepted for the conference TRISTAN 8, 2013.
- [20] Harilaos N. Psaraftis and Christos A. Kontovas. Co2 emission statistics for the world commercial fleet. *WMU Journal of Maritime Affairs*, 8(1):1–25, 2009.
- [21] Koichi Shintani, Akio Imai, Etsuko Nishimura, and Stratos Papadimitriou. The container shipping network design problem with empty container repositioning. *Transportation Research Part E: Logistics and Transportation Review*, 43(1):39–59, 2007.
- [22] Martin Stopford. *Maritime Economics*. Routledge, third edition, 2009.
- [23] Candace A. Yano, Thomas J. Chan, Lori K. Richter, Theodore Cutler, Katta G. Murty, and David McGettigan. Vehicle routing at quality stores. *Interfaces*, 17-2:52–63, 1987.
- [24] Gang Yu, Michael Argüello, Gao Song, Sandra M. McCowan, and Anna White. A new era for crew recovery at continental airlines. *Interfaces*, 33:5–22, 2003.

Chapter 2

A base integer programming model and benchmark suite for liner shipping network design

Berit D. Brouer* J. Fernando Alvarez[†] David Pisinger* Christian E. M. Plum*, [‡] Mikkel M. Sigurd[‡]

*Department of Management Engineering, Technical University of Denmark,
Produktionstorvet, Building 426, DK-2800 Kgs. Lyngby, Denmark
blof@man.dtu.dk, pisinger@man.dtu.dk, cemp@man.dtu.dk

[†]Business Optimization, Uptimize!
Gøteborggata 9, Oslo N-0566, Norway
jfa2@cornell.edu

[‡]Maersk Line
Esplanaden 50, DK-1098 København K.
mikkel.muhldorff.sigurd@maersk.com

1

Abstract The liner shipping network design problem is to create a set of non-simple cyclic sailing routes for a designated fleet of container vessels, which jointly transports multiple commodities. The objective is to maximize the revenue of cargo transport, while minimizing the costs of operation. The potential for making cost effective and energy efficient liner shipping networks using Operations Research (OR) is huge and neglected. The implementation of logistic planning tools based upon OR has enhanced performance of airlines, railways and general transportation companies, but within the field of liner shipping applications of OR are scarce. We believe that access to domain knowledge and data is a barrier for researchers to approach the important liner shipping network design problem. The purpose of the benchmark suite and the paper at hand is to provide easy access to the domain and the data sources of liner shipping for OR researchers in general. We describe and analyze the liner shipping domain applied to network design and present a rich integer programming model based on services, which constitute the fixed schedule of a liner shipping company. We prove the liner shipping network design problem to be strongly NP-hard. A benchmark suite of data instances to reflect the business structure of a global liner shipping network is presented. The design of the benchmark suite is discussed in relation to industry standards, business rules and mathematical programming. The data is based on real life data from the largest global liner shipping company, Maersk Line, and supplemented by data from several industry and public stakeholders.

¹Accepted for *Transportation Science* (2013)

Computational results yielding the first best known solutions for 6 of the 7 benchmark instances is provided using a heuristic combining tabu search and heuristic column generation.

Keywords: liner shipping, mathematical programming, network design

2.1 A benchmark suite for research on liner shipping network design problems

Operations research (OR) is widely used within the transportation sector to provide a cost efficient and competitive organization. However, application of OR within containerized liner shipping is scarce (Christiansen et al. [13]). The potential impact of OR on this billion dollar industry is enormous especially given the large concentration of players in the business. Maritime shipping produces an estimated 2.7% of the world's CO_2 emissions, of which 25% is attributable to container ships alone [16]. An energy efficient liner shipping network is becoming increasingly important to all stakeholders. OR can help in designing effective and energy efficient liner shipping networks to mitigate the carbon footprint of the liner shipping industry. We believe that the lack of OR within liner shipping is partly due to barriers for new researchers to engage in the liner shipping research community. Constructing mathematical models and creating data for computational results requires profound knowledge of the domain and data sources. The benchmark suite presented in this paper aims to make liner shipping network design problems (hereafter LSNDP) approachable for the research community in general. We wish to create a platform where methods can be compared on a set of known data instances. By disseminating our knowledge of the liner shipping domain into real world network data instances for mathematical programming, we hope to diminish this formidable hurdle. Our hope is to enable research on the field of liner shipping network design to develop like the Vehicle Routing Problems (VRP) have through the benchmark instances of Solomon [48]. The benchmark suite is seen as the root of a tree where new branches will appear as our ability to solve more complex interpretations of the liner shipping problem grows.

The initial model of the diverse field of VRP models was the Capacitated Vehicle Routing Problem (CVRP). We present a reference model for the LSNDP, which is an extended version of the model in Alvarez [5]. The reference model may be used for comparing various methods applied to the problem and serve as an inspiration to future model development. Our goal is to enable the benchmark suite to support future model and method development, and hence the data of the benchmark suite encompasses several attributes such as transit times not applicable to the reference model. In Section 2.2 relevant literature within the field of liner shipping economics will underpin the importance of various costs and restrictions within network design. A review of the literature within OR and the various models presented is provided. We review methods for solving liner shipping network design related problems and the computational results in the literature to illustrate the computational hardness of the problem. Section 2.4 provides an introduction to the liner shipping network design domain. The domain discussion is complemented by the strategic business and domain knowledge of one of the major global operators within liner shipping. We discuss the data provided in the benchmark suite in relation to the data needed in current mathematical models and the open research questions we try to support. In Section 2.5, we discuss the data objects and data generation. The goal for the instances generated is to capture real life cost structures, trade imbalances, market shares and scale for a global liner shipping company. We discuss the diversity of instances needed to qualify and challenge development of both exact and heuristic methods for solving the LSNDP. Finally, we solve the generated instances by using a new tabu search heuristic based on the framework by Alvarez [5]. The algorithm has been extended to handle more complex routes (so-called butterfly routes) and to ensure a weekly or bi-weekly service frequency. Moreover, an improved MIP neighborhood is used to generate new candidate routes. Computational results are presented in Section 2.6 to qualify the benchmark suite with experimental results and establish the currently best known solutions for the liner shipping network design problem and hopefully spur competition and interest into our field of research. Section 2.8 contains conclusions and directions for further research into liner shipping

network design problems. Sections 2.4 and 2.5 are a shortened version of the technical report [32]. The reader is referred to this report for additional information.

2.2 Literature review

The two papers, Christiansen et al. [13] and Christiansen et al. [14] provide excellent surveys of OR in maritime transportation. They point to the limited research of planning problems in liner shipping in contrast with the size and the possible impact of optimization within the industry. There are a number of contributions on the individual economic and structural factors in designing liner shipping networks, but no comprehensive description of the general liner shipping domain related to network design has been published to date. One contribution of the present paper is to give an overview of these factors related to the LSNDP. In the first part, we give an overview of contributions related to some of the factors affecting network design such as *network configuration*, *bunker price*, *transit time*, *competitive position*, *repositioning of empty containers*, *frequency*, *port call sequence*, *schedule*, and *fleet deployment*. In the second part we review contributions on the LSNDP.

The *network configuration* of carriers is explored in Notteboom [39] concluding that a global network will not have a pure hub-and-spoke structure or a pure multi-port structure. The economy of deploying super panamax vessels on either a multi-port or hub-and-spoke network structure is investigated in a case study by Imai et al. [27]. They conclude that the multi-port structure is superior for the Asia-North America and Asia-Europe trades, whereas the hub-and-spoke structure is advantageous in the European trades.

The impact of *bunker price* on the network configuration of liner shipping companies has been explored by Stopford [49], Notteboom and Vernimmen [41] and Cariou [11]. Seen from the carriers' point of view, reducing speed and making effective use of the capacity deployed in the network is economically attractive as the general analysis of vessel costs by Stopford [49] reveals bunker fuel as the dominant cost in operating a liner shipping network. According to Notteboom and Vernimmen [41] managing the bunker consumption in the network gives the carriers strong incentive to reduce speed, deploy additional vessels to services, and increase buffer time in the schedule to avoid having to increase speed to accommodate for delays and port congestion. The incentive to reduce speed depends on the actual bunker price because the additional vessels deployed to maintain the frequency come at an increased capital cost. Cariou [11] argues that a bunker price exceeding 350 USD per ton will ensure the sustainability of "slow steaming", which means sailing at a reduced average speed. Cariou [11] calculates the CO_2 reductions from slow steaming from 2008-2010 to 11% attesting that slow steaming is a very effective way of reducing the carbon footprint of the liner shipping industry. Recently, Wang and Meng [54] investigate sailing speed optimization for each individual port-to-port voyage on a liner shipping service to reduce bunker consumption using an outer-approximation method.

In liner shipping there is an inherent trade off between reducing bunker consumption through speed reduction and achieving competitive *transit time* for cargo. Notteboom [40] explores the time factor in liner shipping network designs related to transit time of a cargo routing and schedule reliability. An important analysis on the relation between the number of port calls on a service and competitive transit times is reported. The conclusion is that the time spent at ports is very significant and hence the number of port calls is decisive for the transit time of direct connections at the end points of a service. Wang and Meng [53] evaluate the schedule design along with container routing for a fixed network with predefined paths for container shipments in order to minimize transshipment cost and transit time.

Gelareh et al. [24] explore the possible *competitive positions* of a carrier in a market with an incumbent carrier. The relation between time and cost on the market share is modeled and investigated. Gelareh et al. [24] underpin transit times and level of service as important factors in the construction of a liner shipping network design. Wang and Meng [55] propose a tactical model for schedule design capturing the trade off between speed optimization and transit time levels taking into account time uncertainty at sea and in the port times. We believe there will be

a surge in sailing speed optimization taking level of service and schedule reliability into account where the liner shipping service network has already been fixed to a large degree. We are not aware of published models of the LSN DP incorporating distinct cargo transit times or level of service, but we believe that future research on the LSN DP will incorporate these factors.

Another important factor in a liner shipping network is *repositioning of empty containers* due to trade imbalances. Shintani et al. [47] design a model for a single service of a carrier. The experimental results indicate that empty repositioning is significant for the port calling sequence and the cargo handling costs incurred. Dong and Song [18] investigate the proportion of empty repositioning given current global container trade using the existing global liner shipping network. Their results conclude that 27% of all container traffic is empty repositioning. Brouer et al. [10] show that joint optimization of demanded cargo and empty repositioning in a fixed network is viable. Meng and Wang [35] incorporate empty repositioning and port productivity into the evaluation and selection of candidate shipping routes. Empty repositioning is so far not incorporated into the cargo allocation in models where the set of candidate routes are not fixed as in the LSN DP.

Liner shipping generally adheres to a schedule based on *weekly or bi-weekly frequency of port calls* and deploying vessels of similar characteristics to each service [49, 39, 40]. In Alvarez [6] the cost of the practice of a weekly frequency is questioned, and level of service expressions are formulated in order to investigate the impact of maintaining a weekly frequency as opposed to another fixed frequency. Several contributions on the optimal *port calling sequence* may be found for a single route, e.g., Lu [33], Chu et al. [15], Shintani et al. [47], and Hsu and Hsieh [26].

Scheduling of liner shipping services is a related optimization problem. An example of scheduling with fixed port call sequence may be found in Yan et al. [57] as well as the more recent papers of Wang and Meng [53, 55], whereas Agarwal and Ergun [1] decides scheduling to a certain weekday and routing simultaneously.

The fleet deployment problem (FDP) is closely related to LSN DP as fleet deployment is often implicitly considered in the network cost. FDP assumes that the service in terms of actual vessel voyages is fixed and hence that the routing is already decided upon. FDP decides on how to assign the carrier's own vessels to the actual vessel voyages along with the options of chartering in (leasing) vessels or chartering out (forward leasing) own vessels to minimize the cost of maintaining the schedule. Fleet deployment is generally described in Christiansen et al. [14] with reference to the main paper within fleet deployment, Powell and Perakis [44]. Fagerholt et al. [21] is a recent paper and case study on liner shipping fleet deployment using Höegh autoliners as a case study. Another recent paper by Gelareh and Meng [23] considers a liner shipping fleet deployment problem, where the frequency of sailing is decided upon in relation to demand coverage and speed on the individual voyage. The paper gives a thorough review of fleet deployment literature to date. The model is solved using a Mixed Integer Programming (MIP) solver for three and four services using three to five vessel types. Recently, the fleet deployment problem in conjunction with transit time levels was described by Wang and Meng [56]. Meng and Wang [36] expanded Wang and Meng [56] with weekly frequency considerations.

Literature on models and methods for *liner shipping network design* within OR is still limited although recent years show increasing activity in the field. Kjeldsen [30] provides a classification scheme for routing and scheduling within liner shipping. The classification scheme entails 24 references in total for routing and scheduling problems within liner shipping. The classification scheme of [30] shows that there is a large variation in the models with respect to cost structures, constraints and scope unveiling that we are dealing with a young research field with many possibilities of scoping network design to various decisions such as scheduling, fleet sizing, fleet deployment, route generation, speed optimization and many more.

An early contribution is that of Rana and Vickson [45], who present a model for liner shipping with non-simple routes by using a head- and back-haul structure. The model does not allow transshipments, which are at the core of liner shipping today. Fagerholt [20] develops a model and solution method for a regional carrier along the Norwegian coast. The model assumes the carrier loads at a single port and finds optimal routes of vessels to service the unloading facilities. The problem may be dealt with as a VRP, given that a designated depot is known and transshipments are not allowed. The solution method is a MIP solver. Similarly, Karlaftis et al. [29] solve a

problem for the region of the Aegean sea using a genetic algorithm related to the coastal freight problem described in Sambracos et al. [46]. These models do not deal with the important concept of transshipments at multiple ports and the resulting interaction between different services.

Agarwal and Ergun [1] are the first to include a *weekly frequency constraint* by grouping vessels into vessel classes in the *simultaneous ship scheduling and cargo routing model*. The model creates routings for a set of vessel classes with a rough schedule. The weekly frequency constraint determines the number of vessels deployed to a service according to the service duration, and they introduce a time-space graph spanning each weekday to reflect the availability of transportation on a certain weekday. The model allows for transshipments, but the cost of transshipments cannot be derived and hence the overall cost structure excludes the important cost of transshipment. The time-space graph gives a temporal aspect in terms of a rough schedule based on weekdays, but the time-space graph is not utilized to reflect transit time restrictions. Blander Reinhardt and Pisinger [9] present a model which is among the first to include transshipment cost and present a branch-and-cut method for solving a liner shipping network design problem with butterfly routes to optimality. The model selects a route for each individual vessel in a fleet. This results in a distinct network layer for each vessel and the grouping of vessels to a single service is not explicit in the model. The configuration is suitable for smaller carriers. Global carriers tend to group a set of vessels with similar characteristics to a single service to reduce the complexity of the network design [49, 41] and to provide simpler schedules to their customers.

Alvarez [5] is the most recent publication on LSNDP considering the *joint routing and fleet deployment model*. It bundles a service with a vessel class, the number of vessels deployed to the service, a target speed, and a non-simple cyclic port sequence. The cost of a service in the model reflects the deployment cost of the vessels in the service and the estimated bunker consumption adjusting for the difference in bunker consumption when sailing at different speeds or idling at a port. The total fleet cost accounts for whether vessels are deployed to services or are chartered out at market prices. Cargo revenues and handling costs are accounted for in the model along with a penalty for cargo that has been forfeited due to insufficient capacity or insufficient revenue in the view of total network cost. The model assumes a planning horizon with stable demands and does not impose restrictions on the frequency of service or the actual scheduling of the services. A case study with discussion on data generation and port/vessel class incompatibilities is provided. The transshipment cost of non-simple routes is, however, not calculated coherently, as the model cannot detect such a transshipment.

Agarwal and Ergun [1] prove the NP-completeness of the *simultaneous scheduling and cargo routing problem*. The LSNDP in general may be viewed as a VRP with split-pickups, split-deliveries, multiple cross docking, no depots and a heterogeneous vessel fleet and hence the LSNDP is strongly NP-hard as proven in section 2.6.4. The methods deployed to solve the problem are mainly heuristics based on integer programming and decomposition techniques. The branch-and-cut method of Blander Reinhardt and Pisinger [9] and a MIP formulation of Alvarez [5] solve smaller instances to optimality. Remaining solution methods are heuristic methods. An overview of the reviewed methods on liner shipping network design is given in Table 2.1.

Computational results are scarce and based on individual data sets, but it is clear from the literature that solving large scale instances is a hard task even for heuristics. The benchmark suite presented in this paper will aid comparison between different methods if the public benchmark data is used for computational results. In Christiansen and Fagerholt [12] the importance of creating public benchmark instances for maritime transportation problems is advocated in order to drive basic research on maritime optimization problems forward.

Alvarez [5] identifies several problems with exact optimization methods. Firstly, the feasible solution space of a single rotation is vast. Combining this with a vector of different capacities and cost for every feasible rotation increases the feasible solution space significantly. Due to economies of scale, a solution to the LP-relaxation will prefer to use a fraction of large vessels, leading to highly fractional solutions. This is amplified by the cost structure, where the majority of the network cost is placed on the integer variables, while the revenue follows the fractional variables representing the cargo flow. Therefore, convergence of decomposition techniques such as column generation might be slow.

Article	Routes	Fleet	Routing	Method	Optimal	Constraints	vessels/ports
Rana and Vickson [45]	multiple	heterogen	in/out bound	Lagrange, Benders	No	Capacity, time, connected	3v, 20p
Shintani et al. [47]	single	vessel class	in/out bound	Genetic algorithm	No	Empty repositioning, bunker cost and handling cost	3v, 20p
Karlaftis et al. [29]	multiple	heterogen	simple cycle - VRP	Genetic algorithm	No	capacity and time windows on demands	not fixed, 26 p
Fagerholt [20]	multiple	heterogen	simple cycle - VRP	Route generation/ MIP solver	Yes	capacity, route duration	20v, 40p
Blander Reinhardt and Pisinger [9]	multiple	heterogen	butterfly cycle	Branch-and-Cut	Yes	capacity, time, connected, transshipment	6v, 15p
Agarwal and Ergun [1]	multiple	vessel class	simple cycle	Greedy, column generation, Benders	No	weekly frequency, transshipment	50v, 10p
Alvarez [5]	multiple	vessel class	non-simple	MIP solver/tabu search	Yes/No	transshipment, bunker	100v, 7p (MIP solver), 120p (Tabu search)

Table 2.1: Overview of the models on liner shipping network design published in international journals. *Routes*: refers to whether the problem is defined for a single or multiple routes/services. A *Fleet* is homogeneous if only a single type of vessel (in terms of capacity and technical specifications) and heterogeneous for multiple vessel types. *Routing* denotes the way routes are perceived and hence constructed by the model, see section 2.4.2. *Method* refers to the algorithmic solution method, *optimal* to whether or not the method proves optimality of a solution, *Constraints* are a summary of the constraints included. Lastly, *vessels/ports* states the number of vessels and ports in the computational study.

Recent developments in modelling the LSNDP have tried to overcome this division by coupling service generation with the cargo load as seen in Jepsen et al. [28] and Kjeldsen [31]. No computational results are reported for Jepsen et al. [28] as the pricing problem has a very complex structure, which cannot be solved efficiently using current state of the art methods such as resource constrained shortest path problem. Kjeldsen [31] solve the pricing problem heuristically and some limited computational results are reported, which confirm the hardness of this problem even for heuristic methods.

2.3 Contribution

The present paper gives an overview of the domain of liner shipping network design related to mathematical formulations of the LSNDP and other network related optimization problems. It is evident that modelling the LSNDP in a compact formulation, which can exploit state-of-the-art mathematical programming techniques is still a challenge for our community and we expect further development in this area. In this paper we present a reference model, which is an extension of Alvarez [5], where transshipment on butterfly routes is correctly accounted for in the objective function without introducing increased complexity and (bi)-weekly frequency is imposed in the route generation. The model presented is a simplified, but realistic description of the LSNDP as described in this paper. We believe that the model along with the benchmark suite can create a platform for the development of heuristic and exact solution methods. We provide a heuristic solution method for solving the LSNDP and the first computational results for our benchmark suite. Any model adhering to the constraints and the objective of the reference model of this paper, should be able to compare itself to the computational results presented and improve the research direction of the LSNDP. The benchmark suite can cater for extensions such as service level requirements, port productivity and sailing speed optimization to support future research directions within the field of mathematical models of the LSNDP. The benchmark suite is also applicable to other related and promising research problems within liner shipping network design, where the routes or the cargo flows have been fixed to some extent as seen in, e.g., [35, 53].

2.4 Domain of the liner shipping network design problem

The market and business of liner shipping is thoroughly described and analyzed in Stopford [49] and Alderton [3]. Complemented by the experience of network planners and optimization managers at Maersk Line, this section introduces the business context of liner shipping network design.

2.4.1 The liner shipping business

The business of liner shipping is often compared to public transit systems such as bus lines, subways and metro. A *service* is a *round trip* sailed at a given *frequency*. The schedule can be consulted to find the next call of the service at a port. The *round trip* is the set of designated stops, which for a liner shipping company is a set of *ports*. The round trip in liner shipping may be divided into a *head* and a *back* haul to distinguish between different demands according to their regional specifications. The head haul is usually the demand intensive direction. Some services are dedicated to connecting many origins and destinations, whereas others serve a smaller distinct market. In liner shipping a distinction is made between *trunk services* serving central main ports with several demands and *feeder services* serving a distinct market typically visiting a single main port and a set of smaller ports. Like in public transit, a transport may include the use of several services to connect between the origin and the destination of the transport. In liner shipping we refer to transits as *transshipments*. A *fleet* of *vessels* with varying capacity and speed is deployed to the services according to the demand. The *transit time* of a cargo denotes the time a cargo travels from origin to destination. Transit time is counted in days and transit times may vary from a single day to 90 days.

2.4.2 Network

The network must be *competitive* and *efficient*. A *competitive* network may accommodate several routings for one origin-destination pair varying on transit time and cost. Most global liners provide several itineraries for an origin-destination pair by end-to-end services as well as transshipment services with different transit times and freight rates. In order for a network to be *competitive* it must offer low transit times and few transshipments. A *competitive* network serves the main ports of a region frequently with good connections to feeder ports with a high schedule reliability. An *efficient* network facilitates transshipments at terminals with high crane productivity and container capacity to minimize the cargo handling time. Transshipments are also used to get effective utilization of vessel capacity such that the trunk services are fed by several feeder services and direct cargo along the trunk line. Finally, empty containers must be effectively repositioned to ensure availability of containers at the origins of the cargo. This is especially critical for reefer containers as several regions such as South America and New Zealand have a large export of refrigerated goods, but an insignificant import of refrigerated goods. Examples of different liner shipping networks are seen in Notteboom [39], which are included in Figure 2.1. Notteboom [39] argues that global liners are multi-layered networks of different types since they are all competitive for particular circumstances.

Value propositions

A specific liner shipping company is referred to as a *carrier*, whereas the owner of a certain cargo is denoted a *shipper*. The competitive position of a carrier is a combination of port coverage, price, transit time, transshipments, schedule reliability, and, in recent years, corporate social responsibility of the company, where environmentally friendly transport has received a lot of attention. The *freight rate* of transporting some cargo depends not only on the actual network cost, but also on the transit time, the container type needed (e.g., a refrigerated container), special regulations (restricted and dangerous goods), the number of transshipments, cabotage regulations, and naturally the relation between demand and supply for cargo transport on the connection in question. In addition the price covers the administrative overhead incurred by the carrier

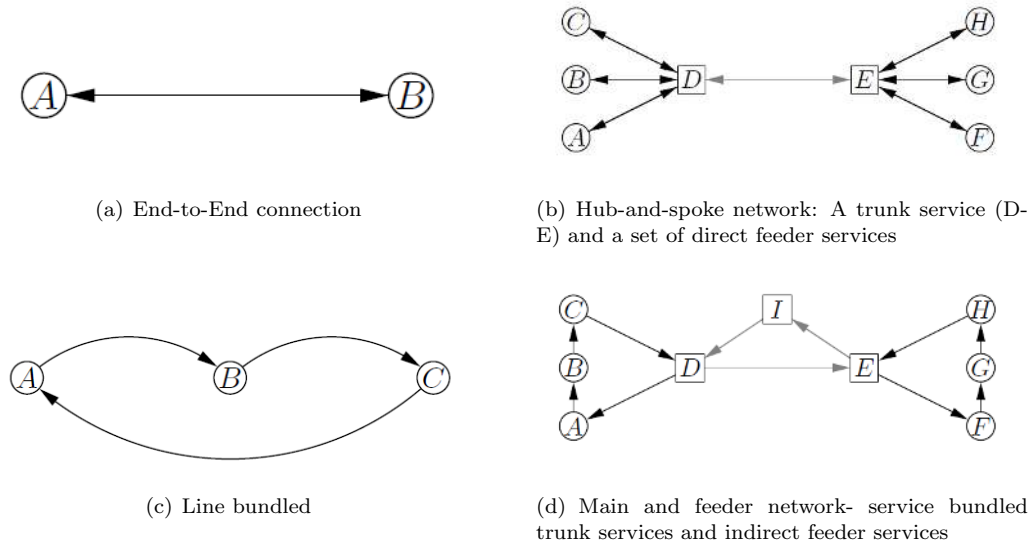


Figure 2.1: Examples of routes and network designs. Square nodes symbolize hubs or main ports, whereas round nodes symbolize feeder ports.

comprising approximately 30% of total cost [49]. Notteboom [40], Notteboom and Vernimmen [41] and Stopford [49] discuss the importance of transit time in pricing cargo transportation as the inventory holding cost is paid by the shipper until the cargo is delivered [40]. Furthermore, some goods are perishable and the time to market becomes crucial. It is becoming increasingly important for companies to ensure transport with the lowest CO_2 emission and hence slow steaming (in particular on back haul connections) to reduce network bunker consumption and overall CO_2 emission.

Forecast and planning horizon

A demand forecast for a given planning horizon is crucial to network design as the ideal network has a perfect fit between demand and capacity [49]. The demand for container transport fluctuates over a year with seasonal variance and peaks at certain times of year. These peaks may be *regional*, if they are related to a crop (e.g., bananas or lemons) or *global* (e.g., Christmas). Hence, the network design is rarely stable over a yearly period. Some structure is fairly stable, but additional structure will reflect seasonality and hence the planning horizon for a network is important to the liner service network design and to the fleet deployment. General economic and financial conditions have a major impact on the liner service network design and fleet deployment, but are hard to predict compared to a seasonal pattern, which may be recognized and accounted for.

Service

A service from the carriers’ point of view consists of a set of port calls with designated arrival and departure times at a given frequency. The individual services are required to be good components of the network with regards to efficient transshipment facilities and fast direct services for critical connections. Notteboom [39, 40] categorize a wide range of rotation patterns as follows:

End-to-End: A direct shuttle service between 2 ports (see Figure 2.1). *Line Bundled:* A rotation visiting a set of ports in a loop. *Trunk Services:* An end-to-end connection between hubs. *Direct feeder services.* Shuttle from feeder port directly to a hub. *Indirect feeder services* - a service bundled rotation to a set of feeder ports and a hub. *Round-The-World (RTW):* A rotation following the equatorial belt visiting hubs in order to service the east-west and north-south trades in a grid. This type of service has a capacity constraint as it must traverse the Panama canal

restricted to panamax ships. *Pendulum*: A service traveling back and forth like a pendulum, e.g., Europe - Far East - US west coast - Far East - Europe. *Butterfly*: Multiple cycles centered around one port (Figure 2.2). Each cycle visiting an alternating sequence of ports. One cycle may be a subset of another. *Conveyor Belt*: A service connecting regional hubs designed for transshipments between continents at the crossing point of trade lanes.

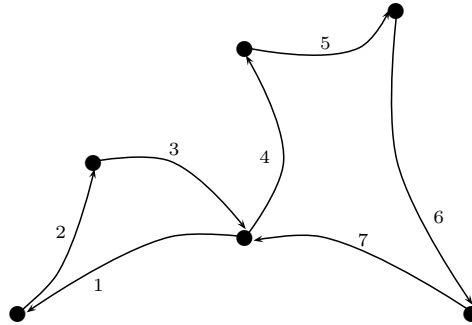


Figure 2.2: An example of a butterfly rotation

Rotation turnaround time varies from a single week up to 20 weeks, although the average rotation is around 8-9 weeks. The rotation turnaround time is composed of voyage time at sea and the service time at ports used for piloting in and out of the port, berthing and loading/unloading cargo. The number of port calls in a rotation is a trade-off between economies of scale and transit times [40]. Visiting many ports leads to a good utilization of the vessel capacity at the cost of long transit times. Another complicating factor is that port stays are very time consuming. Notteboom [40] reports that 21% of the transit time is the accumulated port stay for a COSCO Europe-Far East service. The schedule includes buffer time to account for delays due to weather, port congestion or unusually high terminal handling time. These delays may cause speed increases on individual voyage legs which increases energy consumption and CO_2 emission.

Frequency

As described in Section 2.4.1, the liner shipping business is characterized by having a public schedule. From the carrier's point of view a schedule consists of a set of *services*. The services consist of a fixed itinerary of ports typically called with weekly or bi-weekly frequency [41, 49]. The fixed (bi)-weekly frequency is widely used in container liner shipping due to significant planning advantages for carriers, shippers and terminals:

- **Reliability.** Fixed departures enabling complete integration of customer supply chains.
- **Simplified network planning.** A guiding rule for the carrier when designing the network.
- **Asset planning.** The use of port berths and vessels can be planned for better utilization.
- **Planning routing scenarios.** Synchronization of connecting services may provide timely transshipments for critical connections.

The stable flow of general cargo in containers enables the carrier to maintain a weekly or bi-weekly service. For instance, the weekly frequency of a schedule is achieved by deploying multiple vessels sailing one week apart, as illustrated in Figure 2.3. The speed of each voyage between port pairs is thus closely related to the number of vessels. Therefore, adjusting the number of vessels on a service is a means to adjust speed. At the same time, the decision of speed enables the carrier to adjust to demand fluctuation as the carrier can free up vessels by increasing speed. These vessels can then be deployed elsewhere, to increase capacity.

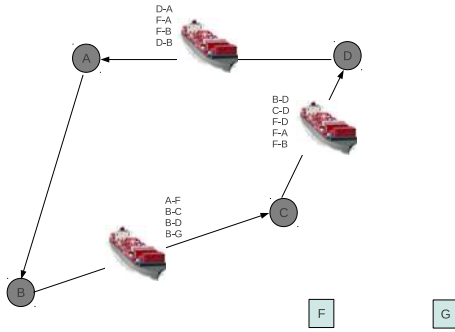


Figure 2.3: A single service connecting ports A,B,C and D. The vessels are depicted along with a cargo list specifying the current cargo load. Total round trip time is 21 days and to provide weekly frequency three vessels are sailing one week apart. The cargo is destined for nodes on the service as well as nodes not belonging to the service. The cargo has multiple origins and destinations.

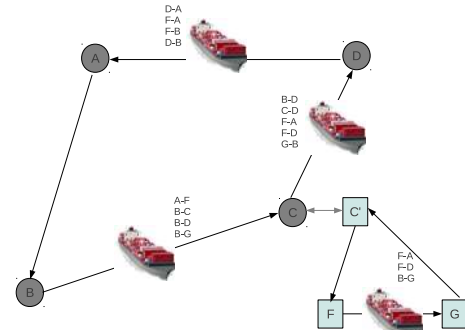


Figure 2.4: Two connecting services. The service from Figure 2.3 and a second service with a round trip time of 1 week illustrated by a single vessels. The cargo composition on board vessels illustrate transshipments at the core of the liner shipping network design.

Business rules for service generation

The generation of services has a high degree of freedom in terms of the number of port calls and average rotation time. In spite of the route flexibility there are certain business restrictions arising from complex real world constraints. The real world constraints may be expressed as “*rules of thumb*” to the combination of the vessel class deployed, the rotation turnaround time and service frequency, according to Network and Product at Maersk Line [38]. Some important business rules are stated as follows:

1. Services deploying a vessel class of capacity of at least 1200 *Forty Feet Equivalent Unit*(FFE) must have weekly frequency. This is a basic structure in *liner* shipping, as customers often plan their production for weekly deliveries. It also facilitates easy management of shared resources between carriers as terminals and canals. Smaller vessels calling smaller ports are exempt from this due to low demand and / or unstable terminal service. Vessels with capacity below 1200 FFE are required to have at least bi-weekly frequency.
2. Services deploying a vessel class of capacity of at least 4200 FFE must have at least 4 week rotation time, as larger vessels require higher volumes to fill and thus be economically competitive.
3. Services deploying a vessel class of capacity of at most 800 FFE must have at least 2 feeder port calls in between every hub-port call.

Business rules 1 and 2 are implemented in our algorithm and thus present in the computational results. Business rule 3 has not been implemented as distinguishing the sequential mix of hub and feeder port calls is outside the scope of the heuristic column generation algorithm.

2.4.3 Assets and Infrastructure

Cost structure and economies of scale

The network cost of a carrier can be divided into fleet cost and cargo handling cost, detailed below (we omit the administrative overhead estimated to around 30% by Stopford [49]):

- **Fleet Cost:**

1. *Bunker Cost* is the cost of bunker, which is the fuel consumed by container vessels.
2. *Capital cost* is the cost of acquiring or financing a single vessel v .
3. *Port Call Cost* is a terminal fee for calling a terminal with a given vessel v . (See below for costs related to cargo handling at ports).
4. *Canal Cost* is the cost of traversing a canal with a given vessel.
5. *Operational Cost (OPEX)* is the operating cost of a vessel including crew, maintenance and insurance.

- **Cargo handling Cost:**

1. *(Un)load cost* is the container handling cost, full or empty, at a given port.
2. *Transshipment cost* is the cost of transshipping a container, full or empty, in a port.
3. *Equipment cost* is the cost of owning / leasing containers.

According to Stopford [49, Table 13.9], the bunker cost is 35-50% of a vessel's cost, capital cost is 30-45%, OPEX is 6-17% and port cost 9-14%. Generally bunker cost exceeds capital costs (apart from the largest vessels).

Bunker consumption depends on the vessel type, the speed of operation, the draft of the vessel (i.e., the actual load), the number of operational reefer containers powered by the vessel's engine and the weather. Bunker consumption for a vessel profile is often approximated by a cubic function of speed per time unit [3, 49]. An empirical study of bunker consumption for container vessels at different speeds is presented in Wang and Meng [54], who conclude that a cubic function of speed is a good approximation of the bunker consumption. During a round trip the vessel may sail at different speeds between ports. The vessel may *slow steam* to save bunker fuel or increase speed to meet a crucial transit time. Speed may be constrained by hard weather conditions or navigation through specific areas.

Both capital and operational cost varies with capacity. Economies of scale means that a large vessel is cheaper to operate per FFE, which is the most common container unit [49]. The market rate of a vessel is called *Time Charter Rate (TC rate)* and represents the cost of leasing (charter in) a container vessel into the fleet or for a carrier to forward lease (charter out) an owned vessel to another carrier. The TC rate fluctuates with seasonality and is highly dependent on the length of the chartering period. A carrier may have an owned fleet supplemented by chartering in and out to meet capacity requirements and to gain flexibility in asset management. TC cost will include daily running costs of the vessel, such as crew costs, repair and maintenance. For vessels owned by the carrier the TC cost will cover operational and capital cost and depreciation of the vessel's value (see [49] page 544). It is assumed that the TC costs represent a market rate, where the carrier will be able to charter out the vessel in case of a surplus of vessels. The methods described herein will consider fixed TC costs, not considering financial asset management of a fleet of container vessels under an expected development of TC costs. For more details on such ideas see [7].

Port call cost and canal costs may be treated identically. The port call cost is a fee paid to the terminal. The fee depends on the size of the vessel, i.e., the capacity booked in the terminal and also on the geographical location and size of the terminal. The cost of traversing the Suez Canal and the Panama canal depends on the size of the vessel and the actual load of the vessel.

(Un)loading and transshipment cost are also known as *cargo handling costs*. The loading and unloading costs are fixed once the cargo is selected for transport, but the transshipment cost depends on the routing of the product and hence the total number of transshipments. A global carrier will not provide direct connections for a significant percentage of the available transport scenarios, which incur transshipment at least once during a voyage. The cargo handling cost can be a non-linear function as some terminals have volume dependent costs.

Equipment cost is the cost of containers. The carrier can own or charter the containers used for cargo transport, [49]. Stopford [49] estimates the daily cost of a FFE dry container to around \$1 per day, whereas a reefer FFE has a daily cost of around \$5.60.

Vessels

An ocean-going container vessel is the core part of a carriers' operations. It can be characterized by specifications as *FFE capacity, weight capacity, speed, length, beam, draft, number of reefer plugs, ice class, age, engine power, etc.* The defining attribute is FFE capacity given as a nominal number. The actual capacity of a vessel depends on the service it is sailing and the actual cargo on board. A vessel cannot accommodate more reefer containers than it has reefer plugs. The draft, length and beam of the vessel dictate which ports, canals and straits the vessel can access. Some waters have special access restrictions like the gulf of Finland which during winter requires ice class vessels for service. With regards to network design, the vessels are grouped in vessel classes with similar properties such as capacity and speed interval, e.g., a Panamax vessel class denoting the maximal width for traversal of the Panama Canal. Other common groupings are according to a capacity band, i.e., vessels with a nominal capacity of 1500-2100 TEU.

Each vessel has a minimum speed S_{min} , and a maximum speed S_{max} , in knots, and a *design speed* S^* , and *design draft* at which its design fuel consumption F^* , is optimized. Still the actual speed largely decides the fuel consumption. Large vessels may use in excess of 200 metric tons (mt) of fuel per sailing day. Additional fuel is consumed by auxiliary engines for other vessel systems (1-12 mt/day) and for electricity for reefer units (a rule of thumb is 0.025 mt/day/reefer, depending on inside/outside temperature). Calculating actual fuel consumption is very complex as vessel draft, wind, waves, currents and date of the last hull scraping affect fuel consumption.

Some seaboards are under sulphur emission restrictions, limiting the percentage of sulphur content in the bunker, denoted LSFO (Low Sulphur Fuel Oil, as opposed to High Sulphur Fuel Oil (HSFO)). LSFO is supplied with a premium to the bunker price and has reduced availability, for details refer to Plum and Jensen [43].

Ports

A port may consist of several terminals competing for the cargo traffic in the corresponding port. A carrier will typically use a single terminal at every port because of connections between services. Ports have a maximum draft and the berths have a maximum length. This results in port-vessel incompatibilities for some port-vessel combinations. A container vessel is piloted in and out of port by a pilot employed at the port authorities. *Pilot times* may be several hours and for ports situated up a river bed it can be 8-10 hours. At some terminals a berthing slot is reserved in advance, whereas others serve vessels by a first-come-first-serve (FCFS) basis. A vessel may have to wait for a given pilot time or wait for an available berthing slot at FCFS ports. To enter/leave some ports a vessel may also have to wait for high tide. A port calling fee, which depends on the vessel's specifications, is levied by the port authority. The port calling fee covers expenses to the pilots, tug boats, the port authorities and the terminal. Once the vessel has entered a terminal the vessel will commence unloading and loading of cargo. Cargo handling is often referred to as *moves* in general. Each move is associated with a cost to cover the expenses of the cranes, terminal crew and terminal administration. A transshipment move will typically cost less than a load and an unload move together. The crane height and length of the terminal may also result in port-vessel incompatibilities. The number of moves per unit time that a port is able to perform is denoted as its *productivity*. There is usually a distinction between main, transshipment, and feeder ports. Feeder ports are usually small and their productivity may vary with the level of technology deployed at the feeder port. Main ports have a large quantity of import/export and some transshipment facilities. Most main ports will have medium to high productivity. Lastly, transshipment ports such as Balboa, Singapore and Algeciras do not have extensive import/exports but serve to transship between services and inter-modal transport. Transshipment ports usually have a high productivity.

Port Stay

In real life port stay times will vary greatly due to the different productivity and strategies of terminals either having reserved berthing slots or perhaps serve vessels on a first come first serve

basis. Furthermore, the port stay depends on the amount of cargo to load and unload in the given port for that particular calling. The benchmark suite does not include measures of productivity, nor do we model the port time on basis of the cargo flow, and therefore it has been chosen to fix the port time to an identical measure for all ports and vessel classes. This is motivated by the observation that larger vessels in practice will have higher productivity, since they call more efficient ports served by more cranes. This means that vessels irrespective of size often have comparable port stay times, an analysis of Maersk Line port stay times roughly supports this.

Canals

The two large canals of Panama and Suez enable fast transport between continents. To traverse the canals, a substantial fee is charged. The canals offer significantly faster transit times and also reduced operational costs due to the reduction of the sailing distance. Some sailing passages may have draft and width restrictions such as the Panama or Kiel Canal resulting in a significant detour for larger vessels. Some, as the Bosphorus straight, are the only entrance to a sea body and the limitations of these straights thus dictate the size of entering vessels. The current restriction on vessel size of the Panama canal is very decisive in the services of the American east coast from the Far East and Australasia. No such restriction exists for the Suez canal with current container vessels.

Equipment

Containers are generally available in lengths 20, 40, and 45 feet each of which exist in different types as: Dry, refrigerating (*reefers*), open top and rack, plus additional specialized containers. A customer will generally require a specific type of container, though in some cases containers are used outside their designated scope. *Dry* equipment covers roughly 80% of a carriers equipment pool [49] with remainder mainly being *reefers*. There is a similar distribution on container sizes with around 80% being 40' containers [49].

2.4.4 Demand and Customers

The goods and customers of containerized transports are plentiful and cover any type of manufactured goods, but can also be bulk-like cargo such as stone, waste-paper or refrigerated commodities like fish or fruits. A service will usually focus on customers in a *trade* (sometimes multiple trades), e.g., *Asia to Europe* or *South America to North America*. They may be grouped into east-west trades, which have larger volumes, allowing for economy of scale deploying huge vessels and into north-south trades, which have much refrigerated cargo requiring specialized vessels and faster transit time. Similarly all trades have special characteristics with regards to volume, service level requirements, etc. that govern how a service can compete in the trade.

Demand

The production and consumption of some commodities vary over the year, some following harvest seasons, others following holidays and festivals or summer vacations. As a result, demand varies, some on a global level others only affecting a single port or trade. The biggest of these is Christmas, which generates the yearly peak season in the third quarter, allowing warehouses to fill up prior to Christmas shopping. Another peak happens prior to Chinese New Year as Chinese factories exports their goods before closing down for the festivities. A customer or shipper will usually only pay for a container transport once it is on the vessel or delivered, i.e., there is no fee for booking and no fee for not submitting a cargo for a booking. This risk-free booking policy for the shipper causes many problems for the carrier and shipper:

- It makes it hard to forecast the demand of cargo for some vessel departures, (although it can be used at ports called later), so, the carrier will overbook departures. This causes delays of cargo when forecasts and bookings fall short.

- Shippers know that overbooking is a risk, and, to counter this, they will intentionally book for more containers than they expect to ship, often spread at different carriers to optimize flexibility. This especially happens in peak seasons as the third quarter, where capacities are most pressed.

Containerized transport has been a growing industry for many years. Between 1983 and 2006 world GDP growth was 4.8% per year, whereas container cargo growth averaged to 10% per year [49]. The trend has continued in recent years looking at the data from The World Bank [51] with an exception of the years 2009 and 2011 where the industry was suffering from worldwide financial crisis. The long history of fast growth in containerized cargo has given an expectation of continuous growth in the industry. The delivery time for new-build vessels is several years and this makes the market slow to adapt to increased demand with resulting huge fluctuations in revenues, as seen in the aftermath of the financial crisis in 2009.

Revenue

In general the revenue of a cargo is closely related to the *demand-supply* balance between the volume of cargo to be transported and the capacity offered by container carriers. But many other aspects come into play. Specialized cargo, requiring specific equipment or different administrative handling will give higher revenue, for instance refrigerated or military cargo. As the demand is not symmetrical (e.g., Asia exports more than is imported) but the supply is symmetrical (vessels return to Asia to be reloaded), hence also the revenue is not symmetric. Transporting a container from Asia to Europe can easily cost three times more than the reverse ([4]).

The demand fluctuations over a year influence the revenues, with the highest revenues in the third quarter, but with variations over trades.

Service level

The specific service a customer is presented with, for transporting a container from port A to port B , the *product*, can be classified by a number of parameters:

- The price: A basic rate, subject to additional fees and surcharges.
- Bunker Adjustment Factor (BAF): A variable price component dependent on the bunker price, making it possible for the carrier to share the risk of oil price volatility with the shipper.
- Transit time: The time to transport a container from A to B .
- Transshipments: The cargo is reloaded onto a new vessel in a third port. A direct product is usually preferred due to the decreased risk of missing a connection.
- Frequency: A weekly service calling a port is preferred to a bi-weekly service due to increased flexibility.
- Reliability: Ocean carriers differ in reliability of services and the products they operate (Notteboom [40], Maersk Line [34]).
- Paperwork: As documentation and administration, different carriers and products require different paperwork, adding complexity for the shipper.
- Equipment: The carrier must supply the customer with a container prior to transporting.
- Weekday: Vessels departing on a Friday can, e.g., be preferable to a departure on Monday, as it can shorten the supply chain with three days (for a factory closed in the weekend).

All these factors can be relevant for a shipper, with price and transit time often being the key factors.

Routing (Products)

As a result of trade policies there may be various *cabotage rules* within countries restricting internal transport within the country to routings on vessels flagged/owned/operated in the country. Cabotage applies to transshipments of cargo destined within the country and also for empty containers. Other special rules may exist, e.g., embargoes: when a country A is embargoing a country B , cargo from/to B can not be transshipped in A . Dangerous goods are subject to *IMO rules* with regards to the quantity and the placement of these goods on the container vessel.

Competition & Partnerships

As mentioned the economies of scale in liner shipping can generate huge savings for larger volumes. This motivates carriers to cooperate on operating services or subcontract. Such partnerships exists in various forms.

Foreign Feeder Agreements (FEF) A service operating between a hub and a few proximate ports is called a *feeder*. These are usually operated by smaller vessels (less than 1250 FFE) and the cargo is transshipping in the hub to some ocean-going service. The feeder carrier will combine volumes from other global carriers to achieve economy of scale.

Alliances and Vessel Sharing Agreements (VSA) The significant economies of scale achieved by deploying larger vessels is the motivation for multiple carriers to enter a VSA. Two or more carriers share a round trip, making it possible to deploy larger vessels and providing higher frequency of service. Deploying large vessels presents a significant saving in operating cost for each carrier (see Figure 13.11 in Stopford [49]) without reducing the service levels presented to their customers. In practice, VSAs are complicated because the partners seldom have matching services and demand, available vessels, etc., but VSAs are widely used and a central part of liner shipping. For more details on the mechanics of VSAs, refer to [2] for game theoretic observations on the subject.

Slot Charter Agreements (SCA) are a combination of FEF's and VSA's. One carrier will enter a contract to use capacity (*slots*) on another carrier's service. This can either be on the full round trip of the service, allowing for the slot to be used differently in head and back haul, or on a specified part of the rotation. The contract is usually for a fixed number of slots which is paid for *used/unused*, sometimes with an option for extra slots as *pay per use*.

2.5 The benchmark suite for LSNDP

The purpose of the benchmark suite is to provide data for prototype implementations of models and methods for the liner shipping network design and fleet deployment problem at a strategic level. The literature review displays three important factors for the development of the benchmark suite:

1. The research community views the problem from different angles in terms of constraints and costs. As a consequence the data must address the important factors seen across the entire research field and also incorporate data on foreseeable extensions to current models.
2. A “base” model is needed in order to compare methods.
3. Currently, exact methods only solve very small instances, while heuristics such as the matheuristic by Alvarez [5] caters for a large instance. As a consequence the benchmark suite should contain small, realistic instances, that may be solved to optimality within foreseeable future and large instances for heuristic development.

There are at least two pressing research questions on the LSNDP: efficient modeling of bunker consumption as a cubic function of speed and incorporating the level of service. Global carriers

such as Maersk Line put great effort into creating a network design of low bunker consumption to meet a demand for environmentally friendly transportation of goods and improve the economy of the network. As noted by Notteboom and Vernimmen [41], the price of bunker has a significant impact on the network design and the fleet deployed. Therefore, formulas for bunker consumption based on vessel class are included in the benchmark suite to allow for calculations of the overall bunker consumption as a function of speed. At the same time, Notteboom and Vernimmen [41] state that customers are demanding a high frequency and low transit times and therefore, level of service and competitive transit times are important factors in LSNDP. The benchmark suite includes a set of maximal transit times on commodities as we believe level of service and transit time incorporation to be important open research questions in the liner service network design problem.

We have chosen to provide vessel data in terms of six fictitious, but realistic vessel classes, where multiple vessels share similar characteristics. We believe a global carrier has several similar vessels assigned to a service [1, 41, 5]. A prototype implementation will contain the most important details, but omit details believed to be non-decisive for the conceptual understanding of the problem. Therefore, the benchmark suite does not contain different equipment types such as reefer and high cube containers. The weight of vessels and cargo in particular are omitted from the benchmark instances.

Assigning specific vessels to a specific service is an operational issue dependent on cabotage rules. Hence, the benchmark instances do not indicate the country of registration of vessels in the present data set. Likewise, IMO rules apply to specific cargoes and stowage plans, which is out of scope of the considered LSNDP.

Weather and currents may have a large influence on travel times and energy consumption in a LSNDP, but we have chosen to ignore these factors to avoid further complexity.

The benchmark instances are not meant for actual scheduling of services. Therefore tidal information on ports is not provided although they may be decisive for a specific berthing window at a port. On the same note, the specific weekday on which a demand becomes available for transport out of a given port is not included in the data of the benchmark suite.

The following section discusses data objects and data generation. A discussion on the range of instances needed to qualify the benchmark suite follows.

The benchmark suite originates from publicly available data and past data for Maersk Line.

External data sources are: “Distances between ports - Publication 151”, National Imagery and Mapping Agency [37], The Hamburg Index, Vereinigung Hamburger Schiffsmakler und Schiffsbagenten [52], “Charter Rates 2000-2010”, Alphaliner charter rates 2000-2010 [4], “Container Freight Rate Insight”, Drewry Shipping Consultants [19].

The data from Maersk Line is from an undisclosed year and has been subject to reasonable random perturbation and measures to anonymize the data in order to protect the confidentiality of Maersk Line.

The purpose of the benchmark suite is to qualify algorithms for scalability, quality, and robustness by providing realistic data for both the scale, structure and complexity of the liner shipping network design problem. We have chosen not to anonymize the actual ports to be able to evaluate the network from a business perspective, but the data still does not represent an actual business case, nor does it contain all relevant data. We hope the benchmark suite can be used to perform strategic analysis on various scenarios. The instances may be found at <http://www.or.man.dtu.dk/English/research/instances> under the name *LINER-LIB 2012*. The following section describes the data generation process in terms of the origin of the data and the logic used to derive the benchmark suite.

2.5.1 Data objects and data generation

An instance consists of the following data: **Port list** with port call cost as a function of vessel capacity (explained in more details in Section 2.5.2). **Fleet list** with design speed and bunker consumption at design speed and idling (Section 2.5.3). **Distances** contains an all-to-all list of travel distances (Section 2.5.4), and **Demands** specifies a set of origin-destination pairs and their

corresponding demands (Section 2.5.5). Finally, the **Graph** file defines a graph representation of the considered problem with way points. The data items are illustrated in Figure 2.5 giving an overview of the attributes of each item.

Port List P	Fleet List V	Cargo Demand List C
Port α^P	Vessel class v^F	Origin port α^C
Name n^P	Capacity c^F	Destination port β^C
Country o^P	TC rate T^F	Quantity q^C
Cabotage a^P	Draft δ^F	Freight rate f^C
Region r^P	Min speed s_{\min}^F	Max transit time t^C
Latitude x^P	Max speed s_{\max}^F	
Longitude y^P	Design speed v_*^F	
Draft δ^P	Fuel consumption f_*^F	
Move cost m^P	Fuel consumption f_0^F	
Transshipment cost t^P	Quantity of vessels q^F	
Fixed portcall cost f^P	Suez fee s^F	
Variable portcall cost v^P	Panama fee p^F	

Graph
Vertex(Port, Waypoint, Canal)
Edge(Source, Target, Distance, Draft)

Distance List D
Origin port α^D
Destination port β^D
Draft δ^D
Distance d^D
Suez traversal s^D
Panama traversal p^D

Figure 2.5: Data objects of an instance

LINER-LIB 2012 does not dictate the length of the planning horizon. Given that weekly service frequency is most often encountered in practice [41, 38], the demand figures provided correspond to weekly volumes. One may obtain total flow volumes for an arbitrary planning horizon by scaling the weekly figures (e.g. multiply the given volumes by 52 to obtain yearly transport demand). The cost and revenues are not meant to reflect the actual cost of any carrier. The costs and revenues are constructed with care to reflect the relative cost structure within the network. This means that a large port is proportionally cheaper to call at than a small port. Likewise, it is cheaper to transship a container in Asia than in Europe or the US. Revenues reflect demand and supply such that visiting a distant port with low demand require higher revenues per FFE than a central port with extensive own demand and supply.

2.5.2 Port List

Ports are identified by UNLOCODE, a unique 5-character identifier [22]). The port list specifies the ports in the instance and contains the following fields: Port (in UNLOCODE); Name; Country; Continent; Cabotage region if applicable; Revenue Region (maps port to the revenue data supplied by Drewry Shipping Consultants [19]); Latitude; Longitude; Maximal berthing length (in meters); Maximal acceptable vessel draft (in meters); (Un)Load cost per FFE (in USD); Transshipment cost per FFE (in USD); Fixed port call cost (in USD); Variable port call cost per FFE (in USD) as seen in table 2.2.

The port list reflects real ports ensuring that a solution can be mapped to a geographical coverage. This allows network planners and others without optimization background to evaluate a proposed network.

The transshipment cost will usually be lower than the sum of loading and discharging as transshipping does not require customs paperwork at the terminal in question. In an optimization model, the load and discharge may be viewed as fixed costs once the demand is chosen for transport.

The geographical location and the port size will determine a fixed cost for visiting the port and a variable cost related to the capacity of the vessel visiting the port. This relation is deduced

Port α^P	Name n^P	Country o^P	Cabotage a^P	Region r^P	Long x^P	Lat y^P	Draft δ^P	Move m^P	Trans t^P	Fixed f^P	Var v^P
CNYTN	Shenzhen	China	China	South China	114.26	22.58	13.5	176.69	78.14	7220.21	3.8

Table 2.2: Example of port entry. Port is given by UNLOCODE. **Draft** is the maximal draft. **Move** is the (un)load cost and **Trans** is the cost of a full transfer for transshipment. **Fixed** is the fixed port call cost and **Var** is the variable port call cost as a function of the capacity.

experimentally from the perturbed Maersk Line data giving a non-zero fixed cost and a variable cost corresponding approximately to a linear function of the capacity.

2.5.3 Fleet List

Six generalized vessel classes are constructed from the fleet list of Maersk Line [8] representing realistic capacity classes in the network. A vessel class will contain the following information: TC rate per day per vessel (in USD), this value includes operational costs; Capacity (in FFE); Design draft (in meters); Minimal speed (in knots); Maximal speed (in knots); Design speed (in knots); Daily bunker consumption in metric tons at design speed; Own consumption in metric tons when idling at ports; Number of vessels in the vessel class currently in the fleet; Suez canal fee (in USD); Panama canal fee (in USD) as seen in Table 2.3.

The minimal- and maximal-speeds referred to in the fleet list are not the technical minimum/maximum speeds. The speed interval is related to the average speed on a service deploying the vessel class in question. This means the vessel will be doing less than the minimal speed on part of the service and hence the minimal speed referred to in the fleet list is higher than the technical minimum speed. The technical maximum speed will also be higher than the maximum speed referred to in the fleet list and likewise refers to the maximal average speed that can be assigned to a service with the vessel class in question deployed. The minimal speed is the “slow steaming” speed of a service deploying the vessel class. The only size measure describing the vessels is draft. Other measures such as length and beam width, should be seen as a proportional function of draft. This generalization is done to simplify the problem, keeping in mind that a full description of the problem can extend to more than this one size measure without losing applicability.

Vessel class v^F	Cap c^F	TC T^F	Draft δ^F	Min s s_{\min}^F	Max s s_{\max}^F	Des s v_*^F	Fuel* f_*^F	Fuel0 f_0^F	Quantity q^F	Suez s^F	Panama p^F
Feeder_450	450	5000	8.0	10	14	12.0	18.8	2.4	38	175769	64800
Feeder_800	800	8000	9.5	10	17	14.0	23.7	2.5	77	218445	115200
Panamax_1200	1200	11000	12.0	12	19	18.0	52.5	4.0	124	267217	172800
Panamax_2400	2400	21000	11.0	12	22	16.0	57.4	5.3	161	413533	345600
Post_panamax	4200	35000	13.0	12	23	16.5	82.2	7.4	91	633007	
Super_panamax	7500	55000	12.5	12	22	17.0	126.9	10.0	10	1035376	

Table 2.3: Example of the fleet list. **Vessel class** is the name of the class. **Cap** is the capacity in FFE. **TC** is the TC rate. **Draft** is a single measure expressing width, length and draft for port/vessel compatibility. **Min s**, **Max s** is the minimum and maximum average speed interval of the vessel whereas **Des s** is the design speed. **Fuel***, **Fuel0** are bunker consumption at design speed and idling respectively. **Quantity** is the available number of vessels in the fleet with these specifications. **Suez**, **Panama** are the canal fees for the respective vessel classes.

The fleet list reflects the demand scenario of the instance as a trade lane has a smaller fleet than a world instance.

TC rates

The TC rates fluctuate with the market and hence capturing TC rate is very dependent on the date it was retrieved. We have compared data from Alphaliner charter rates 2000-2010 [4] and

“HAMBURG INDEX Containership Time-Charter-Rates” Vereinigung Hamburger Schiffsmakler und Schiffsagenten [52], which have charter rates for vessels up to 4000 and 4800 TEU respectively. For vessel classes above 4800 TEU, we have constructed a TC rate based on the percentage increase of TC rate in the Maersk Line data to corresponding capacity intervals of the fictitious vessel classes and have increased the Hamburg index of charter prices with the corresponding percentage. The TC rates have been ceiled to the nearest thousand for vessel classes below post panamax size and to the nearest 5000 for the two largest vessel classes.

Bunker consumption

Bunker cost is a significant cost component of the network. The bunker curves for the generalized vessel classes are based upon the formulas of Stopford [49] and Alderton [3]. There is general agreement that bunker consumption, F , may be estimated by a cubic function

$$F(s) = (s/v_*^F)^3 \cdot f_*^F \quad (2.1)$$

for any speed s between the *min speed* s_{\min}^F and *max speed* s_{\max}^F of the vessel F , where v_*^F is the design speed, and f_*^F is the fuel consumption at design speed. This function disregards draft although it is a significant factor, so the bunker consumption should be the consumption at design draft as well. The design draft is the draft the vessel has with an average load and for which it is designed to have the lowest bunker consumption as a function of speed. The fleet list specifies the design speed and consumption at design speed for the cubic formula for each vessel class. The bunker price is variable and the effect of bunker price could be a scenario for a user of LINER-LIB 2012. In the computational results we use a flat bunker price of 600 USD per ton.

2.5.4 Distances

The distances between ports are based on information from the National Imagery and Mapping Agency (NIMA, 2001). The data from NIMA contain distances from each port to major way points (ocean junctions) and ports in the vicinity. The data enables a mapping of global sailing routes onto a graph of ports and way points. Each arc is given specifications such as the maximal draft and maximal width such as the Panama canal, which has a width restriction. A single-source shortest path algorithm between each pair of ports determines the shortest feasible path for each vessel class based on this data. This means that the distance between Oakland on the US West Coast and Savannah on the US East Coast will differ for vessels able to traverse the Panama canal as opposed to vessels not able to traverse the Panama canal. To account for canal cost, a parameter indicates whether the distance is based on a visit to the Suez Canal and/or the Panama canal. The distance file is a table with multiple entries for each port pair depending on draft limitations and canal usage.

The distance data consists of: Origin port α (in UNLOCODE); Destination port β (in UNLOCODE); Distance (in nautical miles); Draft limit (in meters) δ^D ; Suez traversal (yes/no); Panama traversal (yes/no) as seen in Table 2.4.

Origin α^D	Destination β^D	Distance d^D	Draft limit δ^D	Suez traversal s^D	Panama traversal p^D
GBABD	USOAK	7953	12	No	Yes
GBABD	USOAK	13862		No	No

Table 2.4: Example of distance file. **Origin** is the origin (in UNLOCODE), **Destination** the destination (in UNLOCODE). **Draft limit** is the draft limit in meters (if relevant). **Suez traversal**, **Panama traversal** indicate whether the respective canals are passed to calculate canal fees.

The canal dues for the world’s two most important canals; Panama and Suez, are included in LINER-LIB 2012. These are based on their published cost structure [50, 42] and created for the relevant vessels classes. Additional canals are omitted from LINER-LIB 2012.

After manual inspection of the distances, several entries were identified to be erroneous. We have attempted to verify the correctness of several distances through external data sources and to correct obvious errors in the data. However, the distances are not guaranteed to be the shortest distances due to the selection of way points and due to possible undetected errors.

The distances have been collected and put into a graph $G = (V, E)$. The set of ports P and the set of way points W constitute $V = P \cup W$. The edge set E is from National Imagery and Mapping Agency [37]. Two distance files are generated

1. An all-to-all distance matrix only for the port set $P \times P$. This graph has few vertexes and many edges.
2. A (sparse) graph $G = (V, E)$, of the instance with $V = P \cup W$ nodes.

The graph file contains all vertices's and edges needed to navigate between any pair of ports in P with the fleet F . The files are generated using a shortest path algorithm between all pairs of ports for each vessel class in the fleet F .

2.5.5 Cargo Demands

Realistic demand data that captures the asymmetry in world trade is important when deciding on capacity and port sequences. A maximal transit time for each demand is provided for future models incorporating level of service or maximal transit time constraints.

The demand table contains the following information: UNLOCODE of the origin port α ; UNLOCODE of the destination port β ; Quantity (in FFE); Freight Rate (in USD per FFE); Maximal transit time (in days) as seen in Table 2.5. The demands are assumed to be the expected weekly demand.

Origin α^C	Destination β^C	Quantity q^C	Freight rate f^C	Max. transit time t^C
USLSA	CNYAT	370	1500	15

Table 2.5: Example of demand file. **Origin, Destination** give the origin and destination of the demand. **Quantity** states the number of FFE needed. **Freight rate** is the freight rate in USD per FFE. **Max transit time** states the maximum transit time in days.

For world instances, demands are aggregated onto the main ports.

The revenues are based on the trade lane prices found in “the container freight rate index”, courtesy of Drewry Shipping Consultants [19]. The rates are independent of any carrier and contains reliable market data collected across several carriers and markets. Drewry Shipping Consultants [19] do not keep data for South America West coast and some region pairs. In such cases the official rates from the website of Hapag-Lloyd [25] have been applied. The revenues are taken from a specific period of time, and it should be noted that these fluctuate greatly dependent on volume changes, new buildings, bunker price, etc. The revenue is 70% of the market freight rates as we are only dealing with the network cost and must be able to pay for the administrative overhead estimated to 30% by Stopford [49]. The volume of the origin/destination ports along with the distance to the nearest main or transshipment port are used to vary the trade lane price according to the demand structure, such that a demand from a transshipment/main port is cheaper than a demand from a feeder port and feeder port revenues are dependent on the distance from the main port. Lastly, the freight rates are perturbed within an interval of $\pm 5\%$ to reflect price variance of different markets and cargoes. The maximal transit time is set 30 % higher than the fastest possible in Maersk Line’s network in 2010.

2.5.6 Instance range of LINER-LIB 2012

The instances are meant to challenge both exact and heuristic methods. Exact methods are currently limited to approximately 10 ports, whereas the tabu search heuristic of Alvarez [5]

provides results for 120 ports. Furthermore, the instances are meant to encourage the progress of method development and hence include instances of real life scale, although we cannot yet solve them to optimality. Most of the instances are subsets of a larger network, where the subset is selected to highlight various aspects of the LSNDP such as using a single hub or multiple hubs, the importance of draft restrictions or the capacity in terms of the fleet size is low compared to the demand and vice versa. The provided instances are listed in Table 2.6. In addition to the instances we have generated a set of rules to transform the *base* cases into two additional scenarios, where there is *low capacity* and *high capacity* compared to demand volume. The scenarios are created by adjusting the fleet data in terms of the number of vessels available and the TC rate, which is adjusted to reflect whether there is a deficit or surplus of capacity in the market. For the *low capacity* case TC-rates are multiplied by 1.4 to reflect a 40% increase of TC-rates, while the fleet quantity is decreased by 20% by multiplying the quantity with 0.8 and rounding to nearest integer value. In the *high capacity* case TC-rates are decreased by 20% by multiplying TC-rate by 0.8 and the fleet quantities are increased by 20% multiplying each vessel class quantity by 1.2 and rounding it to the nearest integer.

Category	Instance and description	Ports	Dmnds
Single hub instances	Baltic Baltic sea with Bremerhaven as hub	12	22
	WAF West Africa with Algeciras as hub to West African ports	19	38
Multi hub instances	Mediterranean Mediterranean with Algeciras, Tangier, Gioia Tauro as hubs	39	369
Trade lane instances	Pacific (Asia-US West)	45	722
	AsiaEurope Europe, Middle East and Far east regions	111	4000
World instances	Small 47 Main ports worldwide identified by Maersk Line	47	1764
	Large 197 ports - the majority of ports serviced directly by Maersk Line	197	9630

Table 2.6: The instances of LINER-LIB 2012 with indication of the number of ports and the number of distinct origin-destination pairs. The instances may be found at <http://www.or.man.dtu.dk/English/research/instances>

The Pacific and the single/multiple hub instances should be challenging for the exact approaches and will also be applicable to model development as models of LSNDP can be verified by analyzing the results of LINER-LIB 2012.

To have a medium sized multi hub instance, the Mediterranean case has been constructed. It includes 39 ports between Morocco and Suez and includes ports in the black sea. Several ports have large demand sets throughout the seaboard. The demands in the Mediterranean instances are constructed and not comparable to a real world setting. The AsiaEurope trade lane instance and World instances should be challenging for the heuristics developing new methods, neighborhoods and competing for the current best known objective value. Exact methods solving these instances to optimality are not expected in the near future. We do not provide computational results for the WorldLarge instance as the size is currently considered out of reach even for heuristic methods.

LINER-LIB 2012 can also be used in other optimization models related to liner shipping network design such as evaluation of a set of candidate services, synchronization of services to reduce transit times, sailing speed optimization etc.

2.6 Reference model for the LSNDP

The following model is based on the model presented in Alvarez [5] with extensions to handle butterfly rotations and weekly or bi-weekly frequencies. Butterfly rotations (illustrated in Figure 2.6) make it more difficult to account for transshipment costs. The planning horizon remains open, but in the route generation a constraint has been added such that the number of vessels are aligned with the planning horizon and service duration to provide (bi)-weekly frequency.

2.6.1 MIP formulation

In this section, we present a possible formulation of the LSNDP. In the maritime sector, a service is defined as a sequence of ports to be visited following a published schedule. While we adhere to this convention, the core entity of our model must be more specific in order to capture all transshipment costs properly. We therefore define a *rotation* as a particular configuration of a service, vessel class, number of vessels deployed, and speed.

Our formulation of LSNDP addresses butterfly routes, as these are often used in practice. This means that the flow balance equations need to be extended. In order to balance the flow at butterfly nodes correctly, a three-index formulation is necessary. Consider the rotation in Figure 2.6, where port c is the butterfly port. With a traditional flow balance at port c , cargo that arrives from b might be redirected towards port a under the same conditions as cargo that continues on to port d . However, cargo travelling on the port sequence $b-c-a$ must be unloaded at c as the vessel continues to d , and later reloaded as the vessel arrives from e . Hence, it appears that we must track the latest port visited by the cargo, as well as its immediate destination, which we can encapsulate using a three-index formulation. Notice, that the cardinality of the problem is not affected by this representation, given that one requires the same number of port pairs or port triples to specify a service uniquely.

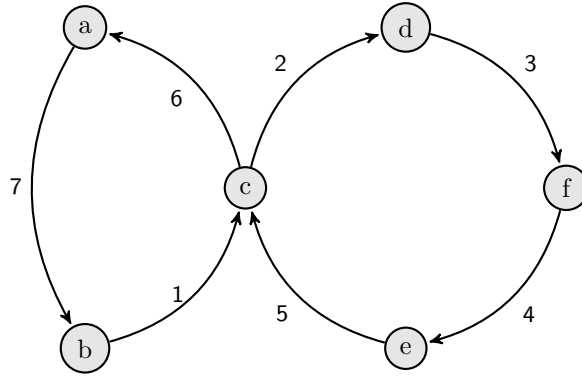


Figure 2.6: An illustration of the flow balancing of butterfly routes. The numbers on the arcs denotes the sequence the vessel sails in the butterfly. Commodities travelling from b to a either follows the vessel on route $b-c-d-f-e-c-a$ or are transshipped at c , where the handling cost of the container must be paid in full.

The following sets are used in the formulation (In square brackets it is described how the sets relate to the graph defined in Section 2.5.4 and data objects defined in Figure 2.5):

- R** All rotations in the model, indexed using r .
- P** All ports in the model. [$\mathbf{P} := P$]
- E** Set of all possible edges in the model [$\mathbf{E} \subset P \times P$]. All edges are directed and uncapacitated.
- E_r** Set of edges used in rotation r .
- Ω_r** Set of ordered port triples (h, i, j) in rotation r . The triples are ordered in the same manner as they will be visited by a vessel in the rotation.
- V** Vessel classes in the model, indexed using v . [$\mathbf{V} := F$]
- G** Set of port pairs with demand for transport, indexed using (o, d) . [$\mathbf{G} \subset P \times P$].

The following parameters represent the known data for the problem (In square brackets it is described how the parameters are calculated from LINER-LIB 2012 data):

a_{ij}^v	Canal costs for vessels of type v when traversing arc (i, j) . [s^F (Suez), p^F (Panama)]
c^v	Capacity (in FFE) of vessels of type v . [c^F]
d_j^v	Port call costs for vessels of type v when entering port j . [$f^P + v^P \times c^v$, where c^v is the number of FFEs the vessel can carry].
e	Fuel price per ton. [\$600 for all instances]
f^v	Daily running costs for vessels of type v over the entire planning horizon. [$f^v = T^F \cdot 180$ as we have a 180 days planning horizon, and daily TC rate is T^F]
\tilde{f}^v	Cost or revenue of excluding a vessel of type v from operations. This can be the revenue obtained from chartering the vessel out over the entire planning horizon, or the cost of laying up the vessel, depending on transportation market conditions. [f^v (vessel in use), $f^v - \tilde{f}^v = 0$ (vessel not in use)]
g^{vs}	Fuel consumption (tons per mile) for vessels of type v steaming at speed s . [A table of possible values of speed and g^{vs} is generated a-priori by the program using equation (2.1)]
h^v	Fuel consumption (tons per day) for vessels of type v when idle at port. [f_0^F]
k_{od}	Total demand (in FFE) to the liner company for transport from o to d over the planning period. [q^C where $o = \alpha^C$ and $d = \beta^C$]
l_{ij}^v	Length in nautical miles of a direct sailing from port i to port j using vessel type v . Notice, for instance, that canal width restrictions will result in different voyage lengths for wider vessels. [d^D where $i = \alpha^D$, $j = \beta^D$ and $\delta^D > \delta^F$]
m_r	Number of round trips performed by a vessel on rotation r during the planning period. We do not impose an integrality restriction on this parameter, as we assume that a fractional portion of the voyage can be completed at the beginning of the following planning period.
p_j^v	Time at port j for vessels of type v . [Based on historical data we set $p_j^v = 24h$ regardless of vessel size]
q_{od}	Revenue from transport of one FFE from o to d . [f^C where $o = \alpha^C$ and $d = \beta^C$]
\tilde{q}_{od}	Penalty for failing to transport one FFE from o to d . [\$1000 USD per FFE in all cases]
s_r	Speed of all vessels used in rotation r . [generated by the algorithm. The code tries several speed-vessel-route combinations (respecting the vessel speed range) and picks the one that leads to a better objective value for that iteration]
u_j	Cost of lifting one FFE at port j . [m^P]
t_j	Cost of transshipping one FFE at port j . [t^P]
v_r	Vessel type used in rotation r .
z^v	Number of vessels of type v available. [q^F]

Finally, the decision variables used in the MIP are:

$X_{(ij)d}^r$	Number of containers travelling to their final destination at port d along edge (i, j) on rotation r .
$U_{(hi)d}^{rs}$	Number of containers travelling to port d which arrive to port i via edge (h, i) of rotation r for transshipment to rotation s .
$W_{(id)}^r$	Number of containers travelling to port d reaching their final destination via edge (i, d) using rotation r .
V_{od}^r	Demand from port o to port d that enters the network for the first time as it is loaded to a vessel in rotation r .
O_{od}	Demand from port o to port d that will not be serviced by the liner company.
Y^r	Number of vessels assigned to rotation r .

The MIP model is presented below, followed by an explanation of the objective function and all the constraints.

$$(VRD) \quad \text{minimize} \quad Z_{MP} = \sum_{r \in \mathbf{R}} f^{v_r} Y^r + \sum_{v \in \mathbf{V}} \tilde{f}^v \left(z^v - \sum_{r \in \mathbf{R}: v_r=v} Y^r \right) + \quad (2.2a)$$

$$\sum_{r \in \mathbf{R}} m_r Y^r \sum_{(i,j) \in \mathbf{E}_r} \left(eh^{v_r} p_j^{v_r} + eg^{v_r s_r} l_{ij}^{v_r} + d_j^{v_r} + a_{ij}^{v_r} \right) + \quad (2.2b)$$

$$\sum_{(o,d) \in \mathbf{G}} \left(\tilde{q}_{od} O_{od} - q_{od} \sum_{r \in \mathbf{R}} V_{od}^r \right) + \quad (2.2c)$$

$$\sum_{r \in \mathbf{R}} \sum_{(h,i,j) \in \Omega_r} \left(u_i \left(W_{(hi)}^r + \sum_{\substack{d \in \mathbf{P} \\ d \neq i}} V_{id}^r \right) + t_i \sum_{\substack{d \in \mathbf{P} \\ d \neq i}} \sum_{\substack{s \in \mathbf{R} \\ s \neq r}} U_{(hi)d}^{rs} \right) \quad (2.2d)$$

subject to

$$X_{(hi)d}^r + V_{id}^r + \sum_{\substack{s \in \mathbf{R}: \\ (k,i,l) \in \Omega_s \\ s \neq r}} U_{(ki)d}^{sr} = X_{(ij)d}^r + \sum_{\substack{s \in \mathbf{R} \\ s \neq r}} U_{(hi)d}^{rs} \quad r \in \mathbf{R}, (h,i,j) \in \Omega_r, d \in \mathbf{P} \quad i \neq dh \neq d \quad (2.3)$$

$$X_{(ij)j}^r = W_{(ij)}^r \quad (i,j) \in \mathbf{E}_r \quad r \in \mathbf{R} \quad (2.4)$$

$$O_{od} + \sum_{r \in \mathbf{R}} V_{od}^r = k_{od} \quad o, d \in \mathbf{G} \quad (2.5)$$

$$\sum_{d \in \mathbf{P}} X_{ijd}^r \leq c^{v_r} \cdot m_r \cdot Y^r \quad r \in \mathbf{R} \quad (i,j) \in \mathbf{E}_r \quad (2.6)$$

$$\sum_{r \in \mathbf{R}: v_r=v} Y^r \leq z^v \quad v \in \mathbf{V} \quad (2.7)$$

$$X_{(ij)d}^r, U_{(ij)d}^{rs}, W_{id}^r \in \mathbb{R}^+ \quad r, s \in \mathbf{R}, r \neq s \quad (i,j) \in \mathbf{E}_r \quad d \in \mathbf{P} \quad (2.8)$$

$$O_{od}, V_{od}^r \in \mathbb{R}^+ \quad r \in \mathbf{R} \quad o, d \in \mathbf{G} \quad (2.9)$$

$$Y^r \in \mathbb{Z}^+ \quad r \in \mathbf{R} \quad (2.10)$$

Expression (2.2a) in the objective function captures the daily running costs incurred by vessels in operation, as well as costs or revenues obtained when excluding some vessels from the operating fleet. The next term (2.2b) obtains the fuel costs (including consumption during steaming and while idling at ports), canal cost, and port calling fees for all active vessels. The term (2.2c) obtains the total revenues from transporting cargo as requested, and the penalties incurred when rejecting carriage requests. Finally, the term (2.2d) computes charges from loading and unloading containers, both at their origin and destination, and at transshipment points.

Constraints (2.3) balance the flow of containers loaded and unloaded from each rotation at nodes other than their final destination. Every container not destined for the port in question must either continue along in the same rotation or be unloaded for transshipment to a different rotation. Constraints (2.4) represent the flow of containers that were accepted for carriage arriving at their final destination. Every container arriving at its final destination is unloaded and grounded.

Constraints (2.5) tally demand from the originating hinterland. The variables O_{od} allow us to obtain a feasible solution for any fleet configuration. When there is not a sufficient number of vessels to transport all the demand, or when it is not economically convenient to transport all containers, the variables O_{od} are set to a positive value, indicating that some demand has been forfeited.

Constraints (2.6) impose restrictions on the total number of containers that can be transported by each edge in a rotation. The total that can be transported is limited by the number of vessels assigned to the rotation, the capacity of those vessels, and the number of trips the vessels perform over the planning horizon. Constraints (2.7) ensure that the number of vessels that are deployed

does not exceed the number of available vessels of each type. Finally, constraints (2.8), (2.9), and (2.10) impose non-negativity and integrality restrictions on the corresponding variables.

Summary of simplifications in the model

In this section we summarize some of the practical constraints and business rules stated in Section 2.4 that are *not* considered in the presented reference model for LSNDP. A subset of these constraints are accounted for in the data of the benchmark suite, while others are considered out of scope for mathematical models on the LSNDP. The reference model does not consider a maximal *Transit time* for each commodity, and transit times are supplied for each commodity, for future model development. Repositioning of empty containers are not dealt with, as real world networks will usually require the flow of containers to balance. *Non-linear handling costs* in ports is not considered in the reference model or the LINER-LIB 2012 data. The *port productivity* and *pilot times* for berthing are not considered in the fixed port stay of the reference model. The reference model considers a single container type and thus reefer capacity and the *bunker consumption of reefer containers* are not included in the reference model. Lastly, we do not consider *equipment cost*, *embargo* and *cabotage rules*.

2.6.2 Metaheuristic

The heuristic that coordinates the overall algorithm is based on a heuristic column generation framework (not to confuse with column generation[17] in LP) where an auxiliary problem is used to generate new rotations (columns) to the overall model (2.2)-(2.10) and a tabu search is applied to select the most promising set of rotations to constitute the network. At any iteration u , we solve a multicommodity flow problem (MCFP) with aggregated origins using the then current set of rotations, \mathbf{R}^u . The MCFP is defined in the $X_{(ij)d}^r$ variables for the edges of the rotation, the $U_{(hi)d}^{rs}$ variables to account for transshipments between rotations with edge cost t_i and $W_{(id)}^r, V_{id}^r$ variables for (un)load to ports at their origin or destination with edge costs u_i . Rejected demand (O_{od}) is modeled by adding an auxiliary edge (o, d) with cost \tilde{q}_{od} for each commodity. In order to explain the transformation to the MCFP we transform the current rotations and their transshipment, (un)loading possibilities into a graph $G = (N, A)$. We define the set of port terminal vertices N_T for every port $p \in P$ and we define a set of port call vertices for every rotation as N_C^r . For every vertex we define a function $p(u)$ that returns the port $p \in P$ represented by vertex u . The unified set of vertices is defined as $N = N_T \cup N_C^r$. The set of edges consist of load edges, A_l , (defined in the variables V_{id}^r with cost u_i and capacity c^{v_r}) connecting $w \in N_T$ to $v \in N_C^r$, where $p(w) = p(v)$. Vice versa unload edges, A_u , (defined in the variables $w_{(ij)}^r$ with cost u_j and capacity c^{v_r}) connect $w \in N_C^r$ to $v \in N_T$, where $p(w) = p(v)$. The set of transshipment edges, A_t , (defined in the variables $U_{(hi)d}^{rs}$) connect $w \in N_C^r$ to $v \in N_C^s$, where $p(w) = p(v)$ (for butterfly routes $r = s$), with cost t_i and capacity $\min\{c^{v_r}, c^{v_s}\}$. Define the set of sailing edges, A_s , (represented by the variables $X_{(ij)d}^r$), with zero cost (as the cost is entailed in the cost of the routing), and capacity c^{v_r} connecting every port pair $w, v \in N_C^r$, where w is the port call preceding v in the rotation defined by the solution. Finally, we represent rejecting demand by omission edges, A_o (defined in the variables O_{od}) is defined for each commodity $g \in G$ with cost \tilde{q}_{od} , and capacity k_{od} . Define edge set $A = A_u \cup A_l \cup A_t \cup A_s \cup A_o$ and $G = (N, A)$ is a capacitated, directed graph for which we can solve a multicommodity flow problem defined with the quantities and revenues of the commodities in the set G from the original problem. The revenue is linked to the V_{od}^r variables for each commodity and the objective function minimizes the cost of the edges subtracting the revenue of each commodity. The graph is illustrated in Figure 2.7.

The heuristic column generation scheme applies a tabu search on a set of rotations iteratively generated by a MIP neighbourhood and is inspired by the approach presented in [5]. The algorithm presented in this paper has a more advanced MIP neighbourhood for generating rotations, which includes more heuristic measures on cargo composition and is designed to generate simple and butterfly routes of high quality. The MIP used for route generation includes (bi)-weekly frequency constraints and respects business rules 1 and 2 from Section 2.4.2.

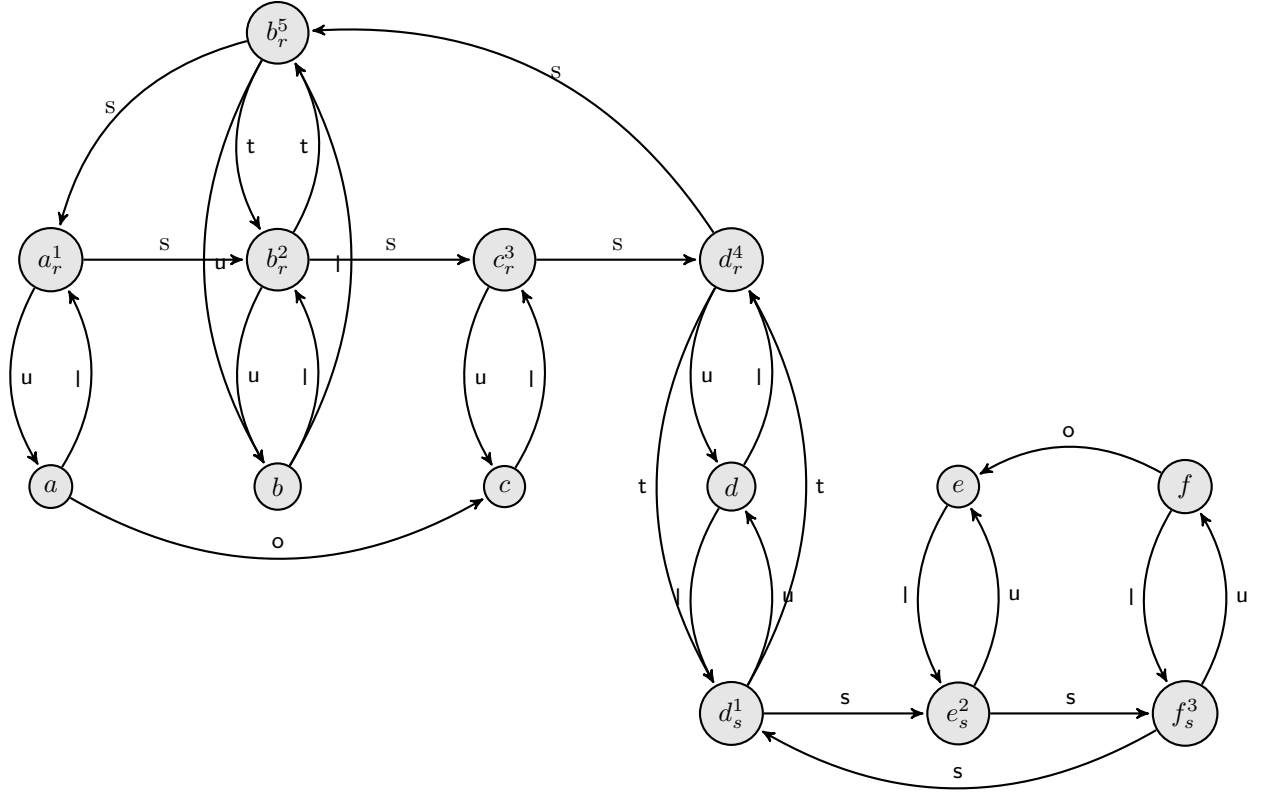


Figure 2.7: An illustration of the multi commodity flow graph of a given solution with separation of terminal vertices and port call vertices for each rotation. Terminal vertices is denoted by a letter designating a given port, whereas a port call vertex is designated by the port it represents, the route it is on and the call number in the rotation. The graph is connected by (un)load edges (indicated by letters u and l), transshipment edges (t), sailing edges (s) and the penalty for rejecting cargo is modelled in the omission edges (o) between port pairs.

The traffic demand that is not assigned to any of the routes in \mathbf{R}^u is the residual traffic demand for the following iteration, \hat{k}^{u+1} . For $u = 0$ there are no rotations in \mathbf{R}^u , and $\hat{k}_{od}^0 = k_{od}$.

Let us assume that, for each vessel family v , a number n_v of vessels are available at the end of iteration u . For each family of vessels where $n_v > 0$, we launch a collection of sub-problems $\text{AUX}(v, s, \kappa)$, for $s \in [s_{min}^F, s_{min}^F + 1, \dots, s_{max}^F]$ and various values of the number of vessels deployed, κ . In practice, this might give rise to a fairly large number of sub-problems. Therefore, during the initial iterations, we focus on a narrow subset of speeds and number of vessels to be deployed.

In solving the sub-problems $\text{AUX}(v, s, \kappa)$, we will typically obtain multiple feasible integer solutions. All these solutions are added to the overall model (2.2)-(2.10), since in any case the sub-problem contains some approximations, and the optimal solution to the $\text{AUX}(v, s, \kappa)$ may not be the solution that advances the overall heuristic the most. Our current implementation parallelizes the sub-problems, and we obtain batches of sixteen sub-problems that run simultaneously.

After the heuristic column generation for iteration u has terminated, we proceed to evaluate each new column ζ_j in conjunction with \mathbf{R}^u . We set $\mathbf{R}_j^{u+1} = \mathbf{R}^u \cup \{\zeta_j\}$ and run the multicommodity flow algorithm on this network. We select from amongst the \mathbf{R}_j^{u+1} the network with the best objective function, and set \mathbf{R}^{u+1} to be this network.

At some iterations, none of the candidate networks will result in a globally improved objective function. We permit the algorithm to exploit this neighborhood for a limited number of iterations. If the overall objective function does not improve after a predefined number of iterations, we trigger a backtracking process. First, we return to the best known configuration \mathbf{R}^* . Then we

delete rotations from \mathbf{R}^* based on three criteria: empty FFE-miles, underutilized legs, and mixed integer programming gap when the column was generated. The number of rotations that are deleted depends on how many times we have returned to this same \mathbf{R}^* – a larger number of returns to \mathbf{R}^* possibly indicates that this is no longer a promising point of departure, and the algorithm will remove a larger subset of rotations from it. The set of routes that results after backtracking is used as the basis for the following iteration.

If at any iteration the heuristic column generation procedure fails to produce any new columns (for instance, because no vessels are available) we delete a small number of rotations from \mathbf{R}^u . In order to reduce cycling, we add recently added rotations to the tabu list to avoid it being deleted from \mathbf{R}^u .

2.6.3 Route Generation

The auxiliary problem $AUX(v, s, \kappa)$ generates rotations based on the current network. We allow one butterfly node per rotation. We also designate one port as the "master" port of the service for ease of writing the subtour elimination constraints.

The following additional parameters are required for the formulation of the auxiliary problem.

\hat{k}_{od}	Residual demand (in FFE) to the liner company for transport from o to d . At any iteration in the heuristic, the residual demand is computed by taking the original demand, and subtracting the flow that is carried by existing rotations.
T	Length of the planning horizon, in days.
δ	Empirical parameter, estimates the amount of additional flow flowing through a butterfly node, as compared to a regular node.
$\phi_n^{in}, \phi_n^{out}$	Empirical parameters that capture the importance of a port as an exporter or importer.
κ	Number of sister vessels on the new service.
v	Vessel type to be deployed.
s	Speed of all vessels that will be deployed.

Problem $AUX(v, s, \kappa)$ has the following decision variables:

N_j	Binary variable, indicates whether port j is visited in the rotation.
B_j	Binary variable, indicates whether port j is a butterfly port in the rotation.
I_j	Continuous variable, used to indicate the sequence of port j in a rotation (for subtour elimination).
C_j	Binary variable, indicates if port j is the master port of the route.
A_{ij}	Binary variable, indicates whether edge (i, j) forms part of the new rotation.
Q_{od}	Continuous variable, indicates the number of FFEs with origin at port o and final destination at port d that will be carried per sailing of each vessel in the rotation.
W_1, W_2	Binary variables, respectively indicating whether the new rotation will have weekly or bi-weekly call frequency.
μ	<i>Inverse</i> of the number of trips to be completed over the entire planning horizon by each vessel on the new service.
ω	Estimated cost per sailing, per vessel, of the new service.

The auxiliary column generation model is then given by:

$$AUX(v, s, \kappa) \quad \text{maximize} \quad Z_{AUX(v, s, \kappa)} = \omega \quad (2.11)$$

subject to

$$\omega = (\tilde{f}^v - f^v)\mu - \sum_{(i,j) \in \mathbf{E}} (eh^v p_j^v + eg^{vs} l_{ij}^v + d_j^v + a_{ij}^v) A_{ij} + \sum_{(o,d) \in \mathbf{G}} (q_{od} - u_o - u_d + \tilde{q}_{od}) Q_{od} \quad (2.12)$$

$$24T\mu = \sum_{(i,j) \in \mathbf{E}} \left(p_j^v + \frac{l_{ij}^v}{s} A_{ij} \right) \quad (2.13)$$

The objective function (2.11) maximizes the net revenue contribution from each sailing of the additional rotation. Constraint (2.12) defines the contribution from each sailing by prorating the daily running cost of the vessel, the opportunity/layup cost of the vessel, sailing and hotel fuel expenses, port calling fees, canal fees, and the revenues generated by the sailing. Constraint (2.13) establishes the number of trips to be completed by each vessel over the decision horizon as a function of sailing time between ports and port stay duration.

The following constraints establish the basic logic for variables A_{ij} , N_i , and B_i .

$$\sum_{j \in \mathbf{P}} A_{jn} = \sum_{j \in \mathbf{P}} A_{nj} \quad n \in \mathbf{P} \quad (2.14)$$

$$A_{ij} \leq (N_i + N_j)/2 \quad (i, j) \in \mathbf{E} \quad (2.15)$$

$$\sum_{j \in \mathbf{P}} B_j \leq 1 \quad (2.16)$$

$$\sum_{i \in \mathbf{P}} A_{ij} \leq N_j + B_j \quad j \in \mathbf{P} \quad (2.17)$$

$$\sum_{j \in \mathbf{P}} A_{ij} \leq N_i + B_i \quad i \in \mathbf{P} \quad (2.18)$$

Constraints (2.14) balance the number of arcs entering and leaving any port in the rotation. Constraints (2.15) allow an arc (i, j) to be part of the rotation only if ports i and j are part of the rotation. The number of butterfly nodes in a rotation is limited to one by constraints (2.16). Constraints (2.17) and (2.18) limit the number of arcs that can enter or leave a port. If a port is a butterfly port, two arcs may enter and leave the port. Otherwise, a single entry and departure are permitted for each sailing.

In order to establish an approximate balance between the cargo that will flow through the new rotation and the capacity of the vessels deployed to the rotation, we have the following constraints:

$$Q_{od} \leq \mu \hat{k}_{od} / \kappa \quad (o, d) \in \mathbf{G} \quad (2.19)$$

$$Q_{od} \leq \hat{k}_{od} N_o \quad (o, d) \in \mathbf{G} \quad (2.20)$$

$$Q_{od} \leq \hat{k}_{od} N_d \quad (o, d) \in \mathbf{G} \quad (2.21)$$

$$\sum_{d \in \mathbf{P}} Q_{od} \leq \phi_o^{out} c^v (N_o + \delta B_o) \quad o \in \mathbf{P} \quad (2.22)$$

$$\sum_{o \in \mathbf{P}} Q_{od} \leq \phi_d^{in} c^v (N_d + \delta B_d) \quad d \in \mathbf{P} \quad (2.23)$$

$$\sum_{(o,d) \in \mathbf{G}} l_{od}^v Q_{od} \leq \sum_{(i,j) \in \mathbf{E}} A_{ij} l_{ij}^v c^v \quad (2.24)$$

Constraints (2.19) limit the amount of cargo that can be carried between any o-d pair by each vessel on each sailing, by pro-rating the total remaining demand amongst all vessels sailing in the rotation, and the number of trips to be performed by each vessel. Constraints (2.20) and

(2.21) require ports o and d to be active in the rotation whenever any part of the residual demand for the corresponding o-d pair is carried by the candidate rotation. Our formulation does not represent the amount of cargo on board the vessels on every leg of the rotation. Rather, we use two approximations to balance the amount of cargo that is targeted for the proposed rotation against the capacity and number of vessels being deployed there. First, constraints (2.22) and (2.23) ensure that the amount of cargo loaded, respectively discharged, at any port d is less than a certain fraction ϕ_d^{out} , respectively ϕ_d^{in} , of the vessel capacity. The factors ϕ_d^{out} and ϕ_d^{in} capture the importance of each port as an importer or exporter of cargo relative to the overall residual demand for transport in the network. For the second approximation, we generate a lower bound (because we use the direct sailing distance for all o-d pairs) for the FFE-miles of cargo being transported. For the second approximation, we estimate the FFE-miles that are required by cargo being transported as well as the maximum FFE-miles that can be provided by the new rotation. The left-hand side of constraints (2.24) represents a (rather weak) approximation on the FFE-miles that will be consumed by the cargo being carried. The approximation is weak because we use the direct distance between origin and destination port l_{od}^v , whereas the cargo will more often travel indirectly, visiting several ports before reaching its destination. The right-hand side of constraints (2.24), however, provides an accurate measure of the FFE-miles that will be provided by the new rotation.

The subtour elimination constraints are:

$$\sum_{j \in \mathbf{P}} C_j = 1 \quad (2.25)$$

$$B_j \leq C_j \leq N_j \quad j \in \mathbf{P} \quad (2.26)$$

$$N_j \leq I_j \leq |\mathbf{P}|N_j \quad j \in \mathbf{P} \quad (2.27)$$

$$1 + I_i - |\mathbf{P}|C_j - |\mathbf{P}|(1 - A_{ij}) \leq I_j \quad (i, j) \in \mathbf{E} \quad (2.28)$$

Constraint (2.25) identifies exactly one port within the rotation as the base port. If the rotation has a butterfly node, constraint (2.26) requires that it must be at the rotation's base port. Constraints (2.27) forces variables I_j away from zero if and only if the corresponding port is to be visited by the rotation. Constraints (2.28) force the value of sequence number I_j to increase along the path of the rotation, except when returning back at the rotation's base port.

In order to ensure that all rotations have weekly or bi-weekly frequency we have the constraints:

$$W_1 + W_2 = 1 \quad (2.29)$$

$$W_2 = 0 \quad c^v \geq 1200\text{FFE} \quad (2.30)$$

$$(W_1 - 1) + \frac{0.91 \cdot 7\kappa}{T} \leq \mu \leq \frac{7\kappa}{T} + (1 - W_1) \quad (2.31)$$

$$(W_2 - 1) + \frac{0.91 \cdot 14\kappa}{T} \leq \mu \leq \frac{14\kappa}{T} + (1 - W_2) \quad (2.32)$$

Constraint (2.30) reflects highly competitive conditions in markets where larger vessels are deployed. This constraint requires that rotations employing vessels with capacity at or above 1200 FFE must have weekly calling frequency. It would be unwise to require rotation frequencies to be exactly seven or fourteen days. Given that the model uses discrete vessel speeds, a strict implementation of the (bi)-weekly frequency requirement would likely discard many routes that are of commercial value, but that may not meet the target frequency by a few hours. Additionally, many carriers build some schedule slack into their routes to address fluctuations in the weather or delays at one of the ports in the rotation. With this in mind, we write constraints (2.31) so that rotations with nominal weekly frequency might in fact call as often as every 6.3 days. Similarly constraints (2.32) permit that rotations with nominal bi-weekly frequency might call as often as every 12.7 days.

Finally we have the integrality and non-negativity constraints:

$$A_{ij} \in \{0, 1\} \quad (i, j) \in \mathbf{E} \quad (2.33)$$

$$N_j, B_j, C_j \in \{0, 1\} \quad j \in \mathbf{P} \quad (2.34)$$

$$W_1, W_2 \in \{0, 1\} \quad (2.35)$$

$$Q_{od} \geq 0 \quad (o, d) \in \mathbf{G} \quad (2.36)$$

$$I_j \geq 0 \quad j \in \mathbf{P} \quad (2.37)$$

$$\mu, \omega \geq 0 \quad (2.38)$$

Our practical experience with problems of type $AUX(v, s, \kappa)$ is that these can be solved to optimality very quickly when the instance includes about 15 ports or less. Instances that entail more than 25 ports are significantly harder to solve, and are therefore not suitable for our overall strategy, where several thousand instances of $AUX(v, s, \kappa)$ may be launched. Our approach is to create clusters of ports that are tightly linked both geographically and by trade volumes. We sort such clusters according to the total demand for transport between the cluster's ports, and select the top cluster. We then formulate problems $AUX(v, s, \kappa)$ using up to 20 ports from the selected cluster. We believe that this approach results in a good compromise between the speed (allowing us to generate many rotations in the heuristic column generation) and solution quality (finding the best rotations).

We also note that the subtour elimination constraints (2.28) are an important source of the difficulty in solving problem $AUX(v, s, \kappa)$. In an effort to build up the network very quickly when the algorithm starts, we suppress constraints (2.28) for ten iterations. In many cases, we obtain solutions that do not contain subtours, allowing us to progress rapidly. As the network becomes more complex, however, it is necessary to reinstate constraints (2.28) in order to obtain valid solutions.

2.6.4 Complexity

In the following we will prove that the LSNDP is strongly NP hard by reduction from the Travelling Salesman Problem (TSP) in the general case and the set-covering problem for the case of the model (2.2)-(2.10). [1] proved their model to be weakly NP hard by reduction from the knapsack problem so the here presented proofs are stronger.

In the general case we may choose rotations arbitrarily. We show that LSNDP is strongly NP-hard by reduction from the TSP. The TSP may be defined as follows: Let $G = (N, A, C)$ be a graph where $N = \{1, \dots, n\}$ is the set of nodes, $A = \{(i, j) | i, j \in N, i \neq j\}$ is the set of edges and $C = c_{ij}$ is a cost or distance matrix associated to the edges A . An optimal solution to the TSP is a minimal cost Hamiltonian cycle covering the nodes in N .

Theorem 2.6.1. *The LSNDP in the general case is NP-hard*

Proof. We reduce from TSP. Let the set of ports P correspond to the set of nodes N in the TSP, and let the travel cost between ports correspond to the cost matrix C between nodes in the TSP. Moreover, set the demand between each pair of ports to 1, and limit the fleet list to one vessel of infinite capacity. The LSNDP will then choose the minimal cost Hamiltonian cycle between the ports P in order to satisfy all demands and is exactly an instance of the TSP. \square \square

If the rotations are given beforehand as in (2.2)-(2.10) it is easy to see that LSNDP is NP-hard by reduction from the set-covering problem. Given a number of nodes N and a family of sets S_1, \dots, S_ℓ , $S_i \subseteq N$ with corresponding costs $\bar{c}_1, \dots, \bar{c}_\ell$, the set covering problem asks to choose a subset of S_1, \dots, S_ℓ covering all nodes in N at the cheapest possible cost.

Theorem 2.6.2. *The LSNDP based on fixed rotations is NP-hard*

Proof. In order to reduce the set covering problem to LSNDP we let the set of ports be $\mathbf{P} = N \cup \{0\}$. For each subset S_i we introduce a rotation $R_i = S_i \cup \{0\}$ which visits the ports in arbitrary order.

The cost of a rotation is set to $\tilde{c}_i = \bar{c}_i$. There is one vessel type, with ℓ vessels each having capacity $|N|$. All transshipment costs are set to zero, and there is a demand of 1 between each port $i \in N$ and port 0. The revenue of delivering one unit is $q_{od} = 0$, while the penalty of not delivering one unit is $\tilde{q}_{od} = \infty$. The LSNDP will now choose the cheapest set of rotations covering all ports which is exactly the set covering problem. \square \square

2.7 Computational results

The model described in Section 2.6.1 has been solved heuristically for the LINER-LIB 2012 data using the algorithm described in Section 2.6.2. The tests were performed on an Intel Xeon E5345, 2.66 GhZ Quad core with 20 GB RAM. As LP solver we have used Gurobi 4.5. The running time of the algorithm varies due to the large difference in the size of the instances. The maximal running times have been set experimentally. As described in Section 2.5.6 on page 32 there are three scenarios of each case representing high, medium and low capacity related to the base case referred to as **Low, Base, and High**. In Table 2.7, column **t sec.** reports the running time per case and reports algorithmic performance for the Median run of the **Low, Base, and High** cases. Ten replications of each instance have been made. The Median is given as the fifth worst value of the ten runs. All figures and remaining tables display the best and median solution with regards to the objective function. Please note that a profit will be negative, as we are minimizing the objective value, but are in reality maximizing revenue.

Several parameters in the heuristic are model and algorithm dependent. The planning period is set to 180 days and all demands are scaled accordingly. A weekly or biweekly frequency has been enforced in (2.31)-(2.32), and the first and second business rules of Section 2.4.2 are imposed. All bunker costs are fixed at 600 USD per metric Ton. Port-draft incompatibilities are also considered in the model implementation.

Instance	t sec.	Columns evaluated			Unique columns			MCF Eval		
		Low	Base	High	Low	Base	High	Low	Base	High
Baltic	300	12912	12586	14046	1521	1719	1866	13680	13255	14809
WAF	900	13120	15343	12248	3858	4637	3951	13481	15343	12602
Mediterranean	1200	1582	2323	3192	1039	1197	1502	1634	2383	3244
Pacific	3600	4987	4203	4726	2552	2361	2660	5060	4268	4769
AsiaEurope	14400	4882	4291	4889	2322	2215	2806	4972	4357	4942
WorldSmall	10800	7313	7209	7628	2543	2937	3424	7425	7304	7731
WorldLarge	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 2.7: Performance of the algorithm and the search progress for the median run with regards to the objective value. **Instance** denotes the instance, **t sec.** is the running time in seconds. **Columns evaluated** is the number of rotations evaluated in the tabu search, **Unique columns** is the number of unique rotations and finally **MCF eval** is the number of times the multicommodity flow solution was calculated. Note that we do not solve WorldLarge due to its size.

Tables 2.7 and 2.8 report the algorithmic performance, Tables 2.9 and 2.10 report the objective values and cost components and finally some network key performance indicators are reported in Tables 2.11 and 2.12.

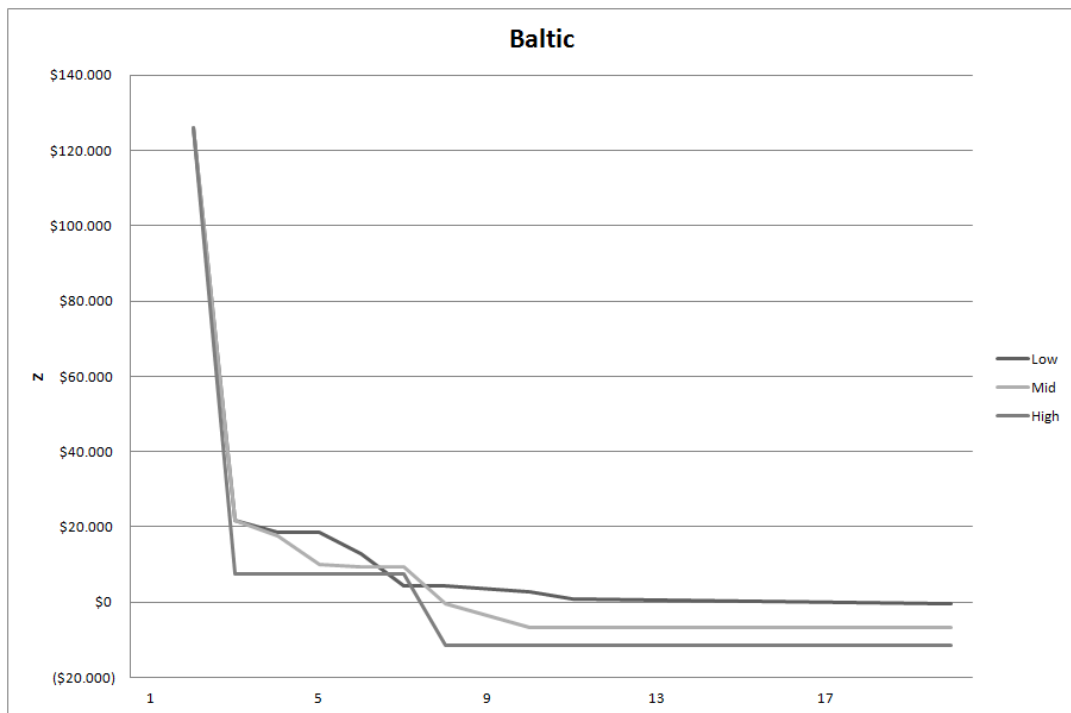
Figures 2.8(a) and 2.8(b) show the objective value for the median solution of the 10 randomized runs, as a function of the computational time, for the Baltic and AsiaEurope scenarios. For the remaining scenarios refer to Appendix 2.9.

The figures show that the objective value is converging for most cases; rapidly for the smaller instances while for the larger instances there are still iterative improvements in spite of the increased run times. The ratio between unique columns (rotations) and total evaluated columns (rotations) is seen to increase with scenario size. This is due to the rotation generating MIP easily finding new unique columns in the larger search space of the larger instances, as opposed to the

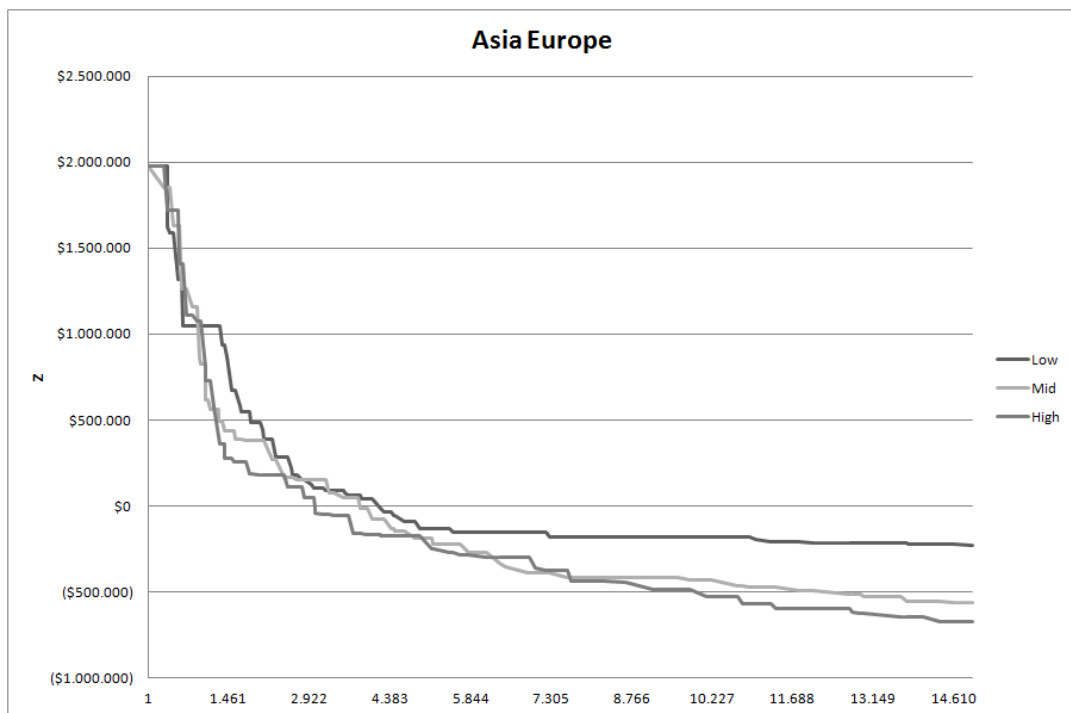
Instance	Iter			Improving Iter			Last Improving Iter		
	Low	Base	High	Low	Base	High	Low	Base	High
Baltic	548	358	459	12	12	9	151	30	12
WAF	267	337	271	19	30	26	115	195	221
Mediterranean	37	39	33	8	15	18	15	31	33
Pacific	57	54	33	31	26	30	57	30	33
AsiaEurope	102	80	87	64	60	58	102	80	85
WorldSmall	75	55	65	41	41	45	75	55	64
WorldLarge	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 2.8: Performance of the algorithm and the search progress for the median run with regards to the objective value. **Instance** denotes the instance **Iter** is the number of iterations, where an iteration refers to \mathbf{R}^u , **Improving Iter** is the number of iterations, where an improving solution was found, and finally, **Last Improving Iter** is the last iteration, where an improving solution was found.

smaller instances. The Mediterranean instance is an exception, where the ratio is rather high. It is believed to be due to a much denser demand matrix compared to the Baltic and WAF case. The number of iterations refers to \mathbf{R}^u of Section 2.6.2 and can be seen in Table 2.8. It can be seen that the number of iterations decreases with scenario size, even though computational time increases for the larger instances. It can be seen that the solution space grows significantly with instance size, as we find more improving solutions even in the last iterations.



(a) Performance for Baltic case. Objective value Z against the running time in seconds.



(b) Performance for AsiaEurope case. Objective value Z against the running time in seconds.

Figure 2.8: The objective value for the median Run Solution of the Low, Base and High instance of the Baltic and AsiaEurope case, as a function of the running time.

Instance			Z	Q	c_v	c_b	c_p	c_c	c_m	c_t	L_v	F
Baltic	Low	best	-6044	96750	7740	9522	8808	0	53780	0	0	10850
		median	838	92840	7740	9405	9509	0	51530	0	0	15490
	Base	best	-8365	98310	6480	10900	10440	0	55820	0	0	6306
		median	-6582	97860	6480	10208	10920	0	54760	62,49	0	8844
	High	best	-15678	103500	5760	11460	10520	0	59140	0	0	942,2
		median	-11509	101200	5760	9894	11670	0	57240	159,4	0	4954
WAF	Low	best	-115361	345800	57420	51750	15150	0	83130	3264	5940	25710
		median	-109470	336500	57420	43624	17270	0	79930	0	3960	32790
	Base	best	-143110	370600	52920	55450	20710	0	90760	3192	5760	10270
		median	-137681	355200	52920	44925	18940	0	84280	1690	8640	23360
	High	best	-159944	382000	48960	60520	16910	5827	94830	3859	10800	1956
		median	-152635	367800	48960	48756	19160	5827	88650	1376	11880	14330
Mediterranean	Low	best	29504	128900	27540	23650	18730	0	64030	10090	0	14360
		median	42636	121700	27540	27546	17320	0	62000	9694	0	20270
	Base	best	12209	136800	26640	22220	20190	0	69620	9475	1440	2311
		median	24934	134200	26640	23591	20610	0	68430	16630	1980	5187
	High	best	6606	137100	26100	22070	19330	0	69920	9861	5400	1807
		median	14267	136700	26100	25450	20260	0	70350	11720	4860	1904

Table 2.9: The objective value separated by cost and income components according to the model in section 2.6 on page 33. **Instance** denotes the name of the instance with a separate row for the **best** and **median** values over 10 replications. The **best** values correspond to the best solution obtained. Z is the objective value, note that we are minimizing expenses such that a negative value of Z is preferable, Q is the total revenue collected, c_v is the total vessel cost, c_b is the total fuel cost, c_p is the total port call cost, c_c are the canal fees, c_m is the total move cost at origin and destination nodes, c_t is the pure transshipment cost. L_v is the income from chartered out vessels and finally F is the sum of goodwill penalties for rejected cargoes. All costs are in k\$. All numbers are presented with a four digit precision. Z may deviate slightly from the sum of the remaining columns according to the objective due to this rounding.

Instance			Z	Q	c_v	c_b	c_p	c_c	c_m	c_t	L_v	F
Pacific	Low	best	84615	1120000	276300	342500	31580	3180	429700	34610	2700	89080
		median	132020	1044000	276300	300400	32460	2989	390100	31450	0	142700
	Base	best	-54087	1197000	247700	348900	31610	6101	451600	34870	13500	35440
		median	-12774	1160000	247700	351800	35210	2967	444400	29110	17460	53710
	High	best	-101671	1208000	236500	350500	38340	6071	456400	33100	36720	22360
		median	-75321	1222000	236500	388300	36000	6016	460400	36180	31140	14150
WorldSmall	Low	best	-551842	4917000	1040000	1019000	100100	255400	1211000	199800	106200	646000
		median	-417100	4902000	1040000	1024000	103100	237900	1220000	247000	48060	661300
	Base	best	-1152761	5366000	942100	1115000	123800	271000	1336000	195200	160000	390500
		median	-888669	5199000	942100	1109000	126300	298400	1296000	213400	164000	489000
	High	best	-1291718	5550000	911700	1202000	129300	287800	1377000	263600	223200	310400
		median	-1168067	5266000	911700	1077000	109400	267100	1319000	244000	261500	431800
AsiaEurope	Low	best	-361041	3109000	718600	672000	117200	233600	609100	128800	17640	286900
		median	-228186	2998000	718600	661100	106800	228800	584400	125800	0	344000
	Base	best	-657972	3283000	648700	690400	134400	255700	637600	132700	81180	206700
		median	-561417	3179000	648700	726100	108800	254800	615600	132500	118100	248700
	High	best	-766385	3390000	625700	834900	130000	276700	656000	127900	176600	149400
		median	-670099	3181000	625700	708900	117700	253900	622400	131400	184900	235300

Table 2.10: The objective value separated by cost and income components according to the model in section 2.6 on page 33. **Instance** denotes the name of the instance with a separate row for the **best** and **median** values over 10 replications. The **best** values correspond to the best solution obtained. Z is the objective value, note that we are minimizing expenses such that a negative value of Z is preferable, Q is the total revenue collected, c_v is the total vessel cost, c_b is the total fuel cost, c_p is the total port call cost, c_c are the canal fees, c_m is the total move cost at origin and destination nodes, c_t is the pure transshipment cost. L_v is the income from chartered out vessels and finally F is the sum of goodwill penalties for rejected cargoes. All costs are in k\$. All numbers are presented with a four digit precision. Z may deviate slightly from the sum of the remaining columns according to the objective due to this rounding.

Tables 2.9 and 2.10 state the objective value for the best and median case. The total objective value of the instance is given by Z and the remaining columns display separate costs for vessels, port calls, cargo handling, omission penalty and incomes from chartering out vessels and transporting cargo. The difference in the objective value between the best and the median case show a large variance in the resulting solution with regards to the objective value. It should be noted that the individual cost and income components of a large network are colossal seen over a planning horizon of 180 days. The large variance may be caused by the heuristic getting trapped in a local minimum, which it cannot escape. The solution space of liner shipping network design is vast and there are many individual components constituting a good solution. As a result fixing a set of rotations to a particular vessel class may result in a profitable and well utilized set of rotations, but a different composition of vessel class to the same rotations could result in a more profitable solution. However, the heuristic will regard the set of rotations to be of high quality and will thus be less likely to remove them and try a different vessel class composition.

Instance		dep%	$ R $	PCpW	BPU%	WPU%	BAU%	WAU%	t/d%	rej%	
Baltic	Low	Best	100,0	5	2,81	100,0	100,0	93,9	47,9	0,0	8,6
		Median	100,0	3	4,36	100,0	100,0	80,3	69,8	0,0	12,3
	Base	Best	100,0	4	4,06	100,0	100,0	75,9	43,2	0,0	5,0
		Median	100,0	5	2,96	100,0	100,0	92,9	48,2	0,3	7,0
	High	Best	100,0	5	3,48	100,0	100,0	98,6	56,6	0,0	0,7
		Median	100,0	4	4,12	100,0	100,0	83,6	61,5	2,7	3,9
WAF	Low	Best	89,4	7	5,19	100,0	95,1	90,4	50,4	7,0	11,7
		Median	92,9	8	4,04	100,0	100,0	79,4	49,7	13,5	14,9
	Base	Best	88,9	11	3,69	100,0	95,9	82,9	44,6	16,5	4,7
		Median	83,3	10	3,38	100,0	83,3	78,6	46,3	8,3	10,6
	High	Best	77,0	12	3,39	100,0	39,1	85,5	24,0	9,2	0,9
		Median	74,7	10	3,64	100,0	39,1	93,1	24,0	1,6	6,5
Mediterranean	Low	Best	100,0	7	5,97	100,0	100,0	94,4	70,1	56,8	7,4
		Median	100,0	7	5,57	100,0	100,0	95,3	76,9	50,4	10,4
	Base	Best	94,6	7	6,83	100,0	100,0	93,7	79,7	46,0	1,2
		Median	91,9	8	5,68	100,0	90,2	90,4	62,3	84,0	2,7
	High	Best	78,4	11	4,75	100,0	76,1	96,3	60,5	48,8	0,9
		Median	80,5	9	5,52	100,0	29,4	89,4	21,0	58,7	1,0

Table 2.11: Key Performance indicators for the smaller instances. **Instance** denotes the name of the instance with a separate row for the **best** and **median** values over 10 replications. The **best** values correspond to the best solution obtained. The **median** value is the median value of all KPI over the 10 replications. **dep%** is the percentage of the fleet deployed, $|R|$ is the number of rotations in the final solution, **PCpW** is the average number of port calls per service per week, **BPU%**, **WPU%** is the best and worst peak utilization percentage respectively, **BAU%**, **WAU%** is the best and worst average utilization percentage respectively. **t/d%** is the percentage of transshipments performed of all delivered units and **rej%** is the percentage of rejected cargo.

Tables 2.11 and 2.12 give some key performance indicators for the networks generated. The *utilization percentage* is the relation between cargo transported and the available capacity on a vessel for a single voyage between two ports. The *Peak utilization* of a rotation relates to the voyage between two ports with the highest utilization percentage. *Best Peak Utilization (BPU)* denotes the peak utilization of the rotation with the highest peak utilization, and *Worst Peak Utilization (WPU)* the rotation with the lowest peak utilization. *The Average Utilization* relates to the average utilization percentage of all voyages on a rotation. *Best Average Utilization (BAU)* describes the average utilization of the rotation with the highest average utilization, and *Worst Average Utilization (WAU)* the average utilization of the rotation with the lowest average utilization.

Instance			dep%	R	PCpW	BPU%	WPU%	BAU%	WAU%	t/d%	rej%
Pacific	Low	Best	99,0	22	5,41	100,0	87,4	100,0	50,0	42,3	7,8
		Median	100,0	21	5,09	100,0	32,8	100,0	18,8	41,7	12,6
	Base	Best	95,0	21	5,73	100,0	36,6	100,0	21,0	46,0	3,1
		Median	93,3	21	5,54	100,0	58,8	100,0	38,5	44,8	4,7
	High	Best	85,1	20	6,53	100,0	100,0	100,0	50,0	47,5	2,0
		Median	87,0	24	5,02	100,0	7,7	100,0	5,3	39,0	1,2
WorldSmall	Low	Best	90,1	35	5,60	100,0	100,0	100,0	50,9	64,7	19,6
		Median	95,5	42	4,04	100,0	48,7	100,0	42,7	74,7	20,0
	Base	Best	83,0	41	4,87	100,0	37,0	100,0	28,3	64,7	11,8
		Median	82,8	40	4,67	100,0	22,1	100,0	21,1	65,3	14,8
	High	Best	76,5	41	4,68	100,0	24,0	100,0	23,1	73,7	9,4
		Median	71,9	37	5,06	100,0	79,6	100,0	52,7	78,3	13,1
AsiaEurope	Low	Best	97,5	30	6,39	100,0	34,0	98,5	22,4	80,2	14,5
		Median	100,0	32	5,50	100,0	32,5	100,0	25,7	75,3	17,4
	Base	Best	87,0	35	6,59	100,0	34,9	96,0	32,2	91,1	10,4
		Median	81,1	38	5,48	100,0	22,9	100,0	16,0	81,6	12,6
	High	Best	71,0	37	6,29	100,0	11,1	100,0	7,5	81,5	7,6
		Median	70,6	34	6,28	100,0	33,4	100,0	23,7	82,9	11,9

Table 2.12: Key Performance indicators for larger instances. **Instance** denotes the name of the instance with a separate row for the **best** and **median** values over 10 replications. The **best** values correspond to the best solution obtained. The **median** value is the median value of all KPI over the 10 replications. **dep%** is the percentage of the fleet deployed, **|R|** is the number of rotations in the final solution, **PCpW** is the average number of port calls per service per week, **BPU%**, **WPU%** is the best and worst peak utilization percentage respectively, **BAU%**, **WAU%** is the best and worst average utilization percentage respectively. **t/d%** is the percentage of transshipments performed of all delivered units and **rej%** is the percentage of rejected cargo.

Fleet utilization and rejected demand The percentage of the fleet deployed, **dep%**, seen in relation to the percentage of rejected demand, **rej%**, reveals lack of capacity to carry all demand for the low capacity instances as expected. Likewise, there is excess capacity in most high capacity instances with a corresponding low rejection rate. AsiaEurope and WorldSmall High instances are an exception to this pattern, which may be due to the profitability of cargo, but as the large instances are continuously finding improving rotations, solutions with a higher deployment and a decrease in rejected demand may exist.

Number of rotations The number of rotations are seen to increase with instance size. WorldSmall has a larger number of rotations than AsiaEurope correlated to the number of average port-calls on a service, **PCpW** being lower in WorldSmall than in AsiaEurope.

Utilization All solutions have a **BPU** of 100 % meaning that at least one service is fully utilized on one voyage between two ports. Low **WPU**s are also seen and can be acceptable for feeder services for outlying profitable cargo. Looking at the average utilization, **BAU%**, **WAU%**, the networks are overall well utilized as some variation must be expected given the asymmetry of world trade and low utilization in some parts of the feeder network.

Transshipment percentage The percentage of number of transshipments over the demand units transported, **t/d%**, reveal that Baltic and WAF instances have very few transshipments and as expected, the Mediterranean case has a larger percentage of transshipments. For the Pacific, WorldSmall and AsiaEurope most demands are subject to one or more transshipments.

2.7.1 Discussion

Overall the solutions obtained are promising. The solutions utilize the available capacity well and the constructed networks transport the majority of the profitable cargoes, as can be seen in Tables 2.11 and 2.12.

Review by Network Planners Network planners at Maersk Line have evaluated that the given solutions are feasible liner shipping networks for the scenarios, but with room for improvement by inspection of experienced eyes. Some of these are dealt with by the incorporated business rules but other such as cabotage rules, commercially driven transit time restrictions, etc. is not dealt with. Realistic traits of the networks can be seen in the construction of shorter feeder services and longer inter continental services. Hubs are used realistically by connecting several feeders and inter continental services to the same hubs.

Baltic and WAF These small instances show very fast convergence in Figures 2.8 and 2.9. The generated networks consist of services with few port calls (Table 2.11) (except lowly capacitated Low WAF instances), as expected of feeder networks with few or no transshipments as seen in Table 2.11.

Mediterranean Has a dense demand matrix which facilitates some hub structure and transshipments. No runs generated a profitable solution but the high capacity cases are close, the best with a loss of 6606 k\$. A study of the revenues of the demands reveal a rather low revenue compared to e.g. the Baltic case and also the actual demand is spread on a large number of demand pairs meaning that the cargo compositions are more complex. The analysis indicates that the Mediterranean case is not very attractive from an economic standpoint.

Pacific and WorldSmall For these large instances complex networks are created with 20 to 40 services calling 5 to 6 ports on average. For Pacific most cargo is transported, except for the lowly capacitated instances, for the WorldSmall a little less is moved, but as vessels are chartered out it must be due to unprofitable demand, but still it is the most complex scenario and improvement are possible. The Pacific case has relatively few transshipments, due to the direct connections over the pacific and the high North American transshipment costs. WorldSmall has realistically, massive use of transshipments.

AsiaEurope The AsiaEurope case is a large challenge due to the size of the instance but also because the revenue is tight. From the utilization indicators (**BAU / WAU**) in Tables 2.11 and 2.12 it can be concluded that traffic flow is very intense in one direction displaying the trade imbalance, which is very decisive for network design today due to the excess capacity on the backhaul. This case is well suited for future work optimizing on empty repositioning. The amount of rejected cargo of the larger instances clearly indicate that some demand is in reality loss giving, a trait any real life case will have.

Cost Structure Looking at the median solution for the largest Base capacity scenario, WorldSmall, the network cost is distributed as 24 % for vessel costs, c_v , 28 % for fuel costs c_b , 10 % for port call and canal costs, $c_p + c_c$ and 38 % for move cost, $c_m + c_t$. Noting that this is a simplified case with demand aggregated to larger ports, reducing the needed number of transshipments and portcalls, this cost distribution can resemble the cost distribution of real liner shipping networks, although this of course fluctuates with cost and market changes.

Transshipments and Hubs The cases vary greatly in the hub-and-spoke structure and hence the desire to transship cargo. The Baltic case has very few transshipments as will be the case for a feeder network. In the AsiaEurope case all cargo transship more than once on average. Seen from a commercial point of view a transshipment carries a risk in terms of disruptions and an

administrative overhead in transshipped cargo. These perceived “costs” of transshipping is not considered in the mathematical model and the found solution may have excess transshipments compared to real life. Adding customer satisfaction costs to the transshipment costs in the model can to some extent account for this problem.

Service Types The solutions clearly display the capacity levels of a feeder network and an intercontinental network, where larger vessels are servicing intercontinental routes. Network planners at Maersk Line confirm that the structure of the networks in terms of capacity deployment and the use of transshipment locations is fairly coherent with the operations of Maersk Line. Also the rotations vary greatly in design and display all of the design structures seen in Figure 2.1 and 2.2, showing the algorithms adaptability for different situations.

Fleet Size The different fleet size instances Low, Base and High capacity, results in different types of networks. The Low capacity instances obviously have less services, which on average are better utilized, as they will fill the vessels with the most profitable cargo and reject more cargo. The high capacity instances will have to create additional services to make a profit of the remaining cargo. Thus the resulting networks of the related fleet size scenarios will result in quite different networks and adds to the complexity of LINER-LIB 2012.

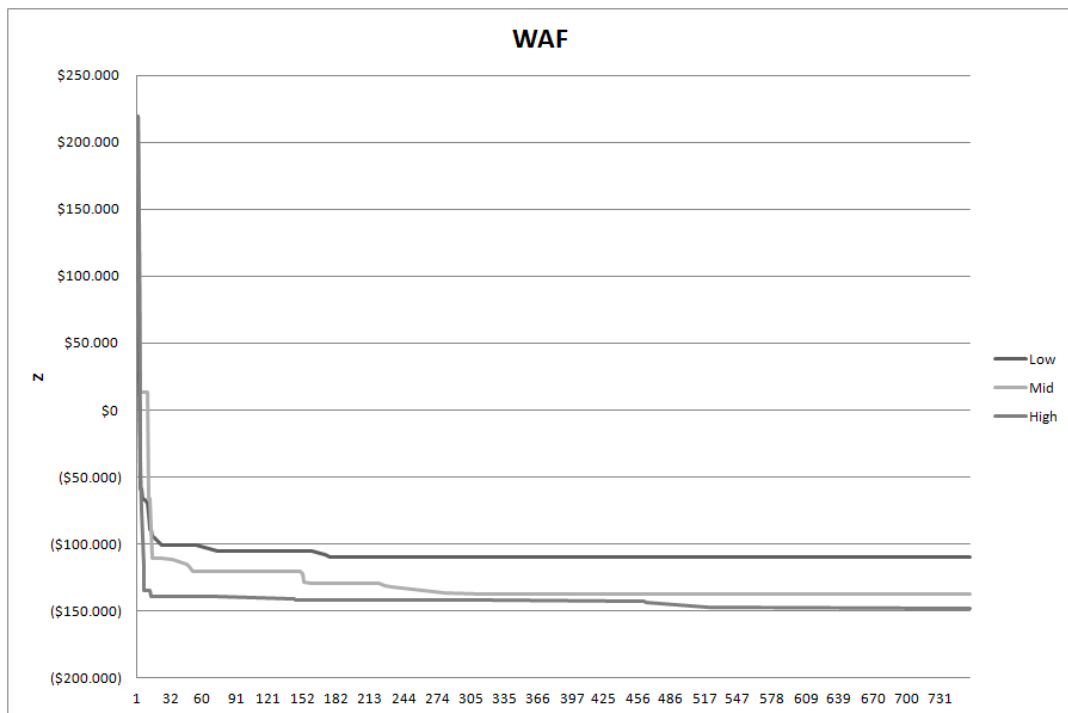
2.8 Conclusion and future work

In this paper we have given an introduction to Liner Shipping Network Design for operations researchers. The problem domain has been described and a discussion of mathematical modelling of the domain and the constraint set has been provided. A set of data instances was introduced as a benchmark suite, resembling real world problems. The problem has been proven to be strongly NP-hard and comparing with optimal methods for solving related Liner Shipping Network Design models, it is apparent that this problem is among the most difficult network design problems. The model of Alvarez [5] has been extended to handle the structure of a complex network, and a heuristic column generation algorithm has been presented using a new method for route generation. Computational results using LINER-LIB 2012 have been presented, and the solutions will be made available on a web-page in order to encourage a competition in developing the best algorithms. The algorithmic performance indicators suggest that we have created a benchmark suite that may be used for the development of both exact and heuristic approaches as the smaller cases should be suitable for benchmarking exact algorithms and the large instances poses a substantial challenge for state of the art heuristics. Overall, LINER-LIB 2012 provides a thorough test of a network design algorithm. The cases challenge the fleet deployment, port call sequence generation, the transshipment structure of the network and the selection of demand to transport in the network. At the same time the cases created closely resembles the real life operations of a global liner shipping company and the market for liner shipping in terms of trade imbalance, geographical transshipment points and fluctuations in demand compared to the capacity available. We also believe that LINER-LIB 2012 will prove valuable to future model development of the liner shipping network design problem. There is a large potential for future work on modelling as several commercial requirements such as low transit times and slow steaming have not been incorporated in the current network design models. The results have been evaluated by network designers of Maersk Line, who affirm that the traits of individual rotations resembles real world rotations, and that the overall network fulfill most of the relevant properties. It is also apparent by inspection, that the heuristic solutions may be improved upon, which encourages the development of better solution methods. We hope that this work will serve as a foundation for unifying and encouraging further work in Liner Shipping Network Design, which is a challenging real world problem having a huge impact on the cost of the world’s supply chains. Moreover, the scientific world needs complex challenges in order to push algorithm development forward.

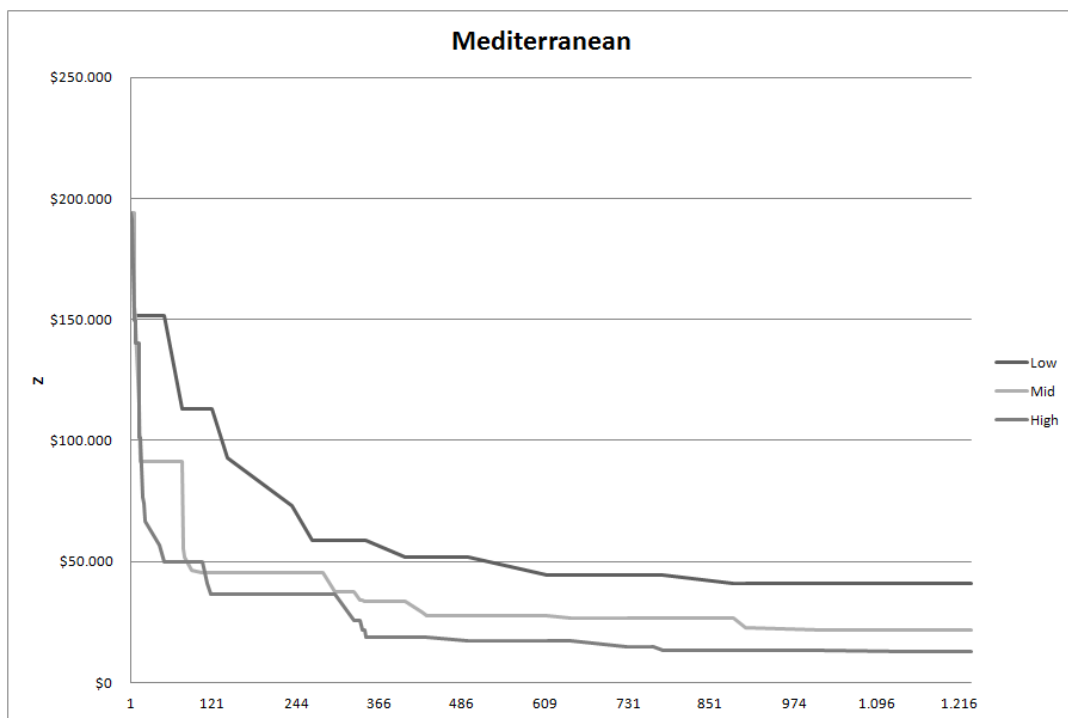
Acknowledgements

We would like to thank Maersk Line for valuable insights and access to real life data for the creation of LINER-LIB 2012. In particular we would like to thank Niels Madsen for valuable comments and suggestions for improvement. This article and the benchmark suite have been greatly improved by his constructive comments and insight into liner shipping network design. We also thank Jørgen Harling and Niels Rasmussen, Maersk Line. This project would not have been possible without their support. Thanks goes to Line Blander Reinhardt for sharing illustrations. We are grateful to Vereinigung Hamburger Schiffsmakler und Schiffsagenten [52], Drewry Shipping Consultants [19], Alphaliner charter rates 2000-2010 [4] and National Imagery and Mapping Agency [37] for providing data to construct the LINER-LIB 2012 data. We thank two anonymous referees, whose comments have improved the paper significantly. This project was supported in part by The Danish Strategical Research Council under the ENERPLAN project.

2.9 Appendix - Figures

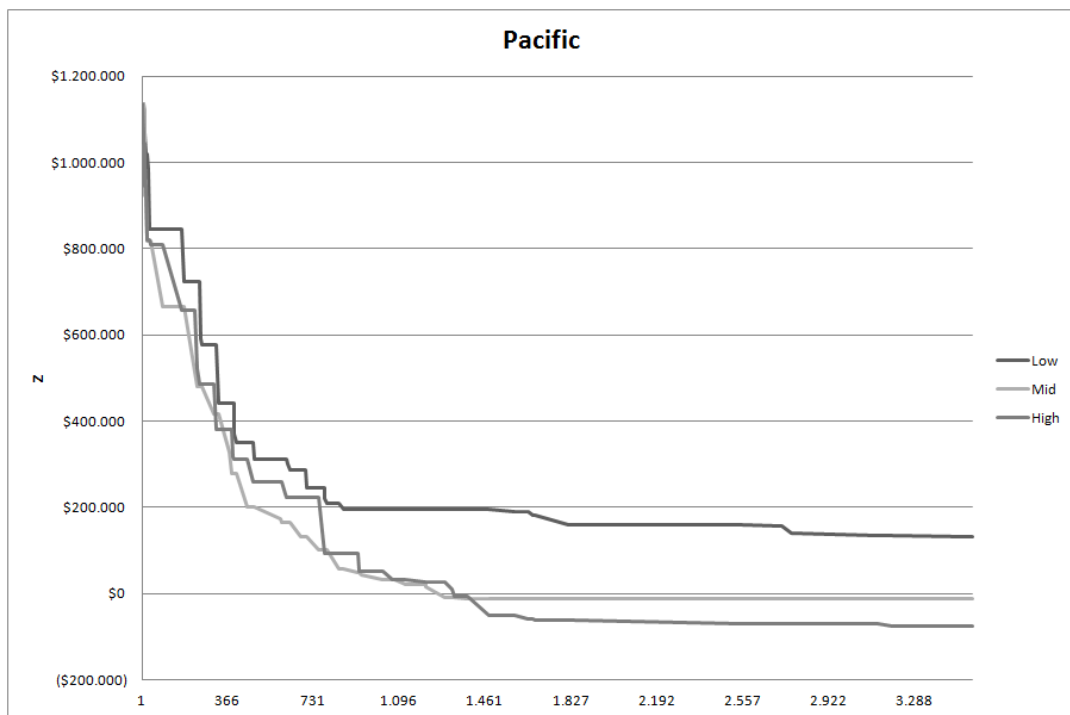


(a) Performance for WAF case

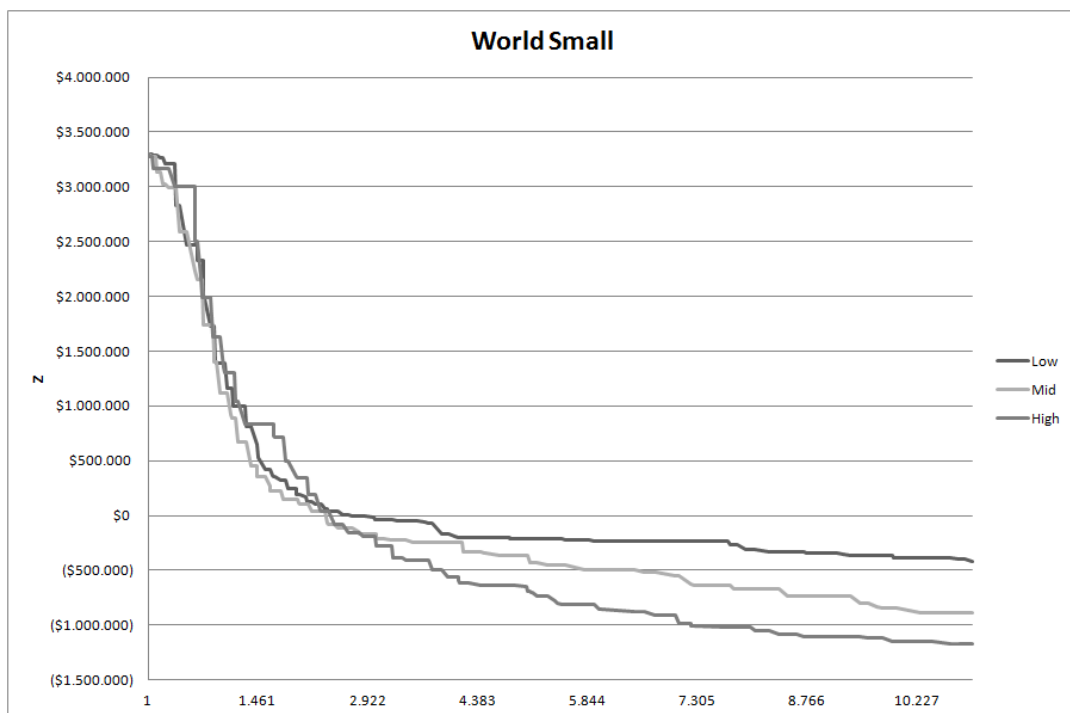


(b) Performance for Mediterranean case

Figure 2.9: The objective value for the median Run Solution of the Low, Base and High instance, as function of the running time in seconds.



(a) Performance for Pacific case



(b) Performance for WorldSmall case

Figure 2.10: The objective value for the median Run Solution of the Low, Base and High instance, as function of the running time in seconds.

Bibliography

- [1] R. Agarwal and O. Ergun. Ship scheduling and network design for cargo routing in liner shipping. *Transportation Science*, 42(2):175–196, 2008.
- [2] R. Agarwal and O. Ergun. Mechanism design for a multicommodity flow game in service network alliances. *Operations Research Letters*, 36(5):520 – 524, 2008.
- [3] P. Alderton. *Reeds Sea Transport — Operations and Economics*. Motorbooks Intl, 2005. ISBN 0713669446.
- [4] Alphaliner charter rates 2000-2010. Alphaliner. <http://www.alphaliner.com>, 2010. Last viewing January 2011.
- [5] J. F. Alvarez. Joint routing and deployment of a fleet of container vessels. *Maritime Economics & Logistics*, 11(2):186–208, 2009.
- [6] J. F. Alvarez. Mathematical expressions for the transit time of merchandise through a liner shipping network. *Journal of the Operational Research Society*, 63(6):709–714, 2012.
- [7] J.F. Alvarez, P. Tsilingiris, E. Engebretsen, and N.M.P. Kakalis. Robust fleet sizing and deployment for industrial and independent bulk ocean shipping companies. *INFOR*, 49(2): 93–107.
- [8] A.P. Møller-Mærsk Fleet List. 2008_apmm_fleetlist.pdf. <http://about.maersk.com>, 2008. Last viewing June 2009.
- [9] L. Blander Reinhardt and D. Pisinger. A branch and cut algorithm for the container shipping network design problem. *Flexible Services and Manufacturing Journal*, 24(3):349–374, 2012.
- [10] B.D. Brouer, D. Pisinger, and S. Spoorendonk. The cargo allocation problem with empty repositioning (caper). *INFOR*, 49(2):109–124, 2011.
- [11] P. Cariou. Is slow steaming a sustainable means of reducing co2 emissions from container shipping? http://www.c4tx.org/ctx/pub/liner_co2.pdf, 2010.
- [12] M. Christiansen and K. Fagerholt. Some thoughts on research directions for the future: Introduction to the special issue in maritime transportation. *INFOR*, 49(2):75–77, 2011.
- [13] M. Christiansen, K. Fagerholt, and D. Ronen. Ship routing and scheduling: Status and perspectives. *Transportation Science*, 38(1):1–18, 2004.
- [14] M. Christiansen, K. Fagerholt, B. Nygreen, and D. Ronen. Maritime transportation. In C. Barnhart and G Laporte, editors, *Handbooks in Operations Research and Management Sciences*, pages 189–284. North Holland, 14 edition, 2007.
- [15] C. W Chu, T. C Kuo, and J. C Shieh. A mixed integer programming model for routing containerships. *Journal of Marine Science and Technology*, 11(2):96–103, 2003.
- [16] The World Shipping Council. The liner shipping industry and carbon emissions policy, 2009, Washington D.C.
- [17] J. Desrosiers and M.E. Lübbecke. A primer in column generation. In G. Desaulniers, J. Desrosiers, and M.M. Solomon, editors, *Column Generation*, pages 1–32. Springer US, 2005.
- [18] J.-X. Dong and D. Song. Container fleet sizing and empty repositioning in liner shipping systems. *Transportation Research Part E: Logistics and Transportation Review*, 45(6):860–877, 2009.

- [19] Drewry Shipping Consultants. Container freight rate insight. http://www.drewry.co.uk/publications/view_publication.php?id=358, 2010. Last viewing January 2011.
- [20] K. Fagerholt. Designing optimal routes in a liner shipping problem. *Maritime Policy & Management*, 31(4):259 – 268, 2004.
- [21] K. Fagerholt, T.A.V. Johnsen, and H. Lindstad. Fleet deployment in liner shipping: A case study. *Maritime Policy and Management*, 36(5):397–409, 2009.
- [22] United Nations Economic Commission for Europe. <http://www.unece.org/cefact/locode/>. Last viewing January 2011.
- [23] S. Gelareh and Q. Meng. A novel modeling approach for the fleet deployment problem within a short-term planning horizon. *Transportation Research Part E: Logistics and Transportation Review*, 46(1):76–89, 2010.
- [24] S. Gelareh, S. Nickel, and D. Pisinger. Liner shipping hub network design in a competitive environment. *Transportation Research Part E: Logistics and Transportation Review*, 46(6): 991–1004, 2010.
- [25] Hapag-Lloyd. Freight rate quotes. <http://www.hapag-lloyd.com/en/schedules/interactive.html>, 2011. Last viewing January 2011.
- [26] C. I Hsu and Y. P Hsieh. Routing, ship size, and sailing frequency decision-making for a maritime hub-and-spoke container network. *Mathematical and Computer Modelling*, 45(7-8): 899–916, 2007.
- [27] A. Imai, K. Shintani, and S. Papadimitriou. Multi-port versus hub-and-spoke port calls by containerships. *Transportation Research Part E: Logistics and Transportation Review*, 45: 740–757, 2009.
- [28] M. K. Jepsen, B. Løfstedt, C. E. M. Plum, D. Pisinger, and M. M. Sigurd. A path based model for a green liner shipping network design problem. In *IMECS 2011 - International MultiConference of Engineers and Computer Scientists 2011*, volume 2, pages 1379–1384, 2011.
- [29] M.G. Karlaftis, K. Kepaptsoglou, and E. Sambracos. Containership routing with time deadlines and simultaneous deliveries and pick-ups. *Transportation Research Part E*, 45:210–221, 2009.
- [30] K. H. Kjeldsen. Classification of ship routing and scheduling problems in liner shipping. *INFOR*, 49(2):139–152, 2011.
- [31] K. H. Kjeldsen. *Routing and Scheduling in Liner Shipping*. Ph.d. thesis, Department of Economics and Business - Aarhus University, 2012.
- [32] B. Løfstedt, J.F. Alvarez, C.E.M Plum, D. Pisinger, and M.M. Sigurd. An integer programming model and benchmark suite for liner shipping network design. Technical Report 19, DTU Management Engineering, 2011.
- [33] H. A Lu. Modelling ship’s routing bounded by the cycle time for marine liner. *Journal of Marine Science and Technology*, 10(1):61–67, 2002.
- [34] Maersk Line. Maersk line still most reliable container carrier. http://www.maerskline.com/link/?page=news&path=/news/story_page/11/container_carrier, 2011. February, 15.

- [35] Q. Meng and S. Wang. Liner shipping service network design with empty container repositioning. *Transportation Research Part E: Logistics and Transportation Review*, 47(5):695–708, 2011.
- [36] Q. Meng and S. Wang. Liner ship fleet deployment with week-dependent container shipment demand. *European Journal of Operational Research*, 222(2):241–252, 2012.
- [37] National Imagery and Mapping Agency. Distances between ports - pub. 151. <http://www.marcon.com/library/resources/DBP/DBP.html>, 2001.
- [38] Network and Product at Maersk Line. Personal communication, 2011.
- [39] T.E. Notteboom. A carrier’s perspective on container network configuration at sea and on land. *Journal of International Trade and Logistics*, 2(1):65–87, 2004.
- [40] T.E. Notteboom. The time factor in liner shipping services. *Maritime Economics & logistics*, 8:19–39, 2006.
- [41] T.E. Notteboom and B. Vernimmen. The effect of high fuel costs on liner service configuration in container shipping. *Journal of Transport Geography*, 17(5):325–337, 2009.
- [42] Panama Canal. The panama canal. <http://www.pancanal.com/eng/maritime/tolls.html>. Last viewing January 2011.
- [43] Christian E. M. Plum and Peter Neergaard Jensen. Minimization of bunker costs. Master’s thesis, 2007.
- [44] B. J. Powell and A. N. Perakis. Fleet deployment optimization for liner shipping: an integer programming model. *Maritime Policy and Management*, 24(2):183–192, 1997.
- [45] K. Rana and R.G. Vickson. Routing container ships using lagrangian relaxation and decomposition. *Transportation Science*, 25(3):201–214, 1991.
- [46] E. Sambracos, J. A. Paravantis, C. D. Tarantilis, and C. T. Kiranoudis. Dispatching of small containers via coastal freight liners: The case of the aegean sea. *European Journal of Operational Research*, 152(2):365–381, 2004.
- [47] K. Shintani, A. Imai, E. Nishimura, and S. Papadimitriou. The container shipping network design problem with empty container repositioning. *Transportation Research Part E: Logistics and Transportation Review*, 43(1):39–59, 2007.
- [48] M.M. Solomon. Algorithms for the vehicle routing and scheduling problems with time window constraints. *Operations Research*, 35(2):254–265, 1987.
- [49] M. Stopford. *Maritime Economics*. Routledge, third edition, 2009. ISBN 0203442660.
- [50] Suez. Suez canal authority. <http://www.suezcanal.gov.eg/calc.aspx>. Last viewing January 2011.
- [51] The World Bank. Container port traffic. <http://data.worldbank.org/indicator/IS.SHP.GOOD.TU/countries?disps> 2010. last viewing May 2011.
- [52] Vereinigung Hamburger Schiffsmakler und Schiffsagenten. The hamburg shipbrokers’ association. <http://www.vhss.de/context.php>. Last viewing January 2011.
- [53] S. Wang and Q. Meng. Schedule design and container routing in liner shipping. *Transportation Research Record*, 2222(1):25–33, 2011.
- [54] S. Wang and Q. Meng. Sailing speed optimization for container ships in a liner shipping network. *Transportation Research Part E: Logistics and Transportation Review*, 48(3):701 – 714, 2012.

-
- [55] S. Wang and Q. Meng. Liner ship route schedule design with sea contingency time and port time uncertainty. *Transportation Research Part B: Methodological*, 46(5):615 – 633, 2012.
- [56] S. Wang and Q. Meng. Liner ship fleet deployment with container transshipment operations. *Transportation Research Part E: Logistics and Transportation Review*, 48(2):470–484, 2012.
- [57] S. Yan, C. Y Chen, and S. C Lin. Ship scheduling and container shipment planning for liners in short-term operations. *Journal of Marine Science and Technology*, 14(4):417–435, 2009.

Part II

Liner Shipping Network Design

Chapter 3

The Single Service Design Problem in Liner Shipping

Christian E. M. Plum^{*, ‡} David Pisinger^{*} Juan-José Salazar-González[†] Mikkel M. Sigurd[‡]

^{*}Department of Management Engineering, Technical University of Denmark,
Produktionstorvet, Building 426, DK-2800 Kgs. Lyngby, Denmark
cemp@man.dtu.dk, pisinger@man.dtu.dk

[†]DEIOC, Facultad de Matemáticas
Universidad de La Laguna, 38721 Tenerife, Spain
jjsalaza@ull.es

[‡]Maersk Line
Esplanaden 50, DK-1098 København K.
mikkel.muhldorff.sigurd@maersk.com

1

Abstract The design of container shipping networks is an important logistics problem, involving assets and operational costs in billions of dollars. To guide the optimal deployment of the ships, a single vessel round trip is considered by minimizing operational costs and flowing the best paying demand under commercially driven constraints. This paper introduces the Single Service Design Problem. Arc-flow and path-flow models are presented using state-of-the-art elements from the wide literature on pickup and delivery problems. A Branch-and-Cut-and-Price algorithm is proposed, and implementation details are discussed. The algorithm can solve instances with up to 25 ports to optimality - a very promising result as real-world vessel roundtrips seldom involve more than 20 ports.

Keywords Traveling salesman problem, Liner shipping, Branch-and-Cut-and-Price, Shortest path, Network design, Green logistics.

3.1 Introduction

Container shipping carriers operate worldwide networks consisting of hundreds of vessels having huge operating costs. Developing methods that can improve the network costs and/or the service level are of huge importance for both the carriers and the customers. Note that most of the market today is based on manufactured products transported on container vessels from distant continents.

¹Submitted for *Computers and Operations Research*

Container shipping networks provide transport of containers from port to port at a fixed (usually weekly) schedule with a predetermined trip duration. The networks consist of a number of *services* and a set of similarly sized vessels sailing on a cyclic itinerary of ports. Services meet at certain hub ports where transshipment of containers can take place. The round trip duration is assumed to be a multiple of a week, and a sufficient number of vessels is assigned to the round trip to ensure a weekly visit to each port. For instance, Figure 3.1 shows a 6 week round trip with 6 vessels to ensure that each port is visited once a week.

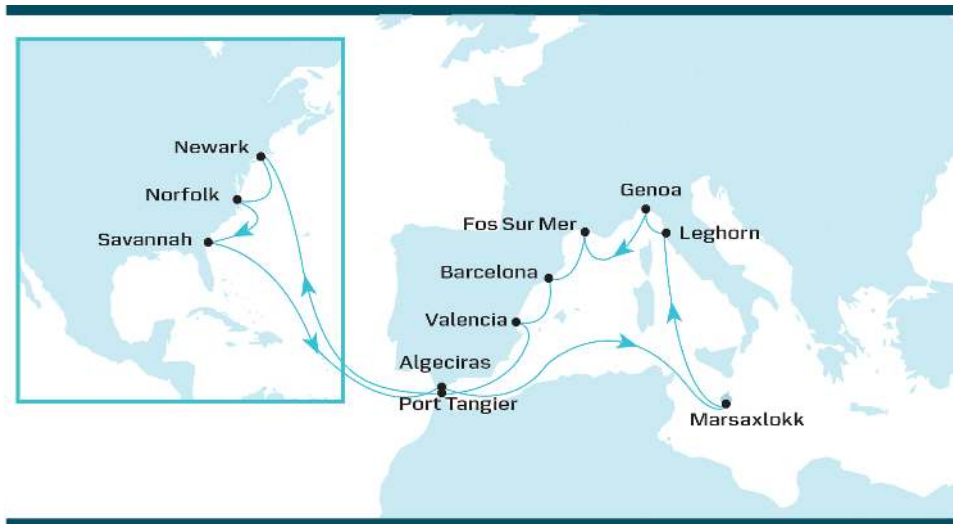


Figure 3.1: The WestMed Service, transporting containers between U.S. east coast and the western Mediterranean.

A given demand is loaded at its origin port to some service, which may bring the demand directly to its destination or unload it at a hub port for transshipment to another service, ultimately bringing the demand to its destination. See Stopford [25] and Notteboom [19] for a more general introduction to the economics of liner shipping.

A usual intercontinental service has between 5-10 port calls for the more direct trades (e.g. Trans-Atlantic or Trans-Pacific) and 15-20 port calls on the longer trades (e.g. Europe-Asia trades), indicating the problem sizes that can be encountered in reality. Stopford [25] has more details on different service types.

The problem investigated in this paper considers the design of a single capacitated service following a simple cyclic rotation where all ports must be visited, i.e. a Hamiltonian tour. A solution approach for this problem is an important tool for a network planner designing a single service as fierce competition between carriers often require low path durations, while the best paying containers must be prioritized to optimize profits. In practice services are seldom Hamiltonian, partly because important ports are called more than once, partly because waterways as canals must be traversed in both directions. An experienced user knows the ports where several visits may be necessary and, by duplicating them, the problem becomes the Hamiltonian variant addressed in this paper. Canals do not cater to demand and hence should be excluded from the port set, but included in the distances between ports.

The problem is then to transport a set of demands on a generated round trip, where the combined sum of these demands must consider the capacity of all edges. A demand has a maximal path duration which must be respected: a demand can be partly fulfilled, but it must still respect the path duration limit. This problem is called the *Single Service Design Problem*, or in short SSDP. To the best of our knowledge this problem has not been addressed before in the literature.

3.1.1 Liner shipping

We refer to Christiansen et al. [7] and Christiansen et al. [8] for an overview of early research on Liner Shipping problems. Since these reviews, a number of articles has been published, with various approaches and scopes for Liner Shipping Network Design Problems (LSNDP). The work of Shintani et al. [24] has a detailed description of the cost structure and includes consideration of repositioning empty containers. The network design problem considered by Agarwal and Ergun [1] generates multiple services and handle transshipment. Bender's and column generation based algorithms are implemented. These algorithms scale well to large instances, but transshipment costs are excluded. The model of Alvarez [2] considers transshipment cost and finds solutions for large instances in a heuristical column generation approach. The Branch-and-Cut method of Reinhardt and Pisinger [23] is the first model considering transshipment while allowing for non-simple rotations (with two calls to a single port, a so-called butterfly route). Small instances are solved optimally. The models of Gelareh et al. [11] and Gelareh and Pisinger [10] use a hub location based approach, generating a main service visiting some ports directly, instances of up to 10 ports are solved to optimality. The work of Brouer et al. [6] describes the domain of LSNDP, discusses the relevant scoping, proposes a model of the problem, and presents a number of benchmark instances for the LSNDP based on real world problems. A novel aggregation of demands was presented in Jepsen et al. [15] giving a new model formulation and decomposition method, though it did not perform well in practice. A heuristic algorithm for a short horizon version of the problem is presented by Wang and Meng [26]. A formulation considering empty container repositioning is found in Meng and Wang [16] and a further model dealing with robust schedule design in Wang and Meng [27], but neither of these consider the order of the port calls and take this as an input. A recent overview to the area is given by Meng et al. [17]. This multitude of publications on LSNDP shows that the interest in these problems has increased. Most of these works considers different models and scopes of the problem and optimal methods can only solve small instances (10–15 ports) and, as real world instances are larger, the problem is still open for research (see Brouer et al. [6]).

3.1.2 Pickup and Delivery problems

The SSDP is related to the well-studied pickup and delivery problems. Parragh et al. [20] and Berbeglia et al. [4] give good introductions to these problems, reviewing existing literature and proposing classification schemes. In the classification of Parragh et al. [20] the SSDP is a Single Dial-A-Ride Problem (SDARP) excluding Time Windows, and with the important difference that no depot is required, i.e. demand can be carried through the depot. In the classification of Berbeglia et al. [4] the SSDP is a **[1-1—PD—1]**: **1-1** as each commodity has one origin and one destination, **PD** as each vertex must be visited exactly once for combined pickup and delivery, and **1** as a single service is generated. An important difference from related problems is the lack of a depot. The multi-commodity one-to-one pickup-and-delivery traveling salesman problem (m -PDTSP) is considered in Hernández-Pérez and Salazar-González [12]: the problem is formulated, and solution methods based on Bender's decomposition are implemented. The SSDP can be seen as an extension of the m -PDTSP with the addition of path duration, and optional demand with associated revenue. An often encountered type of subproblems in pickup and delivery problems are Shortest Path Problems with Resource Constraints (SPPRC) which also appear by decomposing the SSDP. We refer to Irnich and Desaulniers [13], Jepsen et al. [14] or Petersen [21] for a review on SPPRC problems and algorithms.

3.1.3 Overview

The main contributions of this paper are two novel models of the SSDP and a Branch-and-Cut-and-Price solution method for solving the problem. The absence of a depot gives a problem structure not seen in related problems. This requires both the pricing problem and the separation of valid inequalities to be designed in a novel manner. The implemented method solves problem

sizes met in real world instances.

In Section 3.2 an arc-flow model of the SSDP is presented. This model is Dantzig-Wolfe decomposed to a path-flow model to be solved with a column generation algorithm, which effectively handles the multi commodity flow problem with path duration constraints. Details of subtour elimination constraints, pricing problems and branching approaches are given. The proposed algorithm has been implemented and computational results are presented in Section 3.3, where instances of up to 25 nodes can be solved to optimality. Details of the data instances are provided. Finally Section 3.4 concludes on the paper and proposes directions for further research. This work is an extension of Plum et al. [22].

3.2 Mathematical Formulation

In the following we introduce the notation, present an arc-flow model, followed by a path-flow model, and a new solution method for the SSDP.

The service must visit each node $i \in V$ exactly once. Directed arcs $(i, j) \in A$ exist between all nodes, giving the complete directed graph $G = (V, A)$. Let $S \subset V$ be a subset of nodes. Each arc $a = (i, j) \in A$ is associated with a cost c_a representing time charter costs for the vessel, bunker cost for propulsion and port call costs for visiting the port j . Traversing the arc a takes the time t_a . This time depends on the sailed distance and the speed of the vessel. The service has a capacity Q , which must be respected at all traversed arcs. The generated service can transport the commodities $k \in K$. Each commodity k has a source s^k and a destination d^k ($s^k, d^k \in V$), a volume of containers $F^k > 0$, a maximal path duration $t^k > 0$ and unit-revenue for transporting $r^k > 0$. A node i can be the source of one or more commodities, as well as destination for some commodities.

3.2.1 Arc-Flow Formulation

The problem is to find a maximal profit set of paths in G for a set of commodities k , such that the containers can be moved from their origin to their destination in at most t^k time. All the paths should be a subset of a Hamiltonian tour, where each arc has a corresponding cost and traversal time.

Let x_a be a binary variable indicating whether the service travels on arc $a \in A$. Let f_a^k be the flow of commodity k on each arc a . The problem can then be formulated as:

$$\min \sum_{a \in A} c_a x_a - \sum_{k \in K} r^k \sum_{a \in \delta^-(s^k)} f_a^k \quad (3.1)$$

subject to

$$x(\delta^-(i)) = 1 \quad \forall i \in V \quad (3.2)$$

$$x(\delta^+(i)) = 1 \quad \forall i \in V \quad (3.3)$$

$$x(\delta^+(S)) \geq 1 \quad \forall S \subset V \quad (3.4)$$

$$\sum_{a \in \delta^+(i)} f_a^k = \sum_{a \in \delta^-(i)} f_a^k \quad \forall k \in K \text{ and } i \in V \setminus \{s^k, d^k\} \quad (3.5)$$

$$\sum_{k \in K} f_a^k \leq Qx_a \quad \forall a \in A \quad (3.6)$$

$$\sum_{a \in \delta^-(s^k)} f_a^k \leq F^k \quad \forall k \in K \quad (3.7)$$

$$\sum_{a \in A} t_a f_a^k \leq t^k \sum_{a \in \delta^-(s^k)} f_a^k \quad \forall k \in K \quad (3.8)$$

$$f_a^k \geq 0 \quad \forall a \in A \text{ and } k \in K \quad (3.9)$$

$$x_a \in \{0, 1\} \quad \forall a \in A \quad (3.10)$$

The objective minimizes the cost of the traversed arcs subtracted the revenue of flowed demand. Constraints (3.2) and (3.3) ensure that all nodes have one outgoing and ingoing open edge, where $\delta^+(i)$ and $\delta^-(i)$ denotes the set of ingoing, respectively outgoing, arcs to node i . Constraint (3.4) are subtour elimination constraints ensuring that the Hamiltonian tour connects all nodes in a single rotation. The conservation of flow is ensured by (3.5), and the capacity is enforced by (3.6). Constraints (3.7) limit the served demand to the upper value F^k . The path duration is ensured by (3.8) since f_a^k is $\sum_{a' \in \delta^-(s^k)} f_{a'}^k$ for all a in the path moving the demand k . Constraints (3.9) and (3.10) set bounds on the decision variables f_a^k and x_a .

3.2.2 Path-Flow Formulation

Constraints (3.5)–(3.9) can be eliminated through a Dantzig-Wolfe decomposition on the arc-flow model, thus replacing the variables f_a^k by path variables. Let P^k be the set of all feasible paths from s^k to d^k , satisfying the constraints (3.5)–(3.9). This set may have an exponential number of elements. Each path $p \in P^k$ is represented as a set of arcs, i.e. $p \subset A$. Let $t_p = \sum_{a \in p} t_a$ be the duration of this path. Let λ_p be a non negative real variable representing the volume of flow of commodity k using path $p \in P^k$. The SSDP can then be formulated as:

$$\min \sum_{a \in A} c_a x_a - \sum_{k \in K} r^k \sum_{p \in P^k} \lambda_p \quad (3.11)$$

subject to (3.2)–(3.4) and

$$\sum_{k \in K} \sum_{p \in P^k: a \in p} \lambda_p \leq Qx_a \quad \forall a \in A \quad (3.12)$$

$$\sum_{p \in P^k} \lambda_p \leq F^k \quad \forall k \in K \quad (3.13)$$

$$x_a \in \{0, 1\} \quad \forall a \in A \quad (3.14)$$

$$\lambda_p \geq 0 \quad \forall k \in K \quad (3.15)$$

The objective function minimizes the costs of chosen arcs subtracted the revenue of flowed demand. The capacity is enforced by constraint (3.12). Convexity constraints (3.13) ensure that at most the available flow is transported.

The exponential number of subtour elimination constraints (3.4) can be relaxed initially and inserted when violated, as done in Reinhardt and Pisinger [23]. A lower bound on the optimal value of this model can be attained by solving the LP-relaxation, where the integrality constraints (3.14) are replaced with constraints $0 \leq x_a \leq 1 \forall a \in A$. This LP-relaxation can be solved using a cut-and-price algorithm. Due to the exponential number of variables λ_p , a restricted master problem is obtained by considering a subset $\bar{P} \subseteq P$ of paths. Additional columns of negative reduced costs are generated by solving a pricing subproblem. Let $\pi_a \in \mathbb{R}$ be the dual variables for the capacity constraints (3.12) and let $\theta^k \leq 0$ be dual variables for the convexity constraints (3.13). Then the pricing problem becomes:

$$\text{Min: } \sum_{a \in A} \pi_a x_a - \theta^k - r^k \quad (3.16)$$

subject to constraints (3.5)–(3.8).

3.2.3 Separation of Subtour Elimination Constraints

Given a solution x^* of the LP relaxation of the path-flow formulation, we must search for any violated subtour elimination constraint (3.4). A violated constraint exists if and only if a minimum-capacity cut in the solution graph $G(x^*)$ has weight less than one. This can be computed in polynomial time.

3.2.4 Pricing Problem

The pricing problem is an Elementary Shortest Path Problem with Resource Constraints (ESP-PRC), which is strongly NP-hard as shown in e.g. Irnich and Desaulniers [13]. The path must have the lowest cost given by arc weights π_a , while respecting path durations. Without the elementarity requirement, the problem can be solved in pseudo-polynomial time. As the pricing problem (3.16) may contain negative coefficients in the objective, negative cost cycles are likely, but still we apply the non-elementary variant. Negative cycles are broken by introducing a new resource for the number of traversed arcs, by imposing an upper bound of $n - 1$ on this resource. The problem is solved by a labeling algorithm, which has the advantage of returning all Pareto optimal paths, in the resources. All these paths are added to the pricing problem, if they have negative reduced costs. When no negative reduced costs can be found for the problem including the resource on number of traversed arcs, then no path with negative reduced costs for the ESPPRC exists.

3.2.5 Branching

When all violated cuts and negative reduced costs columns have been added to the current node, and fractional binary variables x_a still exists, branching is commenced. Binary branching is used, by selecting the most fractional x_a and adding constraints $x_a \leq 0$, $x_a \geq 1$ to the two branching children, respectively. As this branching is done on variables x_a existing in both the original and reformulated problem space, the branching constraints will be directly imposed in the pricing problem and subtour elimination cuts and no further consideration of this is needed.

A main contribution of this paper lies in the powerful formulation of the flow and path based models for this new problem. These formulations allows for carrying demand through the depot, while selecting which demand to flow and enforcing the path duration limit. The effectiveness of these formulations allows for efficient algorithmic techniques, as branching in the original space of the x_a variables, separating subtour elimination constraints and solving the pricing problem with an efficient labeling algorithm.

3.3 Computational Results

The algorithm has been implemented using the COIN-OR DIP (Galati and Ralphs [9]) framework to implement the Branch-and-Cut-and-Price method and using CPLEX 12.1 as LP solver. Boost's

graph library (Boost [5]) has an implementation of SPPRC, which is used to solve the pricing problem as described above. **Concorde** (Applegate et al. [3]) has an efficient implementation of a min cut algorithm and **boost** also finds connected components, to check if we have a feasible solution. The implementation has been run on a 4 GB Ram, Intel E8400 3.00 GHz using a single core.

3.3.1 Instances

The algorithm has been tested on a set of instances inspired by the *class2* and *class3* instances of Hernández-Pérez and Salazar-González [12], which again are based on the description by Mosheiov [18] for instances of TSP with pickup-and-delivery. These have n random points in the square $[-500, 500] \times [-500, 500]$, one of which is located at point (0,0) (formerly the depot). For each problem size, we have 5 randomly generated instances. The travel cost c_a is the Euclidean distance between the points and the travel time is $t_a = c_a/100 \cdot (0.95 + rand(0, 0.1))$. Hence the travel time is proportional to the cost, but with some small deviation, as seen in real-life problems. The vessel has a capacity Q . The commodities have a path duration limit t^k , a volume F^k and an associated revenue r^k . To test the scalability and properties of the algorithm a number of variations of the instances have been created.

Path duration limit Instances with $t^k \in \{0, 5, 10, 200\}$ have been created. $t^k = 10$ is used as the default setting. All t^k have the same setting in an instance.

Revenue Instances with $r^k \in \{0, 250, 1000, 10000\}$ have been created, $r^k = 1000$ is used as the default setting. All r^k have the same setting in an instance.

Graph Instances with 10, 15, 20 and 25 nodes have been run, all with a complete set of edges. These graph sizes resembles real service design problems.

Commodity Density Instances with fixed (F) number of commodities 5, 10 and 15 have been generated. To test larger commodity sets, instances with (A) sparse commodity density n , (B) populated commodity density $3n$, (C) dense commodity density $(n^2 - n)/2$ and (D) complete commodity density $n^2 - n$ have been tested. Commodity Density (B) is used if nothing else is mentioned.

Capacity Instances with $Q \in \{0, 10, 30, 200\}$ have been created, $Q = 10$ is used as the default setting. These instances are constructed as to resemble problems that could arise in real service design situations by the relation between capacity, revenue and path duration limit, as the interplay between revenue, cost and operational and commercial restrictions come in play.

3.3.2 Results

Tables 3.1-3.4 shows the results of the developed algorithm on the test instances. Each row in the table corresponds to runs on five randomized instances with the same properties. In the tables column n is the number of nodes, m is the number of commodities, $m = 3n$, commodity density (B), if nothing else is stated. The algorithm has been run with a time limit of 3600 seconds, and *Time* is the average computational time of the five runs. *Timeout* is the number of runs which timed out and hence was not solved to optimality. *Gap* is the percentual gap between the upper and lower bound of all computed instances including both optimally solved and timeouts. The number of added subtour elimination constraints is given by *Cuts*. The number of generated path columns are given by *Columns* and the number of search nodes in the Branch-and-Bound tree is given by *B&B Nodes*. All values are averages over the 5 randomly generated instances.

In general it can be seen that the solution time increases with graph size. Some classes of instances are easily solved as they reduce to a standard TSP problem, these are the cases where $t^k = 0$, $r^k = 0$ and $Q = 0$.

Table 3.1 reports the effect of different commodity densities. It shows that the small 5-commodity instances are easily solved, but interestingly the completely dense D -instances are also among the fastest solved for all n . This must be due to the abundance of commodities, making it easier to find the optimal solution, as commodities that fits together without exceeding path duration limit exists.

In Table 3.2, problem complexity can be seen to increase with increasing t^k . The number of $B\&B$ Nodes decreases for the larger instances with high t^k values, as each node becomes harder to solve with non binding path duration constraints.

In Table 3.3 problem complexity also increases with revenue, but not for the largest graphs, where the gap stabilizes or decreases for very large revenues, as the problem shifts from balancing cost against revenues, to maximizing the revenues.

The dependence on the vessel's capacity can be seen in Table 3.4, where the complexity of the instance appears to be proportional with the capacity. The largest instances that can be solved to optimality have 25 nodes and commodity density A or D . These problem types represent real world problems well and it proves the methods applicability in a real world setting as a decision support tool to generate services in a complex operational and commercial setting.

Generation of subtour elimination constraints (*Cuts*) increases with the size of the graph, for the instances solved to optimality. For instances reaching the time limit the time for each iteration increases with graph size, and thus *decreases* the number of subtour elimination constraints generated. The number of path columns generated (*Columns*) increases with increasing *Path duration limit* as more paths become feasible. There is also some dependence with the commodity density and capacity. The number of $B\&B$ Nodes follows the same pattern as the number of *Cuts*, i.e. increasing with larger graph size. As the time of each iteration increases the number of branches decreases due to timeout.

n	Commodity Density	m	Time	Timeout	Gap	Cuts	Columns	B&B Nodes
10	F	5	1	0	0 %	42	124	20
10	F	10	2	0	0 %	79	467	48
10	F	15	3	0	0 %	99	872	58
10	A	10	160	0	0 %	3936	1141	5510
10	B	30	18	0	0 %	291	3988	203
10	C	45	19	0	0 %	300	3511	191
10	D	90	6	0	0 %	130	4197	48
15	F	5	3	0	0 %	111	506	74
15	F	10	885	0	0 %	4465	6336	5519
15	F	15	292	0	0 %	1881	7518	1961
15	A	15	3	0	0 %	50	1225	12
15	B	45	1034	0	0 %	2278	33037	2035
15	C	105	3602	5	7 %	3327	62769	1448
15	D	210	117	0	0 %	484	29509	201
20	F	5	28	0	0 %	246	1454	175
20	F	10	1237	1	4 %	5000	7557	5585
20	F	15	3017	4	5 %	5213	22944	3735
20	A	20	3610	5	inf	601	41362	216
20	B	60	3605	5	12 %	1093	72048	415
20	C	190	3612	5	8 %	893	54188	220
20	D	380	620	0	0 %	702	54569	269
25	F	5	562	0	0 %	2275	7489	2324
25	F	10	3185	4	10 %	2394	35609	1904
25	F	15	3607	5	14 %	1109	46511	497
25	A	25	1713	1	1 %	848	150721	510
25	B	75	3619	5	39 %	457	52158	120
25	C	300	3613	5	12 %	644	79878	112
25	D	600	849	0	0 %	441	45677	189

Table 3.1: Computational results with varying commodity density. Average values of 5 instances.

n	t^k	Time	Timeout	Gap	Cuts	Columns	B&B Nodes
10	0	0	0	0 %	7	30	3
10	5	1	0	0 %	16	101	6
10	10	18	0	0 %	291	3988	203
10	200	3072	3	1 %	6331	162744	4969
15	0	1	0	0 %	11	45	8
15	5	7	0	0 %	124	298	109
15	10	1034	0	0 %	2278	33037	2035
15	200	3604	5	13 %	1445	163834	546
20	0	2	0	0 %	5	60	1
20	5	55	0	0 %	372	856	349
20	10	3605	5	12 %	1093	72048	415
20	200	3614	5	29 %	548	63474	114
25	0	2	0	0 %	1	75	1
25	5	496	0	0 %	2215	3541	2100
25	10	3619	5	39 %	457	52158	120
25	200	3614	5	25 %	529	104782	105

Table 3.2: Computational results with varying path duration limits. Average values of 5 instances.

n	r^k	Time	Timeout	Gap	Cuts	Columns	B&B Nodes
10	0	0	0	0 %	7	30	3
10	250	20	0	0 %	300	4414	210
10	1000	18	0	0 %	291	3988	203
10	10000	22	0	0 %	324	4401	234
15	0	1	0	0 %	11	45	8
15	250	1046	0	0 %	2687	36270	2597
15	1000	1034	0	0 %	2278	33037	2035
15	10000	976	0	0 %	2120	31345	1927
20	0	2	0	0 %	5	60	1
20	250	3606	5	18 %	1051	83624	391
20	1000	3605	5	12 %	1093	72048	415
20	10000	3603	5	11 %	1035	65633	395
25	0	2	0	0 %	1	75	1
25	250	3616	5	58 %	465	61431	123
25	1000	3619	5	39 %	457	52158	120
25	10000	3613	5	21 %	486	53531	127

Table 3.3: Computational results with varying revenue. Average values of 5 instances.

n	Q	Time	Timeout	Gap	Cuts	Columns	B&B Nodes
10	0	1	0	0 %	6	110	4
10	10	18	0	0 %	291	3988	203
10	30	90	0	0 %	665	8757	520
10	200	109	0	0 %	789	8934	628
15	0	1	0	0 %	8	231	4
15	10	1034	0	0 %	2278	33037	2035
15	30	3110	3	11 %	3600	101576	2781
15	200	3146	4	15 %	2957	90291	1926
20	0	2	0	0 %	4	412	2
20	10	3605	5	12 %	1093	72048	415
20	30	3608	5	inf	404	99737	108
20	200	3619	5	76 %	398	107869	102
25	0	5	0	0 %	15	632	10
25	10	3619	5	39 %	457	52158	120
25	30	3607	5	124 %	368	173356	76
25	200	3609	5	153 %	346	168005	67

Table 3.4: Computational results with varying capacity. Average values of 5 instances.

3.4 Conclusion and Further work

We have presented the SSDP, a pickup and delivery problem which differs from related pickup and delivery problems by not considering a depot, having optional demands, and having to respect

path durations for the demand. The inclusion of path durations and optional demand is a new, and probably more realistic, way of seeing liner shipping network design.

A novel arc-flow model as well as a path-flow model have been proposed, and a Branch-and-Cut-and-Price algorithm has been devised for the path-flow model. This algorithm effectively deals with the path duration limits in subproblems for each demand, while it chooses the vessel round trip, demand paths and quantity of each demand to respect the vessel capacity in the master problem. The solution method has been implemented and extensive testing shows that it is able to solve problem instances with $n = 25$ nodes and commodity density (A) or (D) to optimality in less than 3600 seconds. The model and developed solution method is generally applicable to a wide range of problems, as well as for liner shipping specific problems. If one wished to capture more of the rich problems faced in liner shipping network design, the model and solution method could be extended to: include time windows, as a carrier will often have a limited number of berth hours available at some port. Another extension would be to allow multiple port calls to some ports, as port calls both in- and outbound on a service can improve path duration for both imports and exports, this would require a model allowing non-simple cycles. The developed solution method can solve problem instances with up to said 25 nodes, which makes it applicable to the design of real world inter continental services, typically calling 10 to 20 ports.

Acknowledgments

The authors would like to thank Berit Dangaard Brouer for valuable comments, Matthew Galati for his help using **COIN-OR DIP** and The Danish Council for Strategic Research for support to the ENERPLAN project.

Bibliography

- [1] Richa Agarwal and Ozlem Ergun. Ship scheduling and network design for cargo routing in liner shipping. *Transportation Science*, 42(2):175–196, 2008.
- [2] J. Fernando Alvarez. Joint routing and deployment of a fleet of container vessels. *Maritime Economics and Logistics*, 11:186–208, 2009.
- [3] David L. Applegate, Robert E. Bixby, Vasek Chvátal, and William J. Cook. Concorde tsp solver, 2001. Available at www.tsp.gatech.edu/concorde.html.
- [4] Gerardo Berbeglia, Jean-Francois Cordeau, Irina Gribkovskaia, and Gilbert Laporte. Static pickup and delivery problems: a classification scheme and survey. *TOP, Trabajos de Investigación Operativa*, 15:1–31, 2007.
- [5] Boost. The boost graph library. Available at www.boost.org.
- [6] Berit D. Brouer, J. Fernando Alvarez, Christian E. M. Plum, David Pisinger, and Mikkel M. Sigurd. A base integer programming model and benchmark suite for liner shipping network design. *Transportation Science*, 2013. Accepted by Transportation Science.
- [7] Marielle Christiansen, Kjetil Fagerholt, and David Ronen. Ship routing and scheduling: status and perspectives. *Transportation Science*, 38(1):1–18, 2004.
- [8] Marielle Christiansen, David Ronen, Bo Nygreen, and Kjetil Fagerholt. Maritime transportation. In *Handbooks in Operations Research and Management Sciences*, pages 189–284. North Holland, 14 edition, 2007.
- [9] Matthew Galati and Ted .K. Ralphs. Coin-or dip (decomposition for integer programming), 2009. Available at projects.coin-or.org/Dip.
- [10] Shahin Gelareh and David Pisinger. Fleet deployment, network design and hub location of liner shipping companies. *Transportation Research Part E: Logistics and Transportation Review*, 47:947 – 964, 2011.
- [11] Shahin Gelareh, Stefan Nickel, and David Pisinger. Liner shipping hub network design in a competitive environment. *Transportation Research Part E: Logistics and Transportation Review*, 46(6):991–1004, November 2010.
- [12] Hipólito Hernández-Pérez and Juan-José Salazar-González. The multi-commodity one-to-one pickup-and-delivery traveling salesman problem. *European Journal of Operational Research*, 196(3):987 – 995, 2009.
- [13] Stefan Irnich and Guy Desaulniers. Shortest path problems with resource constraints. In *Column Generation*, pages 33–65. Springer US, 2005.
- [14] Mads K. Jepsen, Bjørn Petersen, Simon Spoorendonk, and David Pisinger. Introducing subset row inequalities in a branch-and-cut-and-price algorithm for the vehicle routing problem with time windows. *Operations Research*, 56:497–511, 2008.
- [15] Mads K. Jepsen, Berit Løfstedt, Christian E. M. Plum, David Pisinger, and Mikkel M. Sigurd. A path based model for a green liner shipping network design problem. In *Proceedings of the IAENG, ICOR 2011*, 2011.
- [16] Qiang Meng and Shuaian Wang. Liner shipping service network design with empty container repositioning. *Transportation Research Part E: Logistics and Transportation Review*, 47(5): 695–708, 2011.
- [17] Qiang Meng, Shuaian Wang, Henrik Andersson, and Kristian Thun. Containership routing and scheduling in liner shipping: Overview and future research directions. *Transportation Science*, 2013. To appear in.

-
- [18] Gur Mosheiov. The travelling salesman problem with pick-up and delivery. *European Journal of Operational Research*, 79(2):299 – 310, 1994.
- [19] Theo E. Notteboom. Container shipping and ports: An overview. *Review of Network Economics*, 3, 2004.
- [20] Sophie Parragh, Karl Doerner, and Richard Hartl. A survey on pickup and delivery problems. *Journal fur Betriebswirtschaft*, 58:81–117, 2008.
- [21] Bjørn Petersen. *Shortest Paths and Vehicle Routing*. PhD thesis, 2011.
- [22] Christian E. M. Plum, David Pisinger, Juan-José Salazar-González, and Mikkel M. Sigurd. The multi-commodity one-to-one pickup-and-delivery traveling salesman problem with path duration limits. In *Proceedings of the International MultiConference of Engineers and Scientists 2012 Vol II*, 2012.
- [23] Line B. Reinhardt and David Pisinger. A branch and cut algorithm for the container shipping network design problem. *Flexible Services And Manufacturing Journal*, 2011.
- [24] Koichi Shintani, Akio Imai, Etsuko Nishimura, and Stratos Papadimitriou. The container shipping network design problem with empty container repositioning. *Transportation Research Part E: Logistics and Transportation Review*, 43(1):39–59, 2007.
- [25] Martin Stopford. *Maritime Economics*. Routledge, third edition, 2009.
- [26] Shuaian Wang and Qiang Meng. Schedule design and container routing in liner shipping. *Transportation Research Record: Journal of the Transportation Research Board*, 2222:25–33, 2011.
- [27] Shuaian Wang and Qiang Meng. Robust schedule design for liner shipping services. *Transportation Research Part E: Logistics and Transportation Review*, 48(6):1093 – 1106, 2012.

Chapter 4

A Path Based Model for a Green Liner Shipping Network Design Problem

Berit D. Brouer* Mads K. Jepsen* Christian E. M. Plum*,[‡] Guy Desaulniers[†] David Pisinger* Mikkel M. Sigurd[‡]

*Department of Management Engineering, Technical University of Denmark,
Produktionstorvet, Building 426, DK-2800 Kgs. Lyngby, Denmark
blof@man.dtu.dk, makj@man.dtu.dk, cemp@man.dtu.dk, pisinger@man.dtu.dk

[†]GERAD
Montreal, Canada
Guy.Desaulniers@gerad.ca

[‡]Maersk Line
Esplanaden 50, DK-1098 København K.
mikkel.muholdorff.sigurd@maersk.com

1

Abstract Liner shipping networks are the backbone of international trade providing low transportation cost, which is a major driver of globalization. These networks are under constant pressure to deliver capacity, cost effectiveness and environmentally conscious transport solutions. This article proposes a new path based **MIP** model for the Liner shipping Network Design Problem minimizing the cost of vessels and their fuel consumption facilitating a green network. The proposed model reduces problem size using a novel aggregation of demands. A decomposition method enabling delayed column generation is presented. The subproblems have similar structure to **Vehicle Routing Problems**, which can be solved using dynamic programming. An algorithm has been implemented for this model, unfortunately with discouraging results due to the structure of the subproblem and the lack of proper dominance criteria in the labeling algorithm.

Keywords Liner shipping, Network design, Mathematical programming, Column generation, Green logistics

¹First version published in *Proceedings of The International MultiConference of Engineers and Computer Scientists (2011)*

4.1 Introduction

Global liner shipping companies provide port to port transport of containers, on a network which represents a billion dollar investment in assets and operational costs.

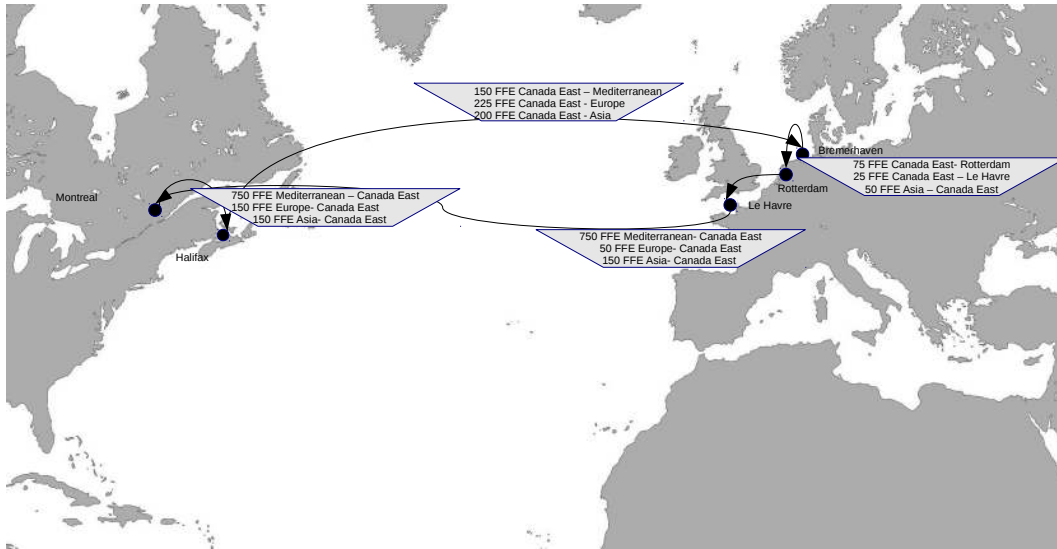


Figure 4.1: A Canada-Northern Europe service. FFE is Forty Foot Equivalent unit container used to express the volume of containers in each cargo category.

The liner shipping network can be viewed as a transportation system for general cargo not unlike an urban mass transit system for commuters, where each route (service) provides transportation links between ports and the ports allow for transshipment in between routes (services). The liner shipping industry is distinct from other maritime transportation modes primarily due to a fixed public schedule with a given frequency of port calls [22]. The network consists of a set of *services*. A *service* connects a sequence of ports in a cycle at a given frequency, usually weekly as an industry standard. In Figure 4.1 a service connecting Montreal-Halifax and Europe is illustrated. The weekly frequency means that several vessels are committed to the service as illustrated in the figure, where four vessels cover a round trip of 28 days placed with one week in between vessels. This round trip for the vessel is referred to as a *rotation*. Note that the Montreal service carries cargo to North Europe, the Mediterranean and Asia, with the two latter transshipping in Bremerhaven. In a similar way cargo headed for Canada has multiple origins. This illustrates that transshipments to other connecting services is at the core of liner shipping. Therefore, the design of a service is complex, as the set of rotations and their interaction through transshipment is a transportation system extending the supply chains of a multiplicity of businesses. Figure 4.2 illustrates two services interacting in transporting goods between Montreal-Halifax and the Mediterranean, while individually securing transport between Montreal-Halifax and Northern Europe, and Northern Europe and the Mediterranean respectively. The Montreal service additionally interacts with a service between Europe and Asia, which is partly illustrated.

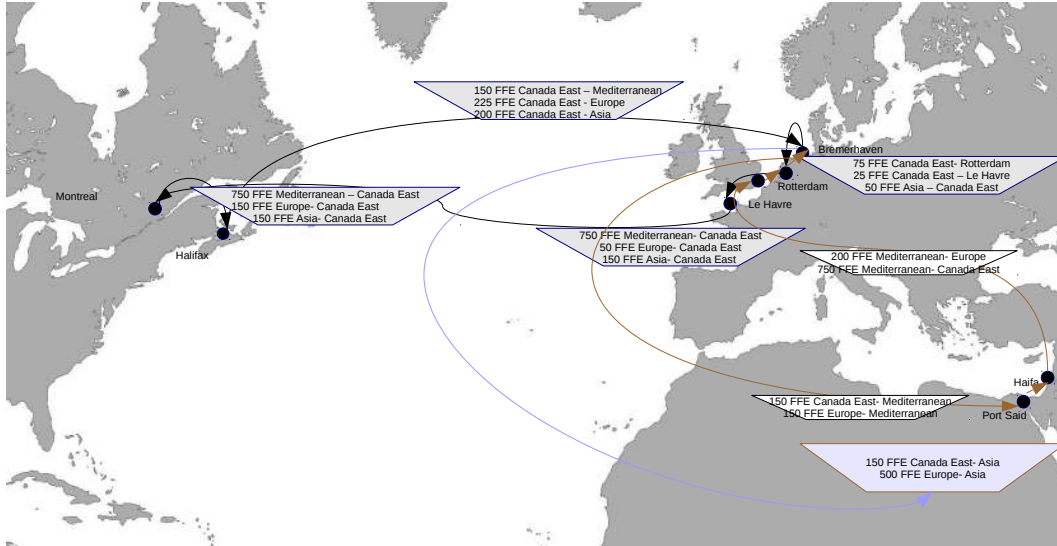


Figure 4.2: Two connecting services. The Montreal service from Figure 4.1 and a Europe-Mediterranean service with a round trip time of 2 weeks illustrated by two white vessels. The cargo composition on board vessels illustrate transshipments at the core of the liner shipping network design. The light blue incomplete service illustrates a larger service transporting cargo between Europe and Asia.

4.1.1 Modelling the Liner Shipping Network Design Problem (LSNDP)

The Liner Shipping Network Design Problem (LSNDP) aims to optimize the design of the networks to minimize cost, while satisfying customer service requirements and operational constraints. The mathematical formulation of the LSNDP may be very rich as seen in Løfstedt et al. [17], where a compact formulation along with an extensive set of service requirements and network restrictions is presented. A rich formulation like the one presented in Løfstedt et al. [17] serves as a description of the LSNDP domain, but is not computationally tractable as the number of feasible services is exponential in the number of ports. Therefore, a formulation of the LSNDP is typically restricted to an interpretation of the domain along with the core costs and constraint structures of the problem. The LSNDP has been modelled as a rich Vehicle Routing Problem (VRP) [4], where transshipments are not allowed and vessels can be assumed to return empty to a single main port of a voyage in, e.g., Fagerholt [8] and Karlaftis et al. [15]. The structure is applicable for regional liner shippers referred to as *feeder* services as opposed to global liner shipping in focus in the present paper. Models where the LSNDP is considered as a specialized capacitated network design problem with multiple commodities are found in Reinhardt and Kallehauge [20], Agarwal and Ergun [1], Alvarez [2], and Plum [19]. The network design problem is complicated by the network consisting of disjoint cycles representing container vessel routes as opposed to individual links. The models allow for transshipments, but transshipment cost is not always part of the objective (e.g., Agarwal and Ergun [1]). The vessels are not required to be empty at any time. The works of Agarwal and Ergun [1], Alvarez [2] identify a two tier structure of constraint blocks: the first deciding the rotations of a single or a collection of vessels resulting in a capacitated network and the second regarding a standard multicommodity flow problem with a dense commodity matrix. The cost

structure of LSNDP places vessel related costs in the first tier and cargo handling cost and revenue in the second tier. The work of Plum [19] has identified two main issues with solving the LSNDP as a specialized capacitated network design problem:

1. Economy of scale on vessels and the division of cost and revenue on the two tiers results in highly fractional LP solutions.
2. The degeneracy of the multicommodity flow problem results in weak LP bounds.

Furthermore, it is well known that the linear multicommodity flow problem and hence capacitated network design problems are increasingly complex to solve with the number of distinct commodities. Computational results for existing models confirm the hardness of this problem and the scalability issues, struggling to solve instances with 10-15 ports and 50-100 commodities.

The model presented in this paper has a single tier and combines revenue with total cost in the service generation problem. The motivation is to ensure efficient capacity utilization of vessels and avoid highly fractional LP solutions. Service generation is based on pick-up-and-delivery of cargoes transported entirely or partly on the service. The cost of a service reflects asset, operational and port call costs of the vessels on the service, along with the cargo handling cost and revenue of collected cargo on the service. The cargo handling cost includes load, unload and transshipment costs. The model is inspired by the Pick-up-and-Delivery VRP problem, but is considerably more complex as we allow transshipments on non-simple cyclic routes, where the vessel is not required to be empty at any point in time.

The degeneracy of the multicommodity flow problem is mitigated both by modeling the flow as assignments to services as opposed to the traditional multicommodity flow formulation, but also by exploiting the liner shipping concept of *trade lanes* to aggregate the number of distinct commodities to a minimum. Trade lanes are based on the geographic distances within a set of ports and their potential to import/export to another region.

Maritime shipping produces an estimated 2.7% of the worlds CO_2 emission, whereof 25% is accounted to container vessels according to the WorldShippingCouncil [23]. Many liner shipping companies focus on the environmental impact of their operation and the concept of *slow steaming* has become a value proposition for some liner shipping companies [16]. Cariou [5] estimate that the emissions have decreased by 11 % since 2008 by slow steaming alone. A break down of the cost of a service to each vessel Stopford [22] state that 35-50% of the cost is for fuel (bunker) whereas capital cost accounts for 30-45%, OPEX (crew, maintenance and insurance) accounts for 6-17% and port cost for 9-14%. Slow steaming minimizes the fuel cost, but comes at an asset cost of additional vessels deployed to maintain weekly frequency [18]. Slow steaming is not always an option as some cargo may have crucial transit times. Current models of LSNDP assumes fixed speed on a service. The model of Alvarez [2] explicitly aims at minimizing the fuel cost and consumption in the network by varying the speed of services in the model. The works of Løfstedt et al. [17], Notteboom and Vernimmen [18], Fagerholt et al. [10] state that the speed on a service is variable on each individual voyage between two ports. Calculating fuel consumption based on an average fixed speed on a roundtrip is an approximation, as the fuel consumption is a cubic function of speed [22]. As a result the actual fuel consumption of a service cannot be estimated until the schedule is fixed. Tramp shipping companies often model their routing and scheduling problem as rich Pick-up-and-Delivery VRP problems with Time Windows [9, 13]. Fagerholt et al. [10] is the first article within tramp shipping with variable speed between each port pair in the routing. The optimization of speed and hence minimizing the fuel consumption and environmental impact is driven by the time windows and the optional revenue of spot cargoes. [10, 11] report significant improvements in solution cost using variable speed. Minimizing the fuel consumption of the network can be a post optimization regarding speed of the liner shipping network, when deciding on the schedule in terms of berthing windows or the transit time of individual cargo routings. The path based model presented in this paper assumes a fixed speed for each vessel class and in the dynamic programming algorithm the number of vessels deployed to a service is rounded up to the nearest integer in order to ensure that a weekly frequency can be maintained on each service.

The path based model is inspired by operations research techniques within the airline industry, where the optimization is divided into faces. Therefore, a solution to the path based model is a generic capacitated network of cyclic services based on a weekly frequency of port calls. The generic network is transformed into an actual network by deciding a specific schedule, deploying vessels and deciding on the speed of the individual voyages and actual flow of all distinct commodities. The slow steaming speed of a vessel is 12 knots and depending on size and age a vessel has a maximal speed of 18 to 25 knots. If the fixed speed is chosen 30-40% above slow steaming speed for each vessel class, rounding up the number of vessels will allow post optimization of the schedule to achieve an energy efficient network with focus on slow steaming, while ensuring the transit time of products. The generic network facilitates the design of a green liner shipping network, while at the same time enabling scalability due to a more general description of the network.

4.1.2 Demand Aggregation

In models of the LSNDP using a specialized capacitated network design formulation the second tier is a standard multicommodity flow problem. The work of Alvarez [2] identifies solving the multicommodity flow problem as prohibitive for larger problem instances due to the large number of commodities considered. In Alvarez [2] the commodities are aggregated by destination, giving a smaller model to solve. This could result in worse LP bounds as identified in Croxton et al. [6], since the LSNDP will have a concave cost function, due to the economies of scales of deploying larger vessels, and high start up costs, as at least one vessel must be deployed.

A contribution of this paper is to formulate a model that considers aggregated aspects of the demand instead of specific origin-destination ($o-d$) pairs. This is motivated by the *trade*-centric view of liner-shipping present in the liner shipping industry instead of the $o-d$ -centric view considered in the literature. As seen in Figure 4.1, the ($o-d$) demand from Halifax to Rotterdam could be considered, but in practice it will be hard to estimate such a specific demand. More realistically one could estimate the volume of exports from Halifax to Northern Europe and reversely the volume of imports from East Coast Canada to Rotterdam (or exports from Mediterranean to Halifax as in Figure 4.2). Each commodity $k \in K$ will then be characterized by a volume d^{XY} from a region X to a region Y i.e. East Coast Canada or Northern Europe as seen in Figure 4.1 on the vessels in deep sea. Each set of X, Y will symbolize a *trade*. Each port $p \in X$ will also have an export and import in the trade: d^{pY}, d^{Xp} , where $\sum_{p \in X} d^{pY} = \sum_{p \in Y} d^{Xp}$ as seen in Figure 4.1 on the vessels in a region. In effect a port as Halifax will be ensured a volume of export to Mediterranean ports and each of these will be insured a volume of imports from East Coast Canadian ports, without specifying the concrete origin-destination pairs. Note the difference in aggregation approach, compared with the models of Croxton et al. [6], as we are now aggregating by *trade* origin-region to destination-region, instead of aggregation by destination port. This should give the benefit of fewer variables due to the aggregation, while we still have quite tight LP-relaxations.

The aggregation of demand may be more or less fine grained according to the definition of ports, regions and trade lanes, enabling both detailed networks for a smaller region and coarse network designs for a larger set of ports that may be refined by subsequent optimization methods. We foresee a computational tractability trade-off between the number of ports and the number of distinct commodities when defining regions for ports.

This can also be seen in the light of forecasting accuracy, usually the more detailed the level of forecasting is the more inaccurate it will be. This allows a forecasting to be done at a more natural level, i.e. on total trade volumes and total port import and export volumes.

In the following we will present a path-based formulation of the LNSDP and a column generation approach generating capacitated, cyclic rotations with assigned flow. We will outline a dynamic programming algorithm to solve the pricing problem. Preliminary computational results of an implementation of the algorithm will be given, which reveals poor performance for solving the pricing problem. This leads us to believe that alternative methods must be developed to efficiently solve the pricing problem, for the approach to be able to solve instances of a significant size. This work is an extension of a contribution to the proceedings of IMECS 2011 of Jepsen, Løfstedt,

Plum, Pisinger, and Sigurd [14].

4.2 Service Based Model

In the following we introduce a model based on a combination of feasible services for each vessel class, into a generic liner shipping network solution. The service based model is based on a Dantzig-Wolfe decomposition of the model presented in Løfstedt et al. [17]. Let S_v denote the set of feasible services for a vessel class $v \in V$ and let $S = \cup_{v \in V} S_v$. Let α_{kps}^{XY} and β_{kps}^{XY} be the amount of respectively load and unload of containers from region X to region Y on the k 'th visit to port p on service $s \in S$. We assume that $\alpha_{kps}^{XY} = \beta_{kps}^{XY} = 0, \forall p \notin X \cup Y \cup G^{XY}$, where G^{XY} is the set of ports where transshipments is allowed for *trade* XY . Let M_p be the maximal number of port visits to port p for each service. Furthermore, let γ_{pq} equal the number of times the service sails between ports $p \in P$ and $q \in P$. The move cost in a port p for a *trade* $XY \in K$ consist of the unload cost u_p^{XY} and load cost l_p^{XY} . For ports $p \in X$ the transshipment cost is included in the unload cost and the revenue is r_p^{XY} . For ports $p \in P \setminus X$ the transshipment cost is included in the load cost. Each vessel of vessel type $v \in V$ has costs c_v for fuel-, crew- and depreciation of vessel value or time-charter-costs per week. The cost of vessel type v calling a port q is c_q^v . The number of vessels used by the service is the round trip distance of the service divided by W_d^v , the weekly distance covered by vessel type v at the predefined speed. This value is rounded up to ensure the vessels can complete the round trip at the predefined speed. The number of vessels used by the service is given as $n_s = \left\lceil \sum_{p \in P} \sum_{q \in P} \frac{d_{pq} \gamma_{pqs}}{W_d^v} \right\rceil$. The cost of a service $s \in S$ is given as:

$$\begin{aligned} c_s = & \sum_{XY \in K} \sum_{p \in X} \sum_{k \in M_p} r_p^{XY} (\alpha_{kps}^{XY} - \beta_{kps}^{XY}) \\ & - \sum_{XY \in K} \sum_{p \in P} \sum_{k \in M_p} (l_p^{XY} \alpha_{kps}^{XY} + u_p^{XY} \beta_{kps}^{XY}) \\ & - c_v n_s - \sum_{p \in P} \sum_{q \in P} c_q^v \gamma_{pqs} \end{aligned}$$

The model based on services is as follows:

$$\max \sum_{s \in S} c_s \lambda_s \quad (4.1)$$

$$\text{s.t.} \quad 0 \leq \sum_{s \in S} \sum_{k \in M_p} (\alpha_{pks}^{XY} - \beta_{pks}^{XY}) \lambda_s \leq d^{pY} \quad \forall XY \in K, \forall p \in X \quad (4.2)$$

$$0 \geq \sum_{s \in S} \sum_{k \in M_p} (\alpha_{pks}^{XY} - \beta_{pks}^{XY}) \lambda_s \geq -d^{Xp} \quad \forall XY \in K, \forall p \in Y \quad (4.3)$$

$$\sum_{s \in S} \sum_{k \in M_p} (\alpha_{pks}^{XY} - \beta_{pks}^{XY}) \lambda_s = 0 \quad \forall p \in G^{XY}, \forall XY \in K \quad (4.4)$$

$$\sum_{s \in S} \sum_{p \in X \cup Y} \sum_{k \in M_p} (\alpha_{pks}^{XY} - \beta_{pks}^{XY}) \lambda_s = 0 \quad \forall XY \in K \quad (4.5)$$

$$\sum_{s \in S_v} n_s \lambda_s \leq |v| \quad \forall v \in V \quad (4.6)$$

$$\alpha_{kps}^{XY}, \beta_{kps}^{XY} \in \mathbb{Z}^+ \quad \forall s \in S, \forall XY, \forall p \in X, \forall k \in M_p \quad (4.7)$$

$$\lambda_s \in \{0, 1\} \quad \forall s \in S \quad (4.8)$$

The objective (4.1) maximizes the profit, constraints (4.2) and (4.3) ensure that the difference between what is loaded and unloaded (unloaded and loaded) by all services in a port is positive and less than the export capacity (import capacity) of the port for the given *trade*. Constraints

(4.4) ensure that the amount of containers loaded equals the amount of containers unloaded in a transshipment port and constraints (4.5) ensure that all containers loaded are unloaded for each *trade*. Constraints (4.6) ensure that the number of available vessels for each vessel class is not exceeded and the binary domain on the variables is defined by (4.8).

The key issue with the service based model is that the set of feasible services S can be exponential in the number of ports. Therefore, we cannot expect to solve instances of significant size. To overcome this issue we propose to write up the model gradually using delayed column generation and then solve the problem through Branch-and-Cut-and-Price. Branching is done by imposing a limit on the number of times an arc can be used by a given vessel class. We will investigate the possibility of applying an enumeration technique similar to the one used within CVRP [3].

4.2.1 Pricing Problem

The pricing problem calculates a non-simple cycle σ centered around any starting node p_s with associated loads and unloads. The cycle respects the capacity of the vessel class, C_v , at every port p , ensures feasibility of a weekly frequency for the vessel class v given the distance of the schedule, and lastly, that port p is visited no more than M_p times $\forall p \in P$. The pricing problem returns a variable representing a load and an unload pattern, which implicitly defines a non-simple cycle starting and ending at the same port $p \in P^v$, deploying n_s vessels to maintain weekly frequency at the fixed speed enforced on the service pattern. The above problem has a similar structure to the pricing problems of Vehicle Routing Problems modelled as a Resource Constrained Shortest Path problem (see Irnich and Desaulniers [12]. The Resource Constrained shortest Path Problem is often solved by label setting algorithms. As it is possible for the demand to be split on different paths, we need to ensure that we allow all possibilities of transshipments. This necessitates that, labels are created for each integral unit of the demand up to the minimum of the available capacity or the demand.

Objective function of the pricing problem

The objective function of the pricing problem is to find the best reduced cost of a master problem variable at the given iteration of the master problem. For each $XY \in K$ a port $p \in P$ is present in at most one of the constraints (4.2) to (4.4). Let $\omega_p^{XY}, \forall XY \in K, \forall p \in X \cup Y \cup G^{XY}$ denote the duals from (4.2) to (4.4). Let δ^{XY} be the dual variables of constraints (4.5) and π^v are the duals of constraints (4.6).

For each vessel class $v \in V$ the reduced cost of a service (column) $s \in S_v$

$$\begin{aligned} \hat{c}_s = & c_s - \sum_{XY \in K} \sum_{p \in X \cup Y \cup G^{XY}} \sum_{k \in M_p} \omega_p^{XY} (\alpha_{kps}^{XY} - \beta_{kps}^{XY}) \\ & - \sum_{XY \in K} \sum_{p \in X \cup Y} \sum_{k \in M_p} \delta^{XY} (\alpha_{kps}^{XY} - \beta_{kps}^{XY}) - \pi_v n_s \end{aligned}$$

Expanding the term \hat{c}_s and rearranging the terms according to load and unload combined with the port belonging to either X, Y or G_k we obtain the following reduced cost:

$$\begin{aligned}
\hat{c}_s = & \sum_{XY \in K} \sum_{p \in X} \sum_{k \in M_p} (r_p^{XY} - l_p^{XY} - \omega_p^{XY} - \delta^{XY}) \alpha_{kps}^{XY} \\
& + \sum_{XY \in K} \sum_{p \in X} \sum_{k \in M_p} (-r_p^{XY} - u_p^{XY} + \omega_p^{XY} + \delta^{XY}) \beta_{kps}^{XY} \\
& + \sum_{XY \in K} \sum_{p \in Y} \sum_{k \in M_p} (-l_p^{XY} - \omega_p^{XY} - \delta^{XY}) \alpha_{kps}^{XY} \\
& + \sum_{XY \in K} \sum_{p \in Y} \sum_{k \in M_p} (-u_p^{XY} + \omega_p^{XY} + \delta^{XY}) \beta_{kps}^{XY} \\
& + \sum_{XY \in K} \sum_{p \in G^{XY}} \sum_{k \in M_p} (-l_p^{XY} - \omega_p^{XY}) \alpha_{kps}^{XY} \\
& + \sum_{XY \in K} \sum_{p \in G^{XY}} \sum_{k \in M_p} (-u_p^{XY} + \omega_p^{XY}) \beta_{kps}^{XY} \\
& - (\pi_v + c_v) n_s - \sum_{p \in P} \sum_{q \in P} c_q^v \gamma_{pq}
\end{aligned}$$

The reduced cost can be rewritten as a cost connected to loading, unloading, and sailing in terms of the number of vessels deployed and the cumulative port call cost. The cost of (un)loading a demand from trade XY depends on the region of the port. If the port is from the origin region X a revenue is obtained for loading and subtracted for unloading at the port. This ensures that revenue is only collected at the initial load. The costs are the (un)load cost, and the dual values from constraints (4.2)-(4.4) concerning the flow conservation and the dual value from the flow balance constraint for the trade (4.5). If the port is from the destination region Y the cost is the (un)load cost, and the dual values from constraints (4.2)-(4.4) concerning the flow conservation and the dual value from (4.5). For a transshipment port $p \in G^{XY}$ the cost is only related to (un)load cost and the dual values of (4.2)-(4.4).

$$\hat{l}_p^{XY} = \begin{cases} r_p^{XY} - l_p^{XY} - \omega_p^{XY} - \delta^{XY} & \forall p \in X \\ -l_p^{XY} - \omega_p^{XY} - \delta^{XY} & \forall p \in Y \\ -l_p^{XY} - \omega_p^{XY} & \forall p \in G^{XY} \end{cases}$$

$$\hat{u}_p^{XY} = \begin{cases} -r_p^{XY} - u_p^{XY} + \omega_p^{XY} + \delta^{XY} & \forall p \in X \\ -u_p^{XY} + \omega_p^{XY} + \delta^{XY} & \forall p \in Y \\ -u_p^{XY} + \omega_p^{XY} & \forall p \in G^{XY} \end{cases}$$

Finally, the port call cost c_q^v is paid upon each sailing/extension onto a new port $p \in P$ and the cost $\hat{c}_v = \pi_v + c_v$ is inferred each time the distance of W_d^v is traveled.

Label setting algorithm for LSNDP

The $|V|$ pricing problems for each vessel class can be formulated as the following graph problem. Given a directed graph $G^v = (N^v, A^v)$ where the node set is $N^v = P^v \cup L^v \cup U^v$. P^v is the set of ports $\in P$ compatible with vessel class v , $L^v = \bigcup_{w \in P^v} L_w$ the set of load nodes. The sets $L_w = \{\rho_w^{XY} | \forall XY \in K, w \in X \vee Y \vee G^{XY}\}$ represents all possible loads at port w , $U^v = \bigcup_{w \in P^v} U_w$ is the set of unload nodes. The sets $U_w = \{\mu_w^{XY} | \forall XY \in K, w \in X \vee Y \vee G^{XY}\}$ represents all possible unloads at port w . In order to correctly identify transshipments and unloads of a *trade* each demand $XY \in K$ is associated with a set of load nodes $L^{XY} \subseteq L^v$ and a set of unload nodes $U^{XY} \subseteq U^v$, where $L^{XY} = \{\rho_w^{XY} | w \in X \cup Y \cup G^{XY}\}$ and $U^{XY} = \{\mu_w^{XY} | w \in X \cup Y \cup G^{XY}\}$.

The arc set is $A^v = A_s \cup A_u \cup A_l$. Define the function $h : U^v \cup L^v \rightarrow P^v, L_q \mapsto q, U_q \mapsto q$ for mapping between the load and unload nodes and the actual port $q \in P^v$ of the (un)load. The set of sailing arcs is defined as follows $A_s = \{(i, j) | i \in L^v \cup U^v, j \in P^v \setminus \{h(i)\}\}$, the set

of unload arcs $A_u = \{(i, j) | i \in P^v, j \in U_i\} \cup \{(i, j) | i \in U^v, j \in U_{h(i)}\}$ and the set of load arcs $A_l = \{(i, j) | i \in P^v, j \in L_i\} \cup \{(i, j) | i \in U^v, j \in L_{h(i)} \setminus \{\rho_{h(i)}^{XY}\}\} \cup \{(i, j) | i \in L^v, j \in L_{h(i)}\}$. The graph topology is illustrated in Figure 4.3. The distance of an arc depends on the arc type:

$$d_{ij} = \begin{cases} d_{h(i)j} & (i, j) \in A_s \\ 0 & (i, j) \in A_l \cup A_u \end{cases} \quad (4.9)$$

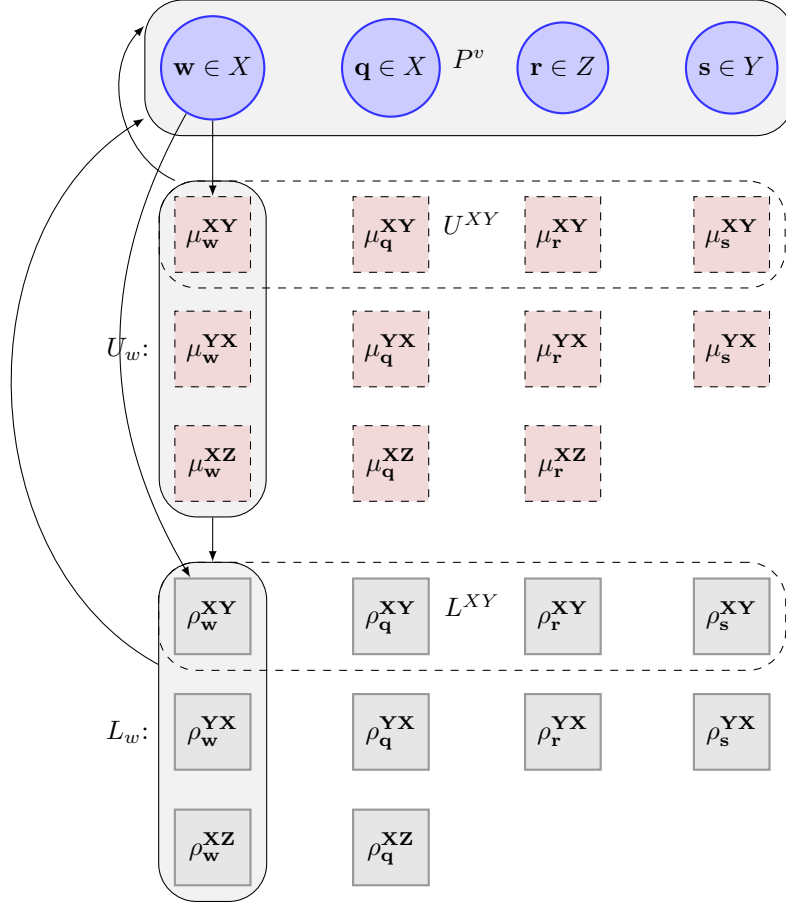


Figure 4.3: A network representation of a graph associated with the label setting algorithm. The set of port call nodes P^v (blue nodes) form a clique. For port $w \in P^v$ the sets U_w (light red nodes), L_w (grey nodes) are illustrated. They represent possible loads and unloads at port w . The sets U_w, L_w form a cliques. A path in the network will follow sequences of $n \in P_v \rightarrow U_n \rightarrow L_n \rightarrow m \in P_v$. It is possible to only unload or load. The load set of a port w is not connected to the unload set of w . Each trade $XY \in K$ is associated with a loadset L^{XY} and an unloadset U^{XY} as illustrated.

In a label setting algorithm, a label E_i is associated with a node i and represents a (partial) path with a (reduced) cost C of the service and a number of resources θ accumulated along the path. A resource may be associated with lower and upper bounds often referred to as a *resource window*. The proposed pricing problem differs significantly from the Elementary Shortest Path Problem with Resource Constraints (ESPPRC) known from VRP:

- The path is not elementary as $M_p \geq 1$.
- The path represents a cycle, σ .
- It is a longest cycle problem as the reduced cost $\hat{c}_s \geq 0$.

- We do not have a designated starting node and hence will have to start the algorithm in every possible port $p \in P^v$.
- The ability to perform a load on the partial path, which can be unloaded at a previous node of the cycle σ . A second pass of all ports in the cycle σ must be performed only allowing the unload extension function to check for load balance.
- There are multiple commodities.
- The route is combined with a loading/unloading pattern not unlike the labelling algorithm for the SDVRPTW in Desaulniers [7].

In the label setting algorithm for LSNDP a label E contains the following information:

- Current port, p_c
- Start port, p_s
- (reduced) cost, t
- Accumulated distance, d_s
- The load of each trade, $F^{XY} \quad \forall XY \in K$
- Current load, $F_c = \sum_{XY \in K} F^{XY}$
- Visit number, $k_p \quad \forall p \in P^v$

The resources are $d_s, (F^{XY})_{XY \in K}, F_c, (k_p)_{p \in P^v}$ i.e. we have $2 + |K| + |P^v|$ resources. The extension function [12] of the distance is defined as $e_{(ij)}^d(E_i) = d(E_i) + d_{ij}$. The feasibility and resource consumption of extending label E_i along an arc depends on the arc type:

- *Case 1: extending along a sail arc $(i, j) \in A_s$*

A feasible extension of label E_i to node j along a sail arc $(i, j) \in A_s$ must satisfy the following conditions:

$$\left\lceil \frac{e_{(ij)}^d(E_i)}{W_d^v} \right\rceil \leq |v| \quad (4.10)$$

$$k_j^i + 1 \leq M_j \quad (4.11)$$

Here, (4.10) ensures the feasibility of the number of vessels deployed to the service and (4.11) ensures the number of port calls to port j does not exceed M_j . If the extension is feasible a new label E_j is created. Define

$$\varpi = \left\lceil \frac{e_{(ij)}^d(E_i)}{W_d^v} \right\rceil - \left\lceil \frac{d(E_i)}{W_d^v} \right\rceil \quad (4.12)$$

ϖ expresses whether the label extension will require an additional vessel on the service to maintain weekly frequency. The following extension functions are applied to create label E_j : $p_c^j = j, p_s^j = p_s^i, t^j = t^i - c_j^v - \hat{c}_v \cdot \varpi, d = e_{(ij)}^d(E_i), F_C^j = F_C^i, F_j^{XY} = F_i^{XY}, k_j^j = k_j^i + 1, k_p^j = k_p^i \quad \forall p \in P^v \setminus \{j\}$

- *Case 2: extending along an unload arc $(i, j) \in A_u, j = \mu_p^{XY}$*

A feasible extension of label E_i to node j along unload arc $(i, j) \in A_u$ must satisfy the following conditions:

$$F^{XY} > 0 \quad (4.13)$$

where (4.13) ensures that the commodity XY is currently loaded on the vessel i.e. that a previous visit to a node in L^{XY} has been performed. To ensure that all possible transshipment and unload patterns are considered all integral unloads in $o \in \{1, \dots, \max\{d^{XY}, F_i^{XY}\}\}$ are created with separate labels.

If the extension is feasible a new label E_j^o is created using the extension functions:

$$p_c^j = h(j), p_s^j = p_s^i, t^j = t^i + \hat{u}_p^{XY} \cdot o, d = e_{(ij)}^d(E_i), F_C^j = F_C^i - o, F_j^{XY} = F_i^{XY} - o, F_j^{ZW} = F_i^{ZW} \forall ZW \in K \setminus \{XY\}, k_p^j = k_p^i \quad \forall p \in P^v.$$

- *Case 3: extending along a load arc* $(i, j) \in A_l, j = \rho_p^{XY}$
A feasible extension of label E_i to node j along a load arc $(i, j) \in A_l$ must satisfy the following conditions:

$$F_c^i < C_v \tag{4.14}$$

(4.14) ensures that the vessel has excess capacity for loading. To ensure that all possible transshipment and load patterns are considered all integral loads in $o \in \{1, \dots, \max\{d^{XY}, C_v - F_C^i\}\}$ are created with separate labels. If the extension is feasible a new label E_j^o is created with the following extension function: $p_c^j = h(j), p_s^j = p_s^i, t^j = t^i + \hat{l} \cdot o, d = e_{(ij)}^d(E_i), F_C^j = F_C^i + o, F_j^{XY} = F_i^{XY} + o, F_j^{ZW} = F_i^{ZW} \forall ZW \in K \setminus \{XY\}, k_p^j = k_p^i \quad \forall p \in P^v.$

A state is feasible when the start node is reached ($p_c = p_s$) and the containers are balanced for all trades ($F^{XY} = 0 \forall XY \in K$) by applying unload extensions to the cycle starting from p_s ending in p_s . To obtain the solution to a service the auxiliary data of what has actually been loaded and unloaded has to be stored and a mapping from L to α and from U to β creates the column entries for (un)load in the master problem. For an exact solution to the pricing problem the service with the best reduced cost ($\max \hat{c}_s$) is added to the master problem. However, the label setting algorithm may find several services, where the cost t is greater than 0 and add several columns in an iteration to accelerate convergence of the column generation algorithm.

4.2.2 Dominance

In order to dominate a label it must hold that the dominating label has the same possibilities for extensions and that no extension of the dominated label can yield a better reduced cost than the dominating label.

A label E_1 dominates a label E_2 if the following holds

- $p_c^1 = p_c^2$
- $p_s^1 = p_s^2$
- $t_1 \geq t_2$
- $d_s^1 \leq d_s^2$
- $k_p^1 \leq k_p^2 \quad \forall p \in P^v$
- $F_c^1 \leq F_c^2$
- $F_1^{XY} = F_2^{XY} \quad \forall XY \in K$

Requiring the cargo loads to be identical gives rise to a weak dominance criteria. This means that the labelling algorithm resorts to being practically brute force and a vast number of labels are generated even for relatively small instances. In recent work on dominance criteria for the Pick-up-and-Delivery problem [21] the dominance criteria for the cargo loads are strengthened by relaxing such that in our case we would have

- $F_1^{XY} \leq F_2^{XY} \quad \forall XY \in K$

if the delivery triangle inequality defined by Ropke and Cordeau [21] as $d_{ij} + d_{jk} \geq d_{ik}$ holds $\forall i, j, k \in V$. Here j is a delivery node. It is however not trivial to see, whether this relaxation holds for the pricing problem in this paper as each commodity may have several delivery nodes attached and there are no precedence relation between pickup and delivery nodes, due to the cyclic nature of a route.

4.2.3 Complexity

Let T denote an upper bound on the distance of a service. The running time of the label setting algorithm can be shown to be

$O((T|P|C|^K| \prod_{p \in X} d^{pY} \prod_{p \in G^{XY}} C)^2)$. Increasing the number of *trades* and the number of transshipment ports will increase the number of states in the Dynamic Programming algorithm. To solve practical problem instances it is therefore important to make a careful choice of the *trades* and the ports, where transshipment is allowed.

4.2.4 Relaxation of pricing problem

In CVRP a pseudo polynomial relaxation is used when solving the strongly NP-hard pricing problem [3] to reduce the practical running time of the algorithm. The method has proven to be very powerful for the CVRP. A pseudo polynomial relaxation of our pricing problem can be obtained as follows: Each port is assigned the minimal load and unload cost and the bounds on the load are removed. In each port the number of different states will then be limited to $T|P||C|$ and a running time of $O(T|P|^2|C|)$ can be obtained. However, defining a strong bound for the minimal load and unload cost for each port is not trivial as several commodities may origin or transship at a given node and further research must be conducted in order to achieve a relaxation with a good bound.

As the pricing problem is very complex, we need not solve the pricing problem to optimality in each iteration, but one could stop once a sufficient amount of columns with positive reduced cost has been found. An easy way to do this is to run the dynamic programming algorithm using a greedy variant adding any reduced cost column instead of the best reduced cost column.

4.3 Preliminary computational Results

The described algorithm has been implemented using CPLEX to solve the master problem and a labelling algorithm to solve the pricing problem. The results are currently not satisfactory for solving the pricing problem. The structure of the labelling algorithm, the lack of proper dominance criteria and especially the need to generate labels for all integral steps of load and unloads ($o \in \{1, \dots, \max\{d^{XY}, F_i^{XY}\}\}$ respectively $o \in \{1, \dots, \max\{d^{XY}, C_v - F_C^i\}\}$) creates a huge number of very similar labels. The combinations of these causes the labelling algorithm to effectively be a brute force algorithm in an extremely large search space. Even for very small graphs ($n = 4$) the number of considered labels are in the 10'ths of millions.

A simplification of the model is to only consider demand paths, which are fully loaded or unloaded either regarding the demand or capacity, i.e. $o = \{F_i^{XY}, d^{XY}\}$ respectively $o = \max\{d^{XY}, C_v - F_C^i\}$. Unfortunately, this approach is inconsistent with the idea of aggregated demands, as these *will* need to split to reach their respective origins / destinations, discouraging this direction.

As a result we have not pursued methods such as bounding to improve upon the current algorithm of the pricing problem as we believe alternative solution methods must be applied for an efficient algorithm to solve the pricing problem. Another alternative is the design and implementation of efficient heuristics to generate variables and subsequently solve a heuristic implementation of a Branch-and-Price algorithm similar to the one seen in Agarwal and Ergun [1].

4.4 Conclusion

We have presented a new model for LSNDP. Among the benefits of the proposed model is a novel view of demands in liner-shipping, which are considered on a *trade* basis. This has the advantage of giving a natural understanding, and requiring fewer variables. The model assigns cargo to routes, which may result in a tighter search space for a branch-and-bound algorithm.

A solution approach using delayed column generation has been presented, where the proposed subproblem is related to the pricing problems in VRP, where Branch-&-Cut-&-Price has been used with great success. We have discussed a pseudo polynomial relaxation to be used as bounding function, when solving the pricing problem in combination with heuristics and other techniques that have been effective in solving VRP problems. In the VRP context resource limitations have proven to be effective for the dynamic programming algorithms in reducing the state space. In the dynamic programming algorithm presented in this paper these resource limitations do not reduce complexity of the subproblem sufficiently, because dominance criteria are different. The proposed algorithm has been implemented but showed disappointing results, due to the lack of dominance criteria and a large search space for the label setting algorithm. We still believe that the main ideas in this paper can be useful to solve the LSNDP, i.e. the thoughts of combining cost and revenue in a single pricing problem and especially the notion of demand aggregation, which lends to a natural understanding in Liner Shipping. However, we must conclude that alternative methods or extensions of the current dynamic programming algorithm will be needed to solve a pricing problem, where cargo load patterns for multiple commodities are combined with a routing.

Further work with richer formulations of LSNDP, considering aspects as transit time limits on paths, and other operational constraints from liner shipping will tighten the search space of the pricing problems. However, it is uncertain whether additional real-life complexity in the pricing problem will allow for effective dominance criteria in a label setting algorithm.

Bibliography

- [1] R. Agarwal and O. Ergun. Ship scheduling and network design for cargo routing in liner shipping. *Transportation Science*, 42(2):175–196, 2008.
- [2] J. F. Alvarez. Joint routing and deployment of a fleet of container vessels. *Maritime Economics & Logistics*, 11(2):186, 2009.
- [3] R. Baldacci, N. Christofides, and A. Mingozzi. An exact algorithm for the vehicle routing problem based on the set partitioning formulation with additional cuts. *Mathematical Programming*, 115(2):351–385, 10 2008.
- [4] R. Baldacci, E. Bartolini, A. Mingozzi, and R. Roberti. An exact solution framework for a broad class of vehicle routing problems. *Computational Management Science*, 7(3):229–268, 2010.
- [5] Pierre Cariou. Is slow steaming a sustainable means of reducing co2 emissions from container shipping? *Transportation Research Part D: Transport and Environment*, 16(3):260 – 264, 2011.
- [6] Keely L. Croxton, Bernard Gendron, and Thomas L. Magnanti. Variable disaggregation in network flow problems with piecewise linear costs. *Operations research*, 55(1):146–157, 2007.
- [7] G. Desaulniers. Branch-and-price-and-cut for the split-delivery vehicle routing problem with time windows. *Operations Research*, 58(1):179–192, January 1 2010.
- [8] K. Fagerholt. Designing optimal routes in a liner shipping problem. *Maritime Policy & Management*, 31(4):259 – 268, 2004.
- [9] K. Fagerholt and H. Lindstad. Turborouter: An interactive optimisation-based decision support system for ship routing and scheduling. *Maritime Economics and Logistics*, 9:214–233, 2007.
- [10] K. Fagerholt, G. Laporte, and I. Norstad. Reducing fuel emissions by optimizing speed on shipping routes. *The Journal of the Operational Research Society*, 61(3):523–529, 08/19 2009.
- [11] K. Fagerholt I. Norstad and G. Laporte. Tramp ship routing and scheduling with speed optimization. *Transportation Research Part C: Emerging Technologies*, 19(5):853–865, 2011.
- [12] S. Irnich and G. Desaulniers. *Shortest Path Problems with Resource Constraints*, chapter 2 in Column Generation (eds. G. Desaulniers and Jacques Desrosiers and M.M. Solomon), pages 33–65. GERAD 25th Anniversary Series. Springer, 2005.
- [13] K. Fagerholt J. E. Korsvik and G. Laporte. A large neighbourhood search heuristic for ship routing and scheduling with split loads. *Computers & Operations Research*, 38(2):474–483, 2011.
- [14] M. Jepsen, B. Løfstedt, C.E.M. Plum, D. Pisinger, and M.M. Sigurd. A path based model for a green liner shipping network design problem. IAENG, 03 2011. Proceedings of The International MultiConference of Engineers and Computer Scientists 2011.
- [15] M.G. Karlaftis, K. Kepaptsoglou, and E. Sambracos. Containership routing with time deadlines and simultaneous deliveries and pick-ups. *Transportation Research Part E*, 45:210–221, 2009.
- [16] Lloyd’s List. Maersk line chief operating officer, morten engelstoft: Slow steaming is here to stay. Lloyd’s List 2010, 2010.
- [17] B. Løfstedt, C.E.M Plum, D. Pisinger, and M.M. Sigurd. An integer programming model and benchmark suite for liner shipping network design. Technical Report 19, DTU Management Engineering, 2011.

-
- [18] T.E. Notteboom and B. Vernimmen. The effect of high fuel costs on liner service configuration in container shipping. *Journal of Transport Geography*, 17(5):325–337, September 2009.
- [19] C.E.M. Plum. The linearized simultaneous string-design and cargo-routing problem. EURO 2010 conference, July 2010.
- [20] L. B. Reinhardt and B. Kallehauge. Network design for container shipping using cutting planes. 12 2007. Conference on Maritime & Intermodal Logistics, Singapore.
- [21] S. Ropke and J-F. Cordeau. Branch and cut and price for the pickup and delivery problem with time windows. *Transportation Science*, 43(3):267–286, 2009.
- [22] M. Stopford. *Maritime Economics*. Routledge, third edition, 2009. ISBN 0203442660.
- [23] WorldShippingCouncil. Liner shipping and co2 emissions policy. www.shippingandco2.org/LinerShippingandCO2EmissionsPolicySeptember.pdf, 2010.

Chapter 5

A Service Flow Model for the Liner Shipping Network Design Problem

Christian E. M. Plum^{*}, [‡] David Pisinger^{*} Mikkel M. Sigurd[‡]

^{*}Department of Management Engineering, Technical University of Denmark,
Produktionstorvet, Building 426, DK-2800 Kgs. Lyngby, Denmark
cemp@man.dtu.dk, pisinger@man.dtu.dk

[‡]Maersk Line
Esplanaden 50, DK-1098 København K.
mikkel.muhldorff.sigurd@maersk.com

1

Abstract Global liner shipping is a very competitive industry, requiring liner carriers to carefully deploy their vessels efficiently, to construct a cost competitive network. This paper presents a novel compact formulation of the liner shipping network design problem (LSNDP) based on service flows. The formulation alleviates issues faced by arc flow formulations with regards to handling multiple calls to the same port, which has not been fully dealt with earlier by LSNDP formulations. This is done by introducing service nodes, together with port nodes in a graph of the problem. Arcs from a port node to a service node represent whether a service is calling some port, and what demand loads and unloads at the port call. This representation allows recurrent calls of a service to the a port, which other optimal LSNDP models have not handled. The model ensures strictly weekly frequencies of services, ensures that port-vessel draft capabilities are not violated, respects vessel capacities and number of vessels available. The profit of the generated network is maximized, i.e. the revenue of flowed cargo subtracted a penalty for not flowed cargo and operational costs of the network. The model can be used to design liner shipping networks to utilize a container carriers assets efficiently and to investigate possible scenarios of changed market conditions. The model is solved as a Mixed Integer Program. Results are presented for the two smallest instances of the benchmark suite LINER-LIB 2012 presented in Brouer et al. [5].

Keywords Liner shipping, Network design, Maritime optimization

¹Submitted for a special issue in *European Journal of Operations Research* on Maritime Logistics

5.1 Introduction

When manufactured goods are transported from one corner of the world to another it is likely to happen in a container. These containers are carried by up to 400 meter long container vessels carrying tens of thousands of containers. Liner shipping carriers operating these vast vessels construct intricate networks of shipping routes that in their interaction provide fast and, relative to any other transport mode, cheap transport that operate at the core of the worlds supply chains.

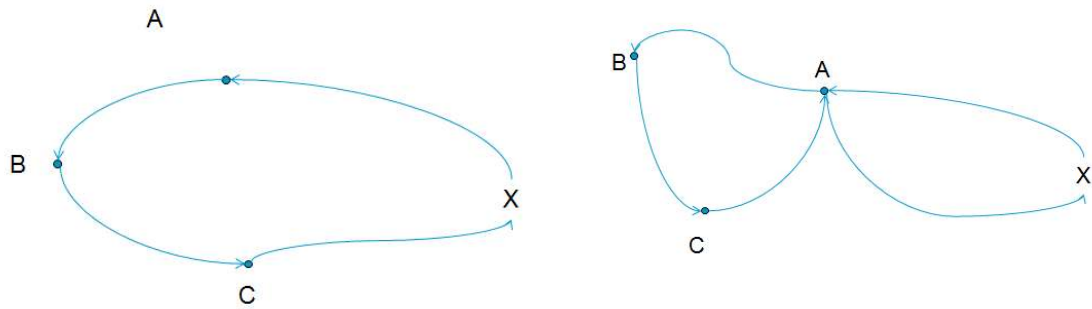
A global container shipping network is extremely costly to operate, with Maersk Line using a two-digit billion USD amount yearly to maintain this. Therefore, even a small improvement of the network's utilization, costs, service levels, etc. can have a significant impact. At the same time the cost structure of the network can be very volatile; by developing models that can investigate an increased cost or reduced demand, the network can rapidly be modified to adapt for these changed market conditions.

The basic cost components of a container shipping network are: vessel costs, bunker fuel costs, port call fees and container move costs in ports. These have been very volatile in the past years, with fuel costs reaching record levels in 2008, to fall again during the financial crisis, followed by a rise in prices during the recovery bun [2]. Vessel charter rates are similarly fluctuating and have fallen dramatically as the financial crisis hit world trade tcc [1]. Furthermore, there is an increasing interest in limiting CO2 emissions, which is related to the fuel consumption. This is already an important aspect of a shipping line company's public profile, which is expected to have increased focus in the years to come. All these factors add to the importance of liner shipping network design. This paper presents a novel model of LSNDP. A liner shipping network consists of a number of *services*, each service being sailed by a fixed number of vessels following the same roundtrip, much like a bus transit network. These services are able to tranship containers between each other at ports. Each port to port sailing by a service is denoted the service's *legs*. Each port call on the roundtrip will be served at a fixed frequency predominantly weekly (as a bus route served every 20'th minute) by a vessel. To utilize vessels best possible, service roundtrip times are thus generally a multiplier of 7 days.

These services constitute the network, that the *demand* must be transported on. Each demand is a fixed weekly volume of containers requiring transport from a specified origin to a specified destination. The demand will pay a revenue for being served, but it is acceptable to leave some demand at the cost of paying a goodwill penalty. The path of the demand will incur additional cost for moving containers on and off the vessels. The goal is then to construct a liner shipping network, consisting of some services, allowing for the transport of containers, with the aim of maximizing the profit of operating the network, e.g. the revenue subtracting costs for services, move cost and goodwill penalties.

In the proposed model, services are constructed by opening links from each service (the number of considered services being fixed) to a number of ports, using specialized service-port arcs. This can be reconstructed to a service traversing between the ports, using auxiliary variables. Likewise the demands will flow from origin port to a service and from a service to destination port (or transshipment), on specialized service-port arcs. The specific port to port path can be reconstructed using auxiliary variables. As compared to LSNDP models of related problems this model has the advantages of allowing non-simple cycles with any number of calls to one or more *butterfly ports*. Services with one or more butterfly ports is often denoted *butterfly routes*, this approach is used in service design for different reasons:

- Increased capacity on the legs between the port calls, as the service can carry less cargo on these legs. This capacity can be used for other cargo. E. g. in Figure 5.1(b) more capacity will be available between the first and second call to port A, as exports from A to region X are not carried here. An example could be Singapore as a butterfly port on a Europe to North Asia / Japan service.
- Two services with non-weekly frequency can be combined to a service with weekly frequency. E.g. Let D, E, F be ports, then service 1 calling ($D \rightarrow E \rightarrow D$) with a 4 day roundtrip time



(a) A part of a service departing from a region X, calling ports A, B and C once each, then returning to region X.

(b) A part of a service departing from a region X, calling ports A, B, C and A again, then returning to region X. A Butterfly service with A as a *butterfly port*.

Figure 5.1: Examples of two services with and without a *butterfly port*.

and service 2 ($D \rightarrow F \rightarrow D$) with a 10 day roundtrip time, can be combined to a service 3 ($D \rightarrow E \rightarrow D \rightarrow F \rightarrow D$) with a 14 day roundtrip time and thus satisfying the weekly frequency requirement, by allowing double calls at port D.

- Draft limits at later ports may require that the vessel is eased before port call, a double port call will alleviate this on port calls in between.
- Improving transit time, as the extra port call will allow for faster imports or exports to remaining ports on the service. E. g. in Figure 5.1(b) exports from port A to region X will be faster with double calls.

Note that draft dependent on vessel load and transit time is not considered in the model, but still it is a cause for using butterfly routes. Thus it is important for a LSN DP model, to accurately model butterfly port calls.

The problem of LSN DP is a strategic one, where network design decisions are made on a 6-12 months horizon, subject to amendments at a later stage. A concern with the LSN DP is that all demand is considered deterministic, with a fixed demand between the worlds ports every week. This assumption is far from reality, as container demand is subject to large fluctuations from week to week. However, a deterministic approach is still relevant, as the demand on an aggregated level gives a more stable picture. E.g. when looking at demand from continent to continent, the forecasts will be more accurate. A deterministically designed network can to some degree be corrected by short term amendments to vessel schedules, using third party tonnage or rolling cargo to the next sailing. An alternate approach to solve liner shipping network design problems could be a stochastic approach. However, stochastic models generally greatly add to the complexity compared to deterministic models. As the deterministic version of LSN DP is very hard to solve, it is unlikely that a stochastic approach will scale to anything reasonable, hence this approach has not been investigated.

This *service flow model for the liner shipping network design problem* will be abbreviated SFM.

The model has been implemented as a MIP, solved by a commercial solver. Results are reported on data from the LSN DP Benchmark instances LINER-LIB 2012, [5].

5.2 Literature

In this section we summarize the literature which has been used directly in our work and is closely related to the considered problem. For a general introduction to the maritime industry please refer to Stopford [13]. For a review of research in maritime optimization refer to Christiansen et al. [6]

and Christiansen et al. [7]. For a detailed introduction to liner shipping network design and details on the structure of the liner shipping business, its operational requirements and cost structure refer to Brouer et al. [5]. A recent overview of research in container routing and scheduling problems in liner shipping is given by Meng et al. [10].

Work on liner shipping network design is particularly relevant for this paper. The work of Shintani et al. [12] has a detailed description of the cost structure of LSNDP and includes consideration of repositioning empty containers. A service is constructed, assuming all demand is flowed if a port is called. The method allows for butterfly ports, by using duplicate virtual ports, but only for a single service.

Two decomposition methods was presented by Agarwal and Ergun [3], a bender's and a column generation based algorithm. The column generation algorithm considers the flows of the containers in a master problem, and the pricing problem considers the generation of vessel rotations by detecting negative cost cycles. Multiple services are generated and transshipment is possible. These algorithms scale well to large instances, but transshipment costs are excluded. The methods are based on a time space graph, which allows for multiple visits to a port as long as the visits do not happen on the same day of the week.

Alvarez [4] proposes a heuristic column generation based method for solving the joint routing and deployment of a fleet of container vessels. The method also considers flow in the master problem and run / service generation in the sub problem. The method is able to solve instances with 7 ports. A solution for a 120 ports instance is reported, using more aggressive heuristic approaches. The algorithm, allows butterfly ports, but can not actually handle butterfly loops, incorrectly calculating towards capacity or transshipment cost as discussed in section 5.3.1.

Reinhardt and Pisinger [11] propose a branch and cut method for the LSNDP. The formulation allows pseudo-simple cycles, meaning a cycle with a single butterfly port with at most two calls. A clover cut for the problem is derived, and some cuts from VRP are also used. This gives a branch and cut algorithm, which outperforms a pure MIP-model solved by CPLEX. Instances with up to 10 ports are solved optimally.

The work of Brouer et al. [5] describes the domain of LSNDP, discusses the relevant scoping, proposes a model of the problem, and presents a number of benchmark instances denoted LINER-LIB 2012, for the LSNDP based on real world problems. A heuristic column generation based algorithm is presented in Brouer et al. [5], which solves a number of the LINER-LIB 2012 instances. The algorithm, which can be seen as an extension of Alvarez [4], have fixed the butterfly issues, but only allows a single butterfly port per service.

Other research in liner shipping network design is Gelareh and Pisinger [8] who proposes a hub and spoke model for which a primal decomposition based solution method is derived. The method does not allow for butterfly ports. A heuristic algorithm for a scheduling and container routing problem is presented by Wang and Meng [14], but does not design the rotations. A formulation considering empty container repositioning, which considers a set of input shipping lines is found in Meng and Wang [9] and dealing with robust schedule design in Wang and Meng [15]. None of these caters for the generation of new services.

To the best of the authors knowledge no model or algorithm have been developed considering transshipment costs, allowing any number of butterfly ports on a service, and none that allows the butterfly ports to have more than two calls. This paper considers this by allowing any number of butterfly ports, with any number of port calls, while considering transshipment costs.

Overview of the paper In section 5.3 we introduce the mathematical notation used in the paper and explain details on the problem structure. The concept of service flow is central for this paper and explained in detail in section 5.3.1. This is used to formulate a model for the LSNDP in section 5.3.2. This MIP model has been implemented and solved with a commercial solver. Computational results are reported in section 5.4, with the key contributions of providing solutions for two instances of LINER-LIB 2012, while considering operational constraints as capacity, port draft and butterfly ports, which no method previously has. We conclude on the paper in section 5.5 and discuss directions for future research.

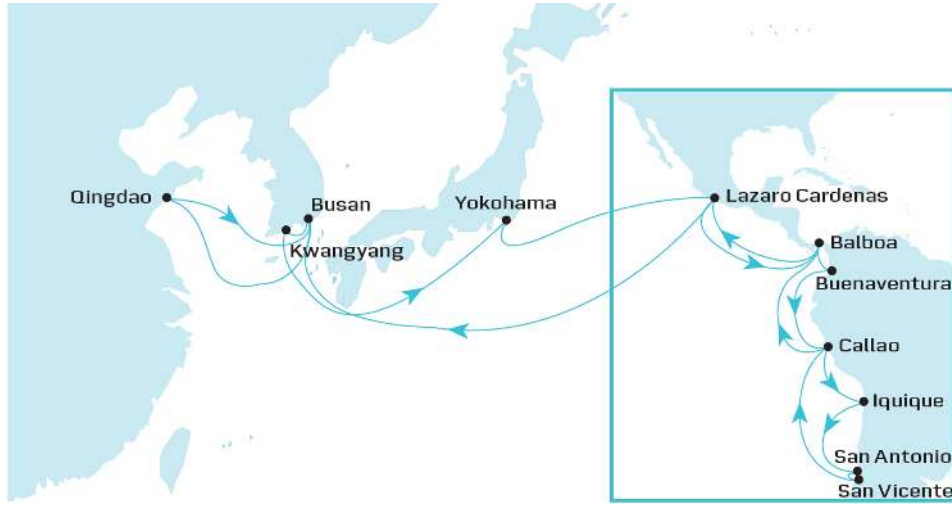


Figure 5.2: The AC3 Service, transporting containers between north Asia and central and south Americas. Note that the ports Busan, Lazaro Cardenas, Balboa and Callao have butterfly calls.

5.3 Mathematical Formulation

The notation used throughout the paper is introduced in this section. Vessels are available in different vessel classes, $v \in V$, characterized by their forty foot container capacity c_v ; a maximal number z_v of available vessels of the vessel class v . The ports $i \in P$ have the following characteristics: a port call cost for a vessel class is $w_{p,v}$; move cost for a container is u_i . The set of directed edges considered, $a \in A \subseteq P \times P$, is a subset of the complete graph spanned by the ports, let $a(i, j)$ denote the arc from port i to port j . All distances $d_{ij} = d_a$ are assumed to satisfy the triangle inequality. $o(a)$ and $d(a)$ respectively is the head and tail ports of the arc. A distance d_a from $o(a)$ to $d(a)$ is assumed not to travel through canals subject to fee's for passing (i.e. Suez and Panama). These canals are assumed to exist in $i \in P$ with appropriate $w_{i,v}$, so a canal traversal will only be possible by visiting this port and thus paying the fee.

The demand $k \in K$ with volume F^k , where $o(k)$ and $d(k)$ (same notation as for arcs) respectively is origin and destination port of the demand, is assumed to flow continuously over the considered time-period, with a steady supply every week. Demand is only to be satisfied if profitable, the revenue for doing so is q^k per container. We have a goodwill penalty for not flowing a demand \tilde{q}^k .

A service $s \in S$ is a directed, possibly non-simple, cycle of ports $i \in P$, traversed by a given vessel class v being intrinsic to s . The port calls of s , can be denoted the *rotation* of s . This rotation is traversed by an integer number τ_s of identical vessels of a vessel class v calling each port in the rotation with a weekly frequency. Services are required to have weekly frequency, i.e. they must follow a fixed schedule with total roundtrip time a multiple of 7 days, i.e. $7 \cdot \tau_s$. Let $s \in v$ imply that s is sailed by vessel class v and $v(s)$ indicate the vessel class used by service s .

A service consists of a number of port calls $b \in B = \{1, \dots, |B|\}$. We will also refer to the b 'th leg of service s , which will be the sailing from the b 'th port call to the $b + 1$ 'th port call. Note that, as the call legs b are cyclic, then $b' = |B| + 1 = 1$ to enforce a cyclic service. A service can have no more port calls than $|B|$.

Vessels are assumed to sail at their design speed s_*^v , measured in nautical miles per hour, during their whole rotation, and port stays are fixed to 24 hours for all port calls, as done in LINER-LIB 2012. These are both very broad assumptions based roughly on actual services, but with large variations. The bunker consumption at speed s_*^v is f_*^v , and the additional idling consumption of the vessel f_0^v is used irrespective of whether the vessel is sailing, for electricity, pumping and other purposes. Both consumption types are measured in metric tonnes per day.

The demand k is then transported on one of these services s either in a **direct path**, where s calls $o(k)$ loads the cargo and carries it all the way to $d(k)$ and unloads the demand. The demand

can also be carried in a **transshipment path**, where a service, s_1 calls port $o(k)$ loads the cargo and carries it to some transshipment port t where the demand is unloaded and loaded by a different service, s_2 carrying it to $d(k)$. A path can involve multiple such transshipments.

Both the demand (weekly number of containers to be flowed) and supply (number of vessels to sail on a service to achieve weekly frequency) is defined in weekly terms, hence the time dimension does not need explicit representation in the problem graph, which can be seen as a recurring representation of the demand / supply balance over a week. The model will then aim to install weekly capacity in the network on which the weekly demand can flow.

This modeling follows industry practice, where all larger and most smaller services have weekly frequency. This has the advantages of providing a stable use of resources such as vessels and port berths, easing transshipments in ports, as it allows connecting time windows of communicating services, to be minimized, and provide customers with an easily understandable service catalogue of the next departure, and its transit time (as a local bus-route with a fixed frequency).

The number of vessels required to enforce weekly frequency on a service, can be bounded by:

$$d_s/s_*^v + 24 \cdot n_s \leq \tau_s \cdot 24 \cdot 7 \quad (5.1)$$

where d_s is the roundtrip distance of s , and n_s is the number of port calls on the service. Using the same assumptions, the weekly running costs c_s of a service s can be calculated as the sum of the daily time charter rate T^v per vessel, the idling bunker consumption per day f_0^v and the sailing bunker consumption $d_s/s_*^v/24 \cdot f_*^v$

$$c_s = 7 \cdot \tau_s \cdot T^v + (7 \cdot f_0^v + d_s/s_*^v/24 \cdot f_*^v) \cdot b \quad (5.2)$$

where b is the cost of bunkers in \$ per metric tonne. Note that the unit of d_s is *nautical miles per week* ($\frac{nm}{week}$) as the whole distance is traversed each week, as the service has a weekly frequency. Following we get the units $\frac{nm}{week} / \frac{nm}{hour} / \frac{hour}{day} \cdot \frac{mt}{day} = \frac{mt}{week}$, for the last term, $d_s/s_*^v/24 \cdot f_*^v$.

Vessels are assumed to sail at design speed s_*^v on all legs. The distance of the service might not exactly fit this distance, as seen in equation 5.1 where the number of vessels is rounded up. This corresponds to the vessels lying still for a while. In practice a vessel will be able to use this time to sail a bit slower than the design speed, and due to the polynomial relation between speed and bunker consumption, it could save bunker cost. Hence the calculated bunker costs are upper bounds on actual bunker costs.

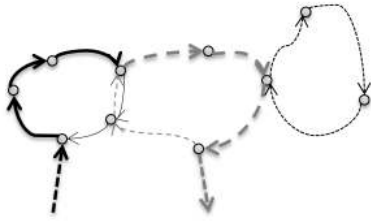
Given a graph $G(P, A)$, a set of demands K , a set of vessel classes V and data on above mentioned cost, volume, capacity, etc. the problem is to find a maximal profit liner shipping network.

5.3.1 Service flow

Current mathematical models of liner shipping network design problems use arc flow formulations, an example is seen in Figure 5.3(a). These have both flow variables $f_{i,j}$ from port i to j and network decision variables, $\delta_{i,j}$ of some form, related to the arcs of the graph. This works well when considering simple cycles. But in real world services it is common practice to have butterfly calls, for reasons mentioned in the introduction.

The issue arising with arc flow formulations, which effect butterfly services, can be illustrated by Figure 5.1(b). A node which is called twice, A , will have flow entering from two directions C and X , and flow leaving in two directions B and X . The flow entering from one of these, can only leave on one of the two leaving arcs. E.g. a demand entering from B should exit towards X , but if the formulation does not explicitly handle this, it can equally well exit towards A . The problem in this (flow entering from B leaving towards A) can be perceived in two ways; The flow cheating on capacity on the legs $A \rightarrow X \rightarrow A$ or cheating on paying transshipment costs in A .

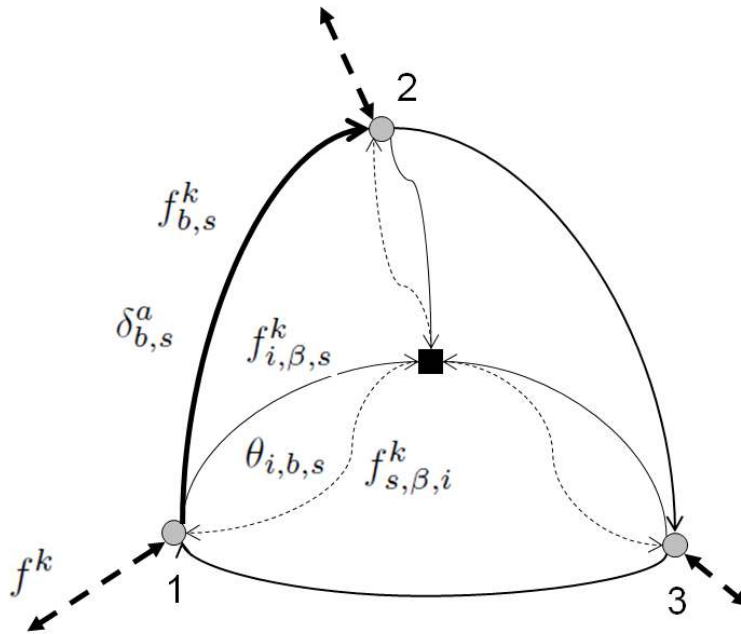
This issue can partly remedied as done in Reinhardt and Pisinger [11], which allows semi simple cycles, i.e. a vessel rotation can have a single butterfly port.



(a) Three services in an arc flow formulation. A flow path entering the first service, transshipping for the second and unloading at the flow destination is shown.



(b) The same three services in an service flow formulation. The service nodes are added in the middle of the three services, which then are connected with their called ports by arcs, opened by binary variable $\theta_{i,b,s}$ and flow by continuous variable $f_{i,b,s}^k$. A flow path entering the first service, transshipping for the second and unloading at the flow destination is shown.



(c) A three port service shown both in an arc and service formulation. A flow path loading in port 1 and unloading in port 2 is shown. Leg flow variables $f_{i,\beta,s}^k$ and $f_{s,\beta,i}^k$ are indicated next to the port to service arcs from port 1 to the service node. Binary variable for opening the arcs $\theta_{i,b,s}$ is shown in between the leg flow arcs. Port to port flow variable $f_{b,s}^k$ is shown next to the arc between port 1 and port 2, where the auxiliary port to port variable $\delta_{b,s}^{a(i,j)}$ indicating whether this arc is open, is also shown.

Figure 5.3: Edge flow modelling 5.3(a) as opposed to service flow modeling 5.3(b). Both are shown in 5.3(c)

The motivation for a service flow formulation comes from the observation that flow, for instance from a feeder port to a hub port, in a arc flow formulation is seen as a path between a number of feeder ports, finally arriving at the hub port. Rather, one could consider it by connecting the feeder port with a service, which then again was connected to the hub port. An example of this

is seen in figure 5.3(b), the square black nodes represent services, the round grey nodes represent ports, arcs pass from port-nodes to service-nodes only.

Flow variables $f_{i,b,s}^k$ represent flow of some demand k from port i to the b 'th call of service s , $f_{s,b,i}^k$ represent flow from the b 'th call of s to i . The auxiliary measured flow variables $f_{b,s}^k$ are the flow of demand k on the b 'th leg of service s . Binary decision variables $\theta_{i,b,s}$ are then set to 1 iff i is the b 'th port on service s .

This formulation allows any number and sequence of recurring calls to a butterfly port. Some problems arise from this formulation that must be tackled in different ways.

Number of port calls per service Since B and S are indices of flow and decision variables, these are input parameters of the models. B can be seen as an upper bound on number of port calls on a service. By using binary variables Ψ_{b_s} in the model, a service can be of any length less than or equal to B .

Weekly frequency Large ocean carriers predominantly use weekly frequency on their services. This is enforced in the generated services, by the equation (5.1), which ensures that the service can traverse the distance d_s with a sailing speed at or less than s_s^v . Thus the cost of the service will be a lower bound of the real cost as the vessels in practise may sail slower, allowing for less fuel spent due to the quadratic fuel consumption curve (see Brouer et al. [5] for details).

Vessel-Port Draught compatibility If vessels have larger draught than some port allows, they are not able to call it, hence the decision variable representing this must be fixed at 0.

Available vessels per vessel class Each vessel class v , has a limited number z_v of vessels available, all services using the same vessel class must adhere to this.

Capacity The capacity of each leg of the service must be respected. Looking at figure 5.3(c), it can be seen that the flow on leg 1 (from port 1 to port 2) of demand k , $f_{1,s}^k$, must adhere to:

$$f_{1,s}^k \geq \sum_i (f_{i,1,s}^k - f_{s,1,i}^k) \quad \text{and} \quad f_{1,s}^k \geq \sum_i (f_{i,1,s}^k - f_{s,1,i}^k + f_{i,3,s}^k - f_{s,3,i}^k) \quad (5.3)$$

as for all demands k and services s , at least one leg b , will have no flow $f_{b,s}^k = 0$, otherwise the flow-path would be taking the full cycle, which can not be optimal. To arrive at the flow on leg b we to need consider the largest of flow deltas ($f_{i,\beta,s}^k - f_{s,\beta,i}^k$) sums, for $\beta \in \text{circ}(b, \dots, B-1)$ previous legs. This can be generalized to:

$$\sum_{\text{circ}(b > \beta \geq \alpha)} \sum_{i \in P} (f_{i,\beta,s}^k - f_{s,\beta,i}^k) \leq f_{b,s}^k \quad \forall s, k, b, \alpha \in B \setminus \{b\} \quad (5.4)$$

where the first summation implies: for each previous call α we need to sum for legs β until the considered call b . The call b can then either be before α as seen from the 0 'th leg, giving: ($b \leq \alpha$) where we sum over β legs: $\beta < b$ and $\beta \geq \alpha$. Or b can be after α as seen from the 0 'th leg, giving: ($b > \alpha$) where we sum β legs: $b > \beta \geq \alpha$. The capacity can then be imposed on variables $f_{b,s}^k$.

The auxiliary binary decision variable $\delta_{b,s}^a$ is set to 1 iff arc a is the b 'th arc used by service s . If the draft of a port i does not allow call by a vessel v : $\text{draft}(i) < \text{draft}(v(s))$, the corresponding variables $\theta_{i,b,s}$, $\delta_{b_s}^{a(i,j)}$ and $\delta_{b_s}^{a(j,i)}$ are set to 0. The non-negative integer variable τ_s is the number of vessels used on service s .

5.3.2 Service flow formulation of the LSNDP

The LSNDP can then be formulated as the following mixed integer programming model:

$$\begin{aligned} \max \quad & \sum_{k \in K} q^k f^k - \sum_{k \in K} \tilde{q}^k (F_k - f^k) - \sum_{i \in P} u_i \sum_{k \in K} \sum_{s \in S} \sum_{b \in B} (f_{i,b,s}^k + f_{s,b,i}^k) \\ & - \sum_{s \in S} \sum_{i \in P} \sum_{b \in B} w_{i,v(s)} \theta_{i,b,s} - \sum_{s \in S} c_s \end{aligned} \quad (5.5)$$

$$\text{s.t.} \quad \sum_{i \in P} \sum_{b \in B} f_{i,b,s}^k = \sum_{i \in P} \sum_{b \in B} f_{s,b,i}^k \quad \forall k \in K, \forall s \in S \quad (5.6)$$

$$\sum_{s \in S} \sum_{b \in B} f_{d(k),b,s}^k = f^k \quad \forall k \in K \quad (5.7)$$

$$\sum_{s \in S} \sum_{b \in B} f_{i,b,s}^k = \sum_{s \in S} \sum_{b \in B} f_{s,b,i}^k \quad \forall k \in K, \forall i \in P \setminus \{o(k), d(k)\} \quad (5.8)$$

$$\sum_{\text{circ}(b > \beta \geq \alpha)} \sum_{i \in P} (f_{i,\beta,s}^k - f_{s,\beta,i}^k) \leq f_{b,s}^k \quad \forall s \in S, \forall k \in K, \quad (5.9)$$

$$\sum_{k \in K} f_{b,s}^k \leq c_{v(s)} \sum_{i \in P} \theta_{i,b,s} \quad \forall b \in B, \forall s \in S \quad (5.10)$$

$$\sum_{k \in K} (f_{i,b,s}^k + f_{s,b,i}^k) \leq 2c_{v(s)} \theta_{i,b,s} \quad \forall i \in P, \forall b \in B, \forall s \in S \quad (5.11)$$

$$f_{i,b,s}^k + f_{s,b,i}^k \leq F_k \theta_{i,b,s} \quad \forall i \in P, \forall b \in B, \forall s \in S, \forall k \in K \quad (5.12)$$

$$\theta_{i,b,s} + \theta_{j,b+1,s} \leq 1 + \delta_{b,s}^{a(i,j)} \quad \forall i, j \in P, \forall b \in B, \forall s \in S \quad (5.13)$$

$$2 * \delta_{b,s}^{a(i,j)} \leq \theta_{i,b,s} + \theta_{j,b+1,s} + \Psi_{b+1,s} \quad \forall i, j \in P, \forall b \in B, \forall s \in S \quad (5.14)$$

$$\theta_{i,b,s} + \Psi_{b+1,s} + \theta_{j,1,s} \leq 2 + \delta_{b,s}^{a(i,j)} \quad \forall i, j \in P, \forall b \in B, \forall s \in S \quad (5.15)$$

$$\sum_{i \in P} \theta_{i,b,s} + \Psi_{b,s} = 1 \quad \forall b \in B, \forall s \in S \quad (5.16)$$

$$\sum_{a \in A} \delta_{b,s}^a + \Psi_{b,s} = 1 \quad \forall b \in B, \forall s \in S \quad (5.17)$$

$$\Psi_{b,s} \leq \Psi_{b+1,s} \quad \forall b \in B \setminus \{|B|\}, \forall s \in S \quad (5.18)$$

$$\sum_{a \in A} \sum_{b \in B} d_a / s_*^v \cdot \delta_{b,s}^a + 24 \sum_{i \in P} \sum_{b \in B} \theta_{i,b,s} \leq 24 \cdot 7 \cdot \tau_s \quad \forall s \in S \quad (5.19)$$

$$\sum_{s \in v} \tau_s \leq z_v \quad \forall v \in V \quad (5.20)$$

$$0 \leq f^k, f_{i,b,s}^k, f_{s,b,i}^k, f_{b,s}^k \leq F_k \quad \forall k \in K, \forall i \in P, \forall b \in B, \forall s \in S \quad (5.21)$$

$$\theta_{i,b,s}, \delta_{b,s}^a \in \{0, 1\} \quad \forall i \in P, a \in A, \forall b \in B, \forall s \in S \quad (5.22)$$

$$\Psi_{b,s}, \tau_s \in \mathbb{Z}_+ \quad \forall b \in B, \forall s \in S \quad (5.23)$$

The objective is to maximize the profit of flowed cargo on the opened network, by optimizing the revenue subtracted a penalty for cargo not flowed, done in the two first terms of (5.5), the container move cost are subtracted in the third term, as are the port call cost in the fourth term and lastly the costs of deployed vessels and bunker cost in c_s .

The flow must adhere to flow conservation in service nodes ensured by constraint (5.6) and port nodes by constraint (5.8). Constraints (5.7) allows flow to enter at the demands source node. The $f_{b,s}^k$ variables must be set by constraint (5.9). The capacity is enforced in constraint (5.10). We

can only flow on open arcs, this is handled by constraint (5.11), using service capacity as big M, which is also tightened by constraints (5.12). Constraints (5.13) - (5.18) ensure proper construction of services, i.e. Arc decision variables $\delta_{b,s}^a$ are set in (5.13). Constraints (5.14) and (5.15) links variables $\theta_{i,b,s}$ and $\delta_{b,s}^a$. Each port call must have exactly one call as handled by constraints (5.16) and (5.17). Constraint (5.18) propagates the variable $\Psi_{b,s}$ through b if set. Weekly frequency is enforced as lower bound on service distance with constraint (5.19). The number of vessels is constrained by (5.20). Variable bounds are given in (5.21) - (5.23).

To further tighten the formulation a number of additional constraints can be added:

$$\sum_{j \in P} \sum_{b \in B} \delta_{b,s}^{a(i,j)} = \sum_{j \in P} \sum_{b \in B} \delta_{b,s}^{a(j,i)} \quad \forall i \in P, \forall s \in S \quad (5.24)$$

$$\sum_{i \in P} \sum_{j \in P} \sum_{b \in B} \delta_{b,s}^{a(i,j)} = \sum_{i \in P} \sum_{b \in B} \theta_{i,b,s} \quad \forall s \in S \quad (5.25)$$

$$\theta_{i,b,s} + \theta_{i,b+1,s} \leq 1 \quad \forall i \in P, \forall b \in B, \forall s \in S \quad (5.26)$$

$$\sum_{k \in K} d_{o(k),d(k)} f^k \leq \sum_{s \in S} \sum_{a \in A} \sum_{b \in B} c_{v(s)} d_a \delta_{b,s}^a \quad (5.27)$$

$$w_{i,v(s)} \theta_{i,b,s} \leq \sum_{k \in K} (f_{i,b,s}^k + f_{s,b,i}^k) (q^k + \bar{q}^k - u_{o(k)} - u_{d(k)}) \quad \forall i \in P, \forall b \in B, \forall s \in S \quad (5.28)$$

where constraints (5.24) - (5.26) tighten the domain of the binary variables. Constraint (5.24) gives that service s for a port i must enter and leave an equal number of times, with regards to the auxiliary $\delta_{b,s}^{a(i,j)}$ variables. Constraint (5.25) gives that the sum of port calls given by $\delta_{b,s}^a$ and $\theta_{i,b,s}$ variables must equal. Constraint (5.26) ensures that two consecutive port calls can not call the same port, i . Constraint (5.27) has the total flowed volume multiplied by its direct distance as a lower bound for all services roundtrip distances multiplied by their capacity, as an optimal capacity allocation would satisfy this. It is required by constraint (5.28) that revenue and goodwill of demand loaded and unloaded in a port call, subtracted the first load and last discharge move, exceeds the port call cost.

5.3.3 Symmetry

The formulation will suffer considerably of degenerate optimal solutions, due to the sets B and S , some of the degeneracy can be dealt with by two types of symmetry breaking constraints.

$$\sum_{k \in K} (f_{i,b,s}^k + f_{s,b,i}^k) \leq \sum_{k \in K} (f_{i,1,s}^k + f_{s,1,i}^k) \quad \forall b \in B \setminus \{1\}, \forall s \in S \quad (5.29)$$

$$\sum_{a \in A} \sum_{b \in B} \delta_{b,s}^a \cdot d_a \leq \sum_{a \in A} \sum_{b \in B} \delta_{b,s+1}^a \cdot d_a \quad \forall s \in v, \forall v \in V, \quad (5.30)$$

The first considers the flow on a service, as the legs of the service are numbered from $\{1, \dots, |B|\}$, an identical service can be generated by shifting the numbering of all legs by one. Constraint (5.29) will break this symmetry by requiring that the 1'th call of the service always has the most load and unload moves. Note that if several calls have the same number of moves, symmetry will still exist. Constraint (5.30) considers services of the same vessel class, requiring these to be ordered by non-decreasing distance, this will alleviate symmetry arising in the formulation.

5.3.4 Bounding of $|S|$ and $|B|$

The solution space of the model will be bounded by the size of the sets S and B , as this will impose artificial limits on how many and long the generated services can be. Hence it is valuable to know upper bounds that these can take in a non constrained version, to be able to set S and

Instance	Scenario	#Ports	#Demands	# Feeder 450	# Feeder 800
Baltic	Low	12	22	3	2
	Base	12	22	4	2
	High	12	22	5	2
WAF	Low	19	38	11	22
	Base	19	38	14	28
	High	19	38	17	34

Table 5.1: The **Baltic** and **WAF** instances from LINER-LIB 2012. # Ports is the number of ports, # Demands is the number of demands, # Feeder 450 and # Feeder 800 is the number of vessels available of the vessel classes.

B accordingly. A lower bound can only be $0 \leq |S|$ and $0 \leq |B|$ as the 0 solution is feasible and will in some instances be optimal. For an upper bound, consider a solution where each demand k had its own dedicated direct service, thus using two legs. This gives the bounds $|S||B| \leq 2|K|$, $|S| \leq |K|$, $|B| \leq 2|K|$.

5.3.5 Problem Complexity

The problem is NP-hard by reduction from the TSP. Given a set of nodes $i, j \in P$ with distances d_{ij} between the nodes the TSP problem asks to find the shortest Hamiltonian cycle visiting all nodes P . For a given instance of TSP we can reduce it to an instance of LSNDP by having only one vessel class v , and only one vessel $z_v = 1$. Let the demand f^k between each pair of nodes P be 1 and the penalty \tilde{q}^k for not shipping the cargo be infinitely large. The sailing bunker consumption between ports i and j is set to d_{ij} and the idling bunker consumption is set to zero. The sailing time between all pair of nodes is set to $24 \cdot 7/|P|$, such that any route visiting all nodes takes exactly one week. Finally, the transshipment move cost is set to infinity and the port call cost is set to 0.

5.4 Computational results

The Mixed Integer Program of (5.5) - (5.28), SFM has been implemented as an MIP-model and solved by CPLEX 12.2 on a Intel i5 @ 2.53 GHZ with 3 GB of RAM. It has been run with the LINER-LIB 2012 instances and compared with the best solutions found in Brouer et al. [5].

Instances The LINER-LIB 2012 benchmark is a set of publicly available instances for the LSNDP presented in Brouer et al. [5]. All relevant data for a network design instance is available, distances between ports d_a ; Vessel classes v described by their capacity, drafts, cost and speed limitations; a list of ports i described by their draft limits, call cost and move cost. Seven instances of varying size is available, each with a fleet of vessels available. Each instance has a demand list K with a volume and revenue that can be flowed. As the size of the SFM is proportional to the size of the number of port calls per service B and the number of services S , this imposes an additional constraint on the model. For the run instances B and S have been set high, to limit the constraining impact, but due to memory shortage for large B and S values a balance has been struck.

SFM results The model has been run on the two smallest instances of LINER-LIB 2012, **Baltic** and **WAF**, Table 5.1 has details on the sizes of these. Both of the instances have one main port (Bremerhaven and Algeiras) and all demands are to or from this main port to feeder ports. Revenue and all costs parameters outlined in the objective Equation (5.5) is included: move cost for loading and unloading containers; port call cost, vessel and bunker costs (both propulsion and idling bunker consumption).

Instance	Scenario	\bar{q}^k	Frequency	Objective	U.B.	Time	$\sum_s B $	$ S $	# vessels
Baltic	Low	0	7	427,485	611,015	3,600	9/20	2/2	4/5
	Base	0	7	408,771	669,774	3,600	10/20	1/2	4/6
	High	0	7	636,152	657,021	3,600	10/20	1/2	4/7
WAF	Low	0	7	1,940,817	6,051,697	10,800	10/24	3/4	11/33
	Base	0	7	3,372,618	6,399,041	10,800	24/24	4/4	26/42
	High	0	7	3,899,767	6,614,613	10,800	24/24	4/4	28/51

Table 5.2: Test results for the **Baltic** and **WAF** instances run by SFM. \bar{q}^k is the used goodwill penalty for rejected cargo. The *Frequency* of the generated services. *Objective* is the objective value of the best found solution by SFM, *U.B.* is the upper bound as given by CPLEX for the SFM formulation, *Time* is the time limit in seconds. $\sum_s |B|$ is the total number of port calls used of the allowed. $|S|$ is the number of services used of the allowed. # vessels is the number of vessels used of the allowed.

Instance	Scenario	Z	Q	c_v	c_b	c_p	c_m	c_t	F
Baltic	Low	427,485	2,389,910	-225,400	-226,833	-273,488	-1,236,703	0	0
	Base	408,771	2,020,328	-140,000	-243,008	-272,637	-954,365	-1,546	0
	High	636,152	2,641,046	-112,000	-225,343	-277,309	-1,390,241	0	0
WAF	Low	1,940,817	4,508,620	-597,800	-833,453	-124,516	-1,012,034	0	0
	Base	3,372,618	9,536,510	-1,162,000	-2,210,227	-322,030	-2,241,530	-228,105	0
	High	3,899,767	9,907,400	-968,800	-2,311,786	-327,024	-2,400,023	0	0

Table 5.3: Revenue and cost details for the best solution for **Baltic** and **WAF** instances found with SFM. Z is the objective value, Q is the revenue of flowed cargo, c_v is vessel cost, c_b is bunker cost, c_p is port call cost, c_m is local move cost, c_t is transshipment move cost and Y the sum of goodwill penalties for rejected cargoes.

The instances have been run with zero goodwill penalty $\bar{q}^k = 0$ for demand not flowed. This resembles a scenario where a carrier solely focuses on optimizing profits. The results are given in Tables 5.2 and 5.3. None of the instances can be solved to optimality, although one has a gap of just 3 %. Profitable solutions are found for all instances. Due to their size the **WAF** instances are considerably harder than the **Baltic** instances. No instance use all vessels, but it cannot be ruled out that B and S are restrictive and thus limits the use of vessels. Table 5.4 shows the considerable size of the MIPs. Referring to Table 5.3 it can be seen that only two instances use transshipment, and one of these on a very small scale. This is reasonable as all demand is from or to a main port. An example of a solution can be seen in Table 5.8.

Benchmark Instance Algorithm In the spirit of the LINER-LIB 2012, the results of the SFM model have been compared with the results from Brouer et al. [5], the Benchmark Instance Algorithm (BIA). The SFM model has some differences from the model and algorithm of BIA, so these are not directly comparable, but by highlighting the differences it is still meaningful to discuss.

BIA allows biweekly services in some cases. The average frequencies of the best BIA result services, can be seen in Table 5.5, where the average frequency is closer to 14 than 7, for all instances. This allows solutions serving low demand ports cheaper, as larger (and thus generally with cheaper unit cost) vessels can serve these less frequently. To make the results of SFM closer

Instance	# Rows	# Cols	# Non Zero's
Baltic	19,949	14,201	621,018
WAF	54,943	46,654	1,001,869

Table 5.4: MIP Instance sizes for **Baltic** and **WAF** with SFM. #Rows is the number of rows. #Cols is the number of columns. #NonZero's is the number of non zero matrix entrances.

Instance	Scenario	$ S $	$\sum_s B_s $	Avg. speed	Avg. frequency	Best Objective
Baltic	Low	5	27	13.6	13.6	235,044
	Base	4	29	13.3	12.6	325,305
	High	5	32	12.9	13	609,700
WAF	Low	7	63	12.3	12.2	4,486,261
	Base	11	68	11.6	11.7	5,565,389
	High	12	68	11.4	11.7	6,220,044

Table 5.5: Details of the best results found by the BIA algorithm. $|S|$ is the number of services used and $\sum_s |B_s|$ is the total number of port calls in the solution. Avg. speed is the averaged speed of the services, weighted by their number of vessels. Avg. frequency is the averaged frequency of the services, weighted by their number of vessels.

to the BIA results, SFM has been modified to allow biweekly services (by halving port call, bunker and vessel cost, halving capacity and doubling the right hand side of Equation (5.19), and the goodwill penalty has been set to $\tilde{q}^k = 1000$, as done in BIA. The results of the modified SFM can be seen in Tables 5.6 and 5.7. Comparing these against the original runs (Tables 5.2 and 5.3), the consequences of halving the vessels capacity can be seen, all port calls B and services S are used in all instances. This is likely one of the reason for the worse solutions. Another is the goodwill penalty \tilde{q}^k , which lowers the overall objective, but also seems to make the problem significantly harder, with much worse gaps. A reason for this is Equation (5.28) being looser with increased goodwill penalty. Details of the best solutions are found in Table 5.3 and 5.7 showing the distribution between revenue and the different cost components.

Still there are other differences between SFM and BIA which can not be mitigated. BIA does not enforce strict (bi)weekly frequency, but allows slight deviations. This gives up to 9 % extra capacity. BIA optimizes on the speeds of the generated services, allowing it to slow steam, where beneficial and thus get lower bunker costs. Or allows the services to speed up to get additional cargo where possible, details on the sailing speeds can be seen in Table 5.5. Lastly BIA is a randomized algorithm, which gives different solutions depending on the seed. We are comparing against the best results from 10 randomized runs. The results of BIA is reported in Brouer et al. [5] over 180 days, and are minimizing the cost subtracted the revenue. The SFM is run on weekly values and is maximizing the revenue subtracted the cost. Hence all values taken from BIA results have been multiplied by $-180/7$ to be comparable with the results of SFM.

On the other hand SFM allows any number of butterfly ports, opening the solution space, where BIA only allows one butterfly port with two port calls. SFM is restricted in its solution space if the sizes of $|B|$ and $|S|$ are limiting. Due to lack of memory the model has not been tested on larger instances.

Comparing results Looking at Tables 5.5 and 5.6 it can be seen that BIA finds a better solution in all cases, but except for one instance the BIA solutions are within the upper bound of SFM. One reason is that the sizes of $|B|$ and $|S|$ in SFM is severely restricting, as BIA uses significantly more port calls and services. The upper bounds of SFM exceeds the best found solutions of BIA except for the instance High WAF instance, which must be due to the differences in the two formulations. In all cases the distance between the BIA solution and the SFM upper bound is small, thus validating the quality of the BIA solutions as giving very good results close to a comparable upper bound.

5.5 Conclusion

A novel model has been presented for the LSNDP. To the best of our knowledge the model is the first method to fully allow any number of butterfly ports, allowing with more than two port calls, a crucial property as many real world services often call a number of ports more than once on the

Instance	Scenario	\tilde{q}^k	Frequency	Objective	U.B.	Time	$\sum_s B $	$ S $	# vessels
Baltic	Low	1,000	14	-265,117	794,524	3,600	20/20	2/2	5/5
	Base	1,000	14	-134,687	830,563	3,600	20/20	2/2	5/6
	High	1,000	14	-183,348	849,487	3,600	20/20	2/2	6/7
WAF	Low	1,000	14	411,317	5,650,061	10,800	24/24	4/4	20/33
	Base	1,000	14	1,059,352	5,665,930	10,800	24/24	4/4	20/42
	High	1,000	14	1,281,583	5,825,944	10,800	24/24	4/4	20/51

Table 5.6: Test results for the **Baltic** and **WAF** instances run by SFM. \tilde{q}^k is the used goodwill penalty for rejected cargo. The *Frequency* of the generated services. *Objective* is the objective value of the best found solution by SFM, *U.B.* is the upper bound as given by CPLEX for the SFM formulation, *Time* is the time limit in seconds. $\sum_s |B|$ is the total number of port calls used of the allowed. $|S|$ is the number of services used of the allowed. # vessels is the number of vessels used of the allowed.

Instance	Scenario	Z	Q	c_v	c_b	c_p	c_m	c_t	F
Baltic	Low	-265,117	3,417,398	-151,900	-298,605	-425,925	-1,841,615	-18	-964,453
	Base	-134,687	3,455,032	-108,500	-284,312	-425,921	-1,870,986	0	-900,000
	High	-183,348	3,439,210	-100,800	-323,047	-489,550	-1,911,161	0	-798,000
WAF	Low	411,317	8,491,320	-622,300	-1,508,075	-197,533	-2,100,095	0	-3,652,000
	Base	1,059,352	8,828,980	-444,500	-1,431,523	-203,320	-2,135,286	0	-3,555,000
	High	1,281,583	9,020,610	-355,600	-1,478,053	-226,845	-2,121,531	0	-3,557,000

Table 5.7: Revenue and cost details for the best solution for **Baltic** and **WAF** instances found with SFM. Z is the objective value, Q is the revenue of flowed cargo, c_v is vessel cost, c_b is bunker cost, c_p is port call cost, c_m is local move cost, c_t is transshipment move cost and Y the sum of goodwill penalties for rejected cargoes.

Service	Vessel Class	Num Vessels	Rotation
Service 1	450	9	<i>Algeciras</i> → <i>Douala</i> → <i>Algeciras</i> → <i>Tema</i> → <i>Algeciras</i> → <i>Conakry</i>
Service 2	450	8	<i>Algeciras</i> → <i>Lome</i> → <i>Algeciras</i> → <i>Luanda</i> → <i>Tema</i> → <i>Dakar</i>
Service 3	800	5	<i>Dakar</i> → <i>Algeciras</i> → <i>Dakar</i> → <i>Algeciras</i> → <i>Dakar</i> → <i>Apapa</i>
Service 4	800	6	<i>Dakar</i> → <i>Algeciras</i> → <i>Cotonou</i> → <i>Dakar</i> → <i>Algeciras</i> → <i>Abidjan</i>

Table 5.8: The best found solution for the **WAF** High case with seven days frequency and no goodwill penalty. Note that Service 3 has both Dakar and Algeciras as butterfly port, Dakar with three port calls.

rotation. The model has been run on two of the benchmark instance of LINER-LIB 2012 providing solutions for these.

Due to the large number of variables and constraints used by the model it is unable to solve the instances to full optimality. Future work could address this in a number of ways, a MIP based heuristic could be developed to find good solutions fast. Decompositions of the model could be investigated, perhaps exploiting the separate services and flow structure to deal with these problems separately, or thirdly, tighter constraints and cuts could be devised for the problem.

Acknowledgements

The authors gratefully acknowledge the support from The Danish Council for Strategic Research to the ENERPLAN project. The authors would like to thank Berit Dangaard Brouer and Christian Vad Karsten for valuable comments. We thank two reviewers for very qualified and constructive comments, this work has been greatly enhanced due to them.

Bibliography

- [1] Time charter rates according to cbs, statistics netherlands. <http://www.cbs.nl/en-GB/menu/themas/dossiers/globalisering/publicaties/artikelen/archief/2009/2009-2801-wm.htm>, 2009.
- [2] Bunker prices from bunkerworld. <http://www.bunkerworld.com/prices/index/bw380>, 2012.
- [3] Richa Agarwal and Ozlem Ergun. Ship scheduling and network design for cargo routing in liner shipping. *Transportation Science*, 42(2):175–196, 2008.
- [4] J. Fernando Alvarez. Joint routing and deployment of a fleet of container vessels. *Maritime Economics and Logistics*, 11:186–208, 2009.
- [5] Berit D. Brouer, J. Fernando Alvarez, Christian E. M. Plum, David Pisinger, and Mikkel M. Sigurd. A base integer programming model and benchmark suite for liner shipping network design. *Transportation Science*, 2013. Accepted by Transportation Science.
- [6] Marielle Christiansen, Kjetil Fagerholt, and David Ronen. Ship routing and scheduling: status and perspectives. *Transportation Science*, 38(1):1–18, 2004.
- [7] Marielle Christiansen, David Ronen, Bo Nygreen, and Kjetil Fagerholt. Maritime transportation. In *Handbooks in Operations Research and Management Sciences*, pages 189–284. North Holland, 14 edition, 2007.
- [8] Shahin Gelareh and David Pisinger. Fleet deployment, network design and hub location of liner shipping companies. *Transportation Research Part E: Logistics and Transportation Review*, 47:947 – 964, 2011.
- [9] Qiang Meng and Shuaian Wang. Liner shipping service network design with empty container repositioning. *Transportation Research Part E: Logistics and Transportation Review*, 47(5): 695–708, 2011.
- [10] Qiang Meng, Shuaian Wang, Henrik Andersson, and Kristian Thun. Containership routing and scheduling in liner shipping: Overview and future research directions. *Transportation Science*, 2013. To appear in.
- [11] Line B. Reinhardt and David Pisinger. A branch and cut algorithm for the container shipping network design problem. *Flexible Services And Manufacturing Journal*, 2011.
- [12] Koichi Shintani, Akio Imai, Etsuko Nishimura, and Stratos Papadimitriou. The container shipping network design problem with empty container repositioning. *Transportation Research Part E: Logistics and Transportation Review*, 43(1):39–59, 2007.
- [13] Martin Stopford. *Maritime Economics*. Routledge, third edition, 2009.
- [14] Shuaian Wang and Qiang Meng. Schedule design and container routing in liner shipping. *Transportation Research Record: Journal of the Transportation Research Board*, 2222:25–33, 2011.
- [15] Shuaian Wang and Qiang Meng. Robust schedule design for liner shipping services. *Transportation Research Part E: Logistics and Transportation Review*, 48(6):1093 – 1106, 2012.

Part III

Operational Liner Shipping Problems

Chapter 6

The Vessel Schedule Recovery Problem (VSRP) - a MIP model for handling disruptions in liner shipping

Berit D. Brouer* Jakob Dirksen* David Pisinger* Christian E. M. Plum* Bo Vaaben*

*Department of Management Engineering, Technical University of Denmark,
Produktionstorvet, Building 426, DK-2800 Kgs. Lyngby, Denmark
blof@man.dtu.dk, jakob@dirksen.nu, pisinger@man.dtu.dk, cemp@man.dtu.dk,
bova@man.dtu.dk

1

Abstract Containerized transport by liner shipping companies is a multi billion dollar industry carrying a major part of the world trade between suppliers and customers. The liner shipping industry has come under stress in the last few years due to the economic crisis, increasing fuel costs, and capacity outgrowing demand. The push to reduce CO_2 emissions and costs have increasingly committed liner shipping to *slow-steaming policies*. This increased focus on fuel consumption, has illuminated the huge impacts of operational disruptions in liner shipping on both costs and delayed cargo. Disruptions can occur due to adverse weather conditions, port contingencies, and many other issues. A common scenario for recovering a schedule is to either increase the speed at the cost of a significant increase in the fuel consumption *or* delaying cargo. Advanced recovery options might exist by swapping two port calls or even omitting one. We present the Vessel Schedule Recovery Problem (VSRP) to evaluate a given disruption scenario and to select a recovery action balancing the trade off between increased bunker consumption and the impact on cargo in the remaining network and the customer service level. It is proven that the VSRP is \mathcal{NP} -hard. The model is applied to four real life cases from Maersk Line and results are achieved in less than 5 seconds with solutions comparable or superior to those chosen by operations managers in real life. Cost savings of up to 58% may be achieved by the suggested solutions compared to realized recoveries of the real life cases.

Keywords: disruption management, liner shipping, mathematical programming, recovery

6.1 Introduction

Disruptions occur often in a global liner shipping network. According to Notteboom [23] approximately 70-80% of vessel round trips experience delays in at least one port. The common causes

¹Published in *European Journal of Operations Research* (2012) [6]

are bad weather, strikes in ports, congestions in passageways and ports, and mechanical failures. More exceptional causes include piracy and crew strikes on the vessels.

Example: The vessel Maersk Sarnia is deployed on a scheduled service providing transport of container cargo between South-East Asia and the west coast of Central America, see Figure 6.1. During the pickup of cargo in South-East Asia the weather conditions cause Maersk Sarnia to suffer a 30 hour delay when leaving Kwangyang in South Korea. The delay can cause the vessel to miss an important scheduled port call in the transshipment port of Balboa in Panama. As a result large parts of the cargo will miss their onward connections and most cargo will not be delivered on time.

In order to mitigate the negative effects of the delay on Maersk Sarnia the operations center at Maersk Line has several options:

- Omit the upcoming port calls at Yokohama, Lazaro Cardenas, or Balboa.
- Speed up significantly to try to reach Balboa on time.
- Swap the port calls of Lazaro Cardenas and Balboa.
- Accept the delay and catch up the schedule returning to South-East Asia from Balboa.

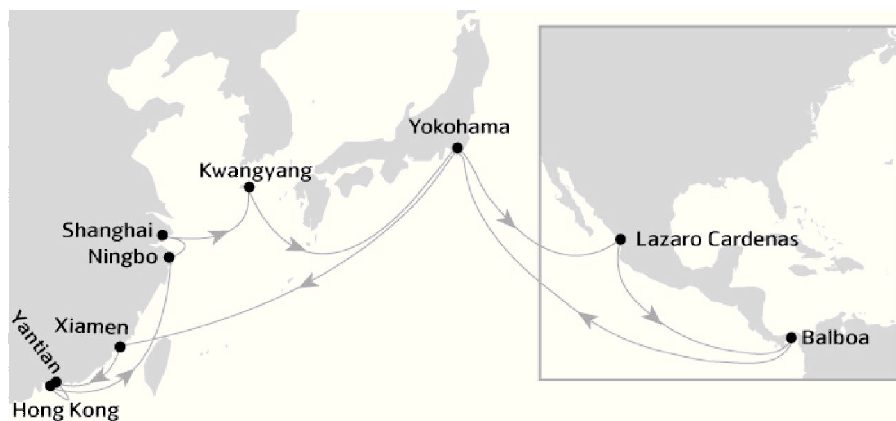


Figure 6.1: A trans-pacific round trip is depicted. Cargo is collected in transshipment ports in Asia and sailed to transshipment ports in Central America. The round trip takes 56 days implying that 8 vessels is required to maintain a weekly service. Feeder vessels are used to connect all ports in a geographical area.



Figure 6.2: A feeder service collect containers in the hub in Bremerhaven and transport them to their destinations in Norway.

Currently when a disruption occur, the operator at the shipping companies manually decides what action to take. For a single delayed vessel a simple approach could be to speed up. However, the cost of bunker fuel is a cubic function of speed [1] and vessels' speeds are limited between a lower and upper limit. So even though an expensive speed increase strategy is chosen, a vessel can arrive late for connections, propagation delays to other parts of the network.

In recent years liner companies have had an increased focus on minimizing the bunker consumption in order to provide environmentally friendly transport and to minimize the operational costs [21]. On the other hand, on time delivery is very important for a global liner shipping company as delayed cargo carries a high cost by customers and key clients. Nevertheless, the negative effects of miss-connections or delaying a key clients merchandise can be hard to measure against a concrete cost of for example bunker. Furthermore, the ripple effect of the recovery on to the remaining network is very complex to overview for a human. In the considered example Maersk Sarnia recovered the situation by a general speed increase with a high bunker cost, but nevertheless the speed increase did not ensure timely delivery of containers to the hub port of Balboa, and final recovery was done returning to Asia. As a result all the cargo was delayed and some cargo missed the onward connection at the hub. The mathematical model presented in this paper suggested omitting the last port call in Asia reaching the transshipment port without increasing the vessel speed and on time. The cost saving, including a delay penalty, of the suggested solution is more than 20 %.

A standardized way of handling disruptions based on mathematical grounded decision support may significantly lower the cost of handling disruptions as seen in the airline industry [25, 37] and simplify implementation of strategic decisions among stakeholders. According to UNCTAD [32] *slow-steaming* has resulted in a significant increase in delays and they expect carriers to resume higher speeds in order to increase reliability and productivity. According to Notteboom [23] reliability is generally achieved by introducing sufficient buffer time into a service. We believe that a mathematical decision support tool as the one presented in this paper may result in sustaining a *slow-steaming* policy, while increasing reliability of service without the need to introduce additional buffer time. In this paper, we introduce a mathematical model for handling the most common disruptions in liner shipping called the Vessel Schedule Recovery Problem, VSRP.

We make four contributions: First, we propose a novel formulation for the VSRP inspired by similar models within the airline industry. To the best of our knowledge the present article is the first to apply optimization to handle disruption management within the domain of liner shipping networks. Secondly, We prove the VSRP to be NP-complete. Third, we report computational results for four cases representing common disruptions, selected by experienced personnel at Maersk Line Operations Center. The recovery options identified by the mathematical model are comparable or superior to the decisions implemented in real life with cost savings of as much as 58%. The model is solved by a MIP solver within seconds for the selected cases. Fourth, a set of generic test instances is used to provide insights into the network sizes that may be handled in seconds by the current model and solution methods.

The remainder of the paper is organized as follows. Section 6.2 introduces disruption management in the liner shipping business. Section 6.3 describes related literature. In Section 6.4 we introduce the *Vessel Schedule Recovery Problem (VSRP)*, the graph topology, and a mathematical model for the VSRP along with proofs of the \mathcal{NP} -completeness of the problem. In Section 6.5 we introduce the four real life cases and the generic test instances and report computational results. Following this section we conclude that a decision support tool based on mathematical optimization of a disruption scenario could greatly aid an operations manager in evaluating the different recovery options.

6.2 The Liner Shipping Business

Liner shipping of containers is the backbone of world trade. Even though containerization simplifies the operations and reduces the cost per transported unit, the earned return is less than 10% on assets [27]. Customers demand fast and reliable delivery, while the shipping companies

constantly search to cut costs. These issues have motivated major investments in improving the daily operations at large shipping companies [23]. The liner shipping company referred to as a *carrier* has a public schedule of services. A service consists of a cyclic route with a scheduled time for each port call en route. Containers travel through the network as passengers in a public transit network, often combining several services. The port calls of a service, must usually happen at a predefined time and place in the port, often called the *berth slot*. This is defined by the physical place that the vessels moors, the berth, and a time window where the vessel is serviced. Most carriers provide weekly frequency of port calls. In recent years major companies are using slow-steaming to lower the variable cost and the CO₂ emission [20, 26, 21]. To stay competitive, research has been focused on designing the network to operate as efficiently as possible. For shipping companies, a division of the ports into *hubs and spokes* is common[8]. The network is not a traditional *hub and spoke network design* with direct links between two hubs or a spoke and a hub. As an alternative large vessels operate *main lines* between a set of hub ports and smaller vessels operate *feeder lines* connecting a set of spokes to a hub. An example of a *main line* service between hub ports is given in Figure 6.1 and an example of a *feeder* service servicing a hub and several spokes is given in Figure 6.2.

The motivation for this hub-and-spoke network design is to benefit from the economies of scale on container vessels [27]. The majority of containers are transhipped at least once during transport adding to the operational complexity and the impact of a disruption. Liner shipping companies operate with a head haul and a back haul direction. In the head haul direction vessels are almost full as opposed to the back haul direction. The head haul generally generates the majority of the revenue retrieved by operating the full service. As described above disruptions are accounted for and handled in the network by adding buffer time. Customer demand for fast delivery results in increased speeds and nearly no buffer time on the head haul, whereas the back haul is slower and has more buffer time. Due to the complexity of recovering from a disruption additional buffer time is included on the back haul with the option of a slight speed increase to catch up with the schedule on the back haul.

The most important variable costs in a liner shipping network is the bunker cost, the cost of using passageways such as the Suez and Panama canals, and the cost of calling ports to load and unload cargo. The fixed cost of operating a network in terms of asset costs on vessels, containers, and equipment are significant. Whenever a vessel fails to operate in accordance with the original schedule it is hurting the shipping company's business (and the business of their customers) [23]. The utilization of vessels will often be affected negatively as containers miss-connect, resulting in a higher cost per transported unit. Furthermore, it might be necessary to arrange alternative transport for the miss-connected units also adding to the cost. Finally, the customers demand a reliable service and expect on time delivery. A major concern is therefore how to handle disruptions when they occur.

For larger liner shipping companies the information about disruptions are gathered in the company's Operational Control Center (OCC), from where decisions are also taken with respect to how the disruptions should be handled. Decisions here are taken in real-time and any system to support this process should support real-time decision making. The reason for this is two-fold. 1) Weather is changing quickly in some parts of the world, which may cause a port to close for a period of time. In such a case it is important to make a reasonable quick decision regarding whether the port should be skipped, which typically will lead to a change of course and the possibility of slowing down and saving on bunker fuel. 2) The other and more important reason is that controllers working in the OCC are in some periods faced with the need for taking many decision and evaluating various alternatives. This is where the requirement for a quick response becomes imperative. For this reason controllers at Maersk have stated 10 seconds as a reasonable response time for a disruption management system.

6.2.1 From airline disruption to liner shipping disruption

Operations research has for many years been applied extensively in the airline industry [4]. Initially OR was mainly used in the planning phase, but during the last two decades OR has also found

its way into the disruption management tools, which are used on the day of operation where the planned schedule is being executed.

This paper focuses on utilizing the findings in disruption management tools for the airline industry in order to construct a mathematical model of the VSRP to handle disruptions in the context of the liner shipping business. The airline and liner shipping businesses have evident similarities, but also some core differences [7]. Larger airlines and larger liner shipping companies both operate a hub and spoke network, where either passengers or containers need to flow from an origin, through one or more hubs to a destination. Here, they need to arrive with the least possible amount of delay. In this way vessels resemble aircraft and containers resemble passengers. While crew recovery is a significant part of disruption management for an airline, this is not the case for a liner shipping company, as crew always follow the vessel and do not have work rules, which significantly limit the utilization of the vessel. Traditional aircraft recovery as described by Thengvall et al. [30] or Dienst et al. [10] makes use of 3 recovery techniques: *Delays*, *Swaps* and *Cancellations*. In addition to these techniques Marla et al. [22] show that a large improvement in the number of passenger miss-connections can be obtained if *speed-changes* are included as a fourth recovery technique. In the following we discuss how each of these techniques can be applied to disruption management in a liner shipping network:

- *Delays*. For an airline the most straight forward way of handling a disruption is to delay flights and let the delays propagate to the subsequent flights of an aircraft. After a number of delay propagations the initial delay will have disappeared due to the fact that the gap between flights is usually a bit longer than the required turn time and most aircrafts are idle over night. For an airline this recovery technique is unfortunately also the one which, when applied alone, often ends up causing a lot of miss-connections [10]. In liner shipping it is also possible to delay the departure of a vessel, but port calls do not have additional slack built into them and container vessels are constantly in service, which means that delay propagation will not be able to resolve a disruption on its own. It will need to be combined with some of the techniques presented below in order to have the desired effect of recovering from a disruption.
- *Swaps*. This is a very efficient recovery technique for an airline, as it can be used to eliminate a lot of delay propagation to subsequent flights. Swaps are possible as an aircraft becomes empty after each flight. As a result one aircraft may be substituted for another. Unfortunately, this technique is not applicable to a liner shipping company, as a container vessel servicing a certain service is never empty and it is both extremely costly and time consuming to empty it completely. While vessels cannot be swapped in the VSRP it is for a liner shipping company possible to swap the order in which ports are being visited, whenever these ports are located geographically close to each other.
- *Cancellations*. This technique is usually not preferred in the aircraft recovery problem, but it is an efficient way of recovering, whenever the airline experience large delays or reduced runway capacity. For a liner shipping company this technique is unfortunately not directly applicable as it would interrupt the service operation of the vessel. In the VSRP it is however possible to cancel or omit a port call. In this case containers, which are destined for the omitted port, are then off-loaded at a subsequent nearby port and containers for on-loading in the omitted port are being held for the next vessel on that service, or another service covering the same ports, which often results in a delay of up to a week.
- *Speed changes*. Including speed changes from a network perspective as an integrated part of disruption management turns out to be a very effective way of balancing passenger delays versus fuel burn for an airline [22]. This is in spite of the fact that a flight usually can only be sped up with 8-10% compared to its planned speed. For a vessel, which is originally scheduled to sail at a slow steaming speed of e.g. 16-18 knots, it is possible to speed up with 40% to e.g. 22-24 knots. This additional speed flexibility may be promising for the application of this technique in a liner shipping network.

As it is seen there are some clear similarities in the techniques, which can be applied in recovering a disruption in an airline network, and the techniques, which can be applied in recovering a disruption in a liner shipping network. The aircraft swapping technique available to an airline provides increased interaction between aircrafts in an airline network as opposed to vessels in a liner shipping network. An additional complication in a liner shipping network is that vessels operate around the clock and cannot naturally recover by using some of the overnight slack, which is often available in an airline network. For this reason recovering from a liner shipping disruption may take days and even weeks as opposed to a typical maximum of 48 hours for airlines. If a container fails to connect to a succeeding vessel the impact will often be more severe in liner shipping. International airports have a number of daily departures for a given destination presenting the option to re-accommodate passengers with a slight delay on a subsequent flight. For liner shipping a missed connection will normally result in a major delay.

We must estimate the effect on the cargo onboard with regards to missed onward connections and delays in order to assess a given recovery plan. Ideally the container groups would be reflowed on the residual capacity of the entire liner shipping network simultaneously with a recovery plan for the delayed vessels. This would significantly increase the graph of our instance as the containers onboard will include services, not considered in the disruption scenario. Additionally, reflowing the cargo is a large scale multicommodity flow problem. Mathematical models incorporating a large multicommodity flow problem such as capacitated network design [14] and liner shipping network design [2] are severely restrained by the size of the problem and excessive solution times for general MIP solvers. We expect similar issues if incorporating the reflow of miss-connected containers into the VSRP and most certainly the application will no longer be able to provide real time suggestions when considering reflowing containers on the residual capacity of the network in a joint optimization. This is furthermore supported by findings in the airline literature where Bratu and Barnhart [5] concludes that a combined model for solving a combined aircraft recovery and passenger re-accommodation model is too complex to solve to make it useful for real time optimization. Similarly the review Clausen et al. [9] shows that full passenger re-accommodation is always handled in a subsequent optimization phase. An approach, which has been useful in the airline industry [22] is not to solve the full passenger re-accommodation problem together with aircraft recovery, but rather let the aircraft recovery be guided towards passenger friendly solutions by penalizing misconnecting passengers. A similar approach could be deployed for disrupted containers.

6.3 Literature review

Notteboom [23] analyze the negative effects of disruptions in liner shipping and the actions taken by liner shipping companies to mitigate them. The recent paper by Notteboom and Vernimmen [24] demonstrates how the increased bunker price has a significant impact on the liner shipping business. The cost of fuel is a dominant cost driver when transporting containers, nevertheless shipping companies are willing to burn extra fuel to arrive according to the schedule. Disruption management is a major concern for liner shippers given this trade-off. Notteboom and Vernimmen [24] argue that the increased price on bunker has resulted in lowering the speed of vessels to save fuel, which in turn gives the vessels more buffer time and the operators more possibilities to recover from a disruption.

Even though the research within maritime transportation has gained increased focus during the last decades, we have encountered no journal papers devoted to disruption management in (liner) shipping. This can be caused by various things; firstly as mentioned the usage of mathematical modeling in maritime transportation is still in its infancy and secondly the market of liner shipping is extremely competitive. The development of decision support software will often be carried out for a major player in the market and therefore not necessarily published. After the submission of this article another model on disruption management in liner shipping was published in the thesis of [17]. A heuristic is presented for solving a relaxed version of the model and computational results are provided for a set of generated disruption scenarios. The work by Yang et al. [35]

and Li et al. [19] addresses disruption management for berth allocation in container terminals. Their papers are focused on how to recover the berthing schedules when vessels are delayed from the terminal point of view. Yang et al. [35] presents an MIP Model and a heuristic solution approach. The problem handled is very different from the VSRP dealing with disruptions from the carriers point of view. The work of Du et al. [12] allocates berths considering fuel consumption and has a good review on other Berth allocation literature. Well-established OR departments at many airlines have addressed the severe economical impact of flight delays and how to mitigate the effects of delays through disruption management based on OR. In 2008 the Joint Economic Committee under the U.S. Congress published a report estimating the infused cost to the American society to more than \$40 billion [16]. The order of magnitude of the cost of disruptions has later been confirmed in a more theoretically profound study by Ball et al. [3] even though their final estimate is $\approx 20\%$ lower. Both Rakshit et al. [25] and Yu et al. [37] document significant savings by implementing real-time decision support systems to handle the disruptions at major US airlines where the later estimates the annual saving to amount to \$40 million for Continental Airlines.

Disruption management research for airlines generally deals with recovering the 3 resource areas aircraft, crew and passengers. The full problem of optimizing all of these areas simultaneously is, however, so complex that no work has been published so far, which cover all 3 areas in one single integrated model. Most of the published models address one single resource. A few of the models focus on one resource area, while including specific aspects of other areas. For a good general introduction to disruption management in the airline industry the reader is referred to Yu and Qi [36] and Barnhart [4]. The paper of Kohl et al. [18] describes a large scale EU-funded project, called Descartes, which addresses various aspects of disruption management for all 3 resource areas. The reader is also referred to an extensive survey of operations research used for disruption management in the airline industry by Clausen et al. [9]. In order to adapt disruption management techniques applied to the airline industry to the liner shipping industry the aircraft recovery problem resembles vessel recovery and the recovery of passenger itineraries resembles container recovery. Since liner shipping companies do not have to deal with crew recovery, this literature review will only focus on aircraft and passenger recovery.

The first model on the *Aircraft Schedule Recovery Problem*, presented in the literature, is a network flow model by Teodorović and Guberinić [28], who contributed by solving small problems with 3 aircraft and 8 flights. This work was extended by Teodorović and Stojković who extended the model in later papers. The solvable problem sizes still remained small with 14 aircraft and 80 flights. Jarrah et al. [15] presented the first work, which were applicable in practice based on instances from United Airlines. They published 2 models, which in combination were capable of producing solutions handling all 3 traditional recovery techniques delays, swaps and cancellations. The drawback of handling this in 2 separate models was that delays and cancellations could not be traded off against each other. This drawback was resolved in the work by Yan and Yang [34] who were capable of trading off delays, swaps and cancellations in one single model based on a time-line network. Thengvall et al. [30] extended this model to also include so-called protection arcs, which serve the purpose of keeping the proposed solutions somewhat similar to the original schedule. This is important for real-life application of the suggested solutions as an unlimited number of changes cannot be applied to the schedule last minute. The work by Dienst et al. [10] extends this model to also cover aircraft specific maintenances and preferences in an aircraft specific recovery model.

The *Passenger Recovery Problem* is an area of disruption management, which has been addressed to a rather limited extent by published research. Our observation from airlines show that most of these use a sequential re-accommodation process, which is carried out after an aircraft recovery schedule has been decided upon. Vaaben and Alves [33] do a comparison of sequential passenger re-accommodation with re-accommodation based on an MIP-model. The main contribution in the area of passenger recovery is done by Bratu and Barnhart [5], who present two models. Both are basically aircraft recovery models with some crew recovery guidance. One of them also includes passenger recovery, but is not solvable in real time. The other one is solvable but does not include complete passenger recovery. Instead it penalizes passenger miss-connections.

The work by Marla et al. [22] extends on the work by Dienst et al. [10] and Bratu and Barnhart

[5] by doing aircraft specific recovery with penalized passenger miss-connections, while at the same time also introducing the additional recovery technique of *speed changes*, which enables the model to balance the trade-off between passenger delay cost and fuel burn cost in a network perspective. The purpose of the present paper is to investigate if the application of similar disruption recovery techniques in a liner shipping context will be beneficial.

6.4 The Vessel Schedule Recovery Problem - (VSRP)

A given disruption scenario consists of a set of vessels V , a set of ports P , and a time horizon consisting of discrete timeslots $t \in T$. The time slots are discretized on port basis as terminal crews handling the cargo operate in shifts, which are paid for in full, even if arriving in the middle of a shift. Hence we only allow vessels arriving at the beginning of shifts. Reducing the graph to timeslots based on these shifts, also has the advantage of reducing the graph size, although this is a minor simplification of the problem. For each vessel $v \in V$, the current location and a planned schedule consisting of an ordered set of port calls $H_v \subseteq P$ are known within the recovery horizon, a port call A can precede a port call B , $A < B$ in H_v . A set of possible sailings, i.e. directed edges, L_h are said to *cover* a port call $h \in H_v$. Each L_h represent a sailing with a different speed.

The recovery horizon, T , is an input to the model given by the user, based on the disruption in question. Inter continental services will often recover by speeding during ocean crossing, making the arrival at first port after an ocean crossing a good horizon, severe disruptions might require two ocean crossings. Feeders recovering at arrival to their hub port call would save many missed transshipments giving an obvious horizon. In combination with a limited geographical dimension this ensures that the disruption does not spread to the entire network.

The disruption scenario includes a set of container groups C with planned transportation scenarios on the schedules of V . A feasible solution to an instance of the VSRP is to find a sailing for each $v \in V$ starting at the current position of v and ending on the planned schedule no later than the time of the recovery horizon. The solution must respect the minimum and maximum speed of the vessel and the constraints defined regarding ports allowed for omission or port call swaps. The optimal solution is the feasible solution of minimum cost, when considering the cost of sailing in terms of bunker and port fees along with a strategic penalty on container groups not delivered “on-time” or misconnecting altogether.

6.4.1 Graph topology

A disruption scenario is conceptualized as a directed graph in a time-space network similar to the one used by Thengvall et al. [29, 30, 31], Marla et al. [22] and Dienst et al. [10]. The horizontal axis corresponds to a point in time within the given planning horizon, and the vertical axis corresponds to a geographical position; a port in the context of VSRP. A simple example of a time-space network is presented in Figure 6.3(a). Here, two geographical positions are given and a vessel can connect from the initial position A to the next position B with three different speeds.

A directed graph $G = (N, E)$ with node set $N = \{p^t \in N | p \in P, t \in T\}$ where p^t denotes port p at time t representing the time-space network. n^- and n^+ denotes the in- and out-going edges of node $n \in N$ respectively. $N_v \subseteq N$ is the set of all nodes for vessel $v \in V$. The set consists of a source node n_s^v corresponding to the current position of the vessel and a sink node n_t^v corresponding to the scheduled position at the end of the recovery horizon. Additional nodes are created for the set of port calls $h \in H_v$ within a time window of $\{a_v^h, b_v^h\}$ defining the earliest and latest arrival time respectively given the vessels minimum and maximum speed, the current position and the remaining set of port calls.

Define the edge set $E = E_s \cup E_g$ where E_s represents a sailing of a vessel $v \in V$ such that $E_s = \{(p^t, q^{t'}) | p^t, q^{t'} \in N, p \neq q, t \leq t'\}$ and $E_g = \{(p^t, p^{t'}) | p^t, p^{t'} \in N, t < t'\}$. The duration of a port call is fixed for each vessel $v \in V$ according to the scheduled port call duration from the original schedule. Because the port call duration is fixed port call edges E_g are included in the sailing edges E_s , thereby removing the set E_g as seen in Figure 6.3(b). Including the edge set E_g

in E_s reduces the number of columns in the mathematical model. For illustrative purposes the port call edges are still visualized in Figures 6.3(c) and 6.3(d), while the remainder of the figures in this paper only visualize the combined edges.

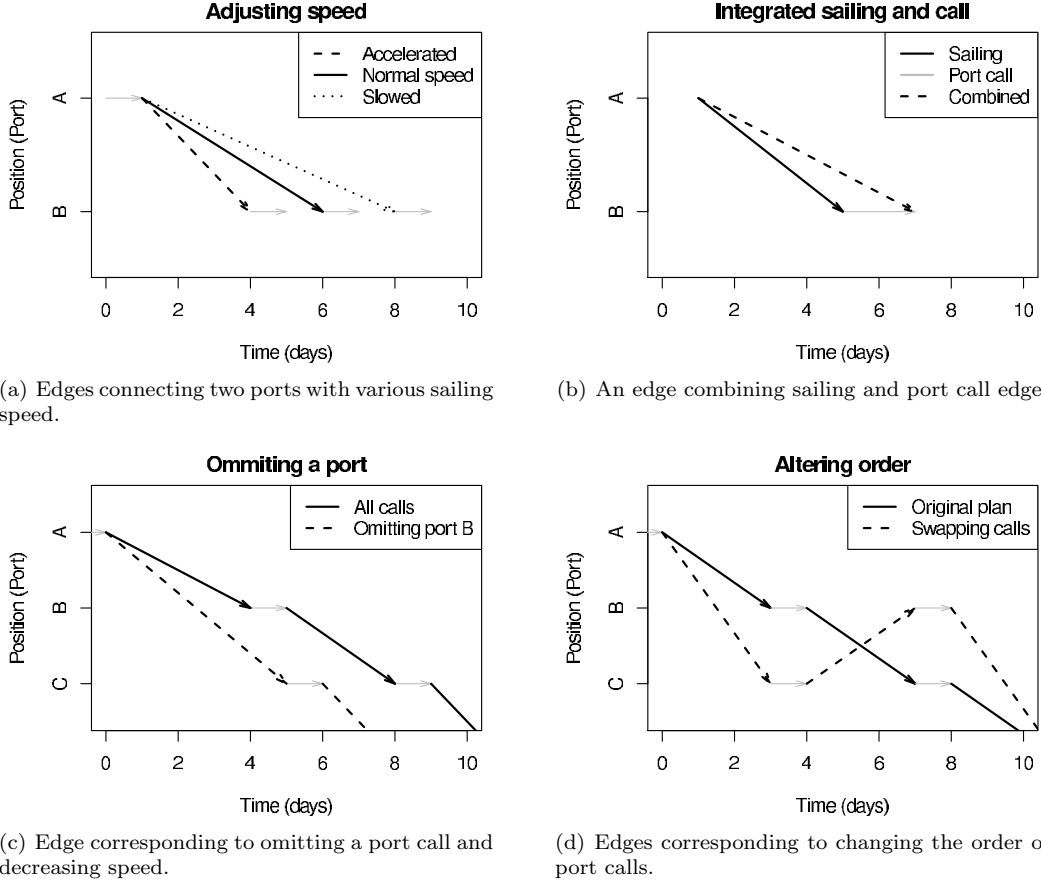


Figure 6.3: Possible moves in the time-space network model. Port call edges are gray.

The edge sets $E_v \subseteq E_s$ are the edges that define feasible sailings among the nodes of N_v for a given vessel $v \in V$. $c_e^v \in \mathbb{R}_+$ is the cost of using edge $e \in E_v$ for vessel type $v \in V$ consisting of the bunker cost at a given speed and port fee for port $p = target(e)$. t_e^v is the time it takes to traverse edge $e \in E_s$ given speed, distance and port call time. The edge set $E_s = \bigcup_{v \in V} E_v$ is defined according to the planned schedule and the possible recovery actions defined below:

- **Adjusting vessel speed** (Figure 6.3(a))
In the span of the minimum and maximum speed of vessel $v \in V$ several edges may connect ports A and B. Define the set of edges $L_h \subset E_v$ covering port call $h \in H_v$ as $L_h = \{(A^t, B^{t'}) | A, B \in H_v, A < B, t \leq b_v^A, a_v^B \leq t' \leq b_v^B, t < t', \forall t' = a_v^B + K \cdot \delta_B\}$ where K is a positive integer denoting the shift and δ_B is the duration of a shift at terminal B.
- **Omitting a port call** (Figure 6.3(c))
Vessels might omit port calls to recover a delay or simply to save the port cost. Omitting port calls will result in miss-connected containers. Allowing to omit port B on a sailing from port A via port B to port C corresponds to having an edge (A^{t_A}, C^{t_C}) where $t_C - t_A$ corresponds to the sailing time. Edges L_h with differing sail speeds must be created as described in above bullet.
- **Swap order of calls** (Figure 6.3(d))

In some cases, a delayed vessel needs to call a number of ports close to each other. It might be possible to swap port calls within a designated geographical area. In the time-space network a swap is included by adding, first an omitting edge, followed by an edge back to the original port call. Again this must be executed for differing vessels speeds, as described in first bullet.

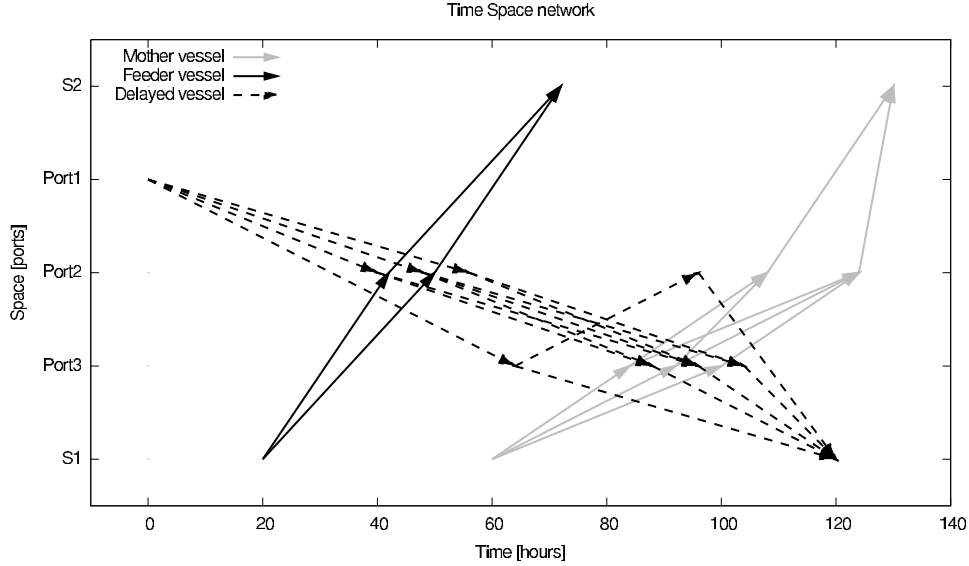


Figure 6.4: Example of a time-space network for a test problem with three vessels, sinks and sources, three ports, speed adjusting edges, and port swap for the *delayed* vessel. In the network only the edges taking part in a feasible path are shown.

Figure 6.4 gives an example of a full time-space network for a small test instance. Three vessels are affected by the delay of the *delayed* vessel.

The set of vessels is

$$V = \{ \textit{delayed}, \textit{feeder}, \textit{mother} \} = \{ d, f, m \}$$

and for each vessel a set of port calls is given. These are

$$H_d = \{ P_1, P_2, P_3, S_1 \}$$

$$H_f = \{ S_1, P_2, S_2 \}$$

$$H_m = \{ S_1, P_3, P_2, S_2 \}$$

where P_i is Port i and S_i corresponds to onward sailing according to schedule. For each of the port calls $h \in H_v$ a set of possible sailings L_h covering the call is given. As an example vessel d has the set of four possible sailings/legs covering the call in Port 2:

$$L_{(d,P_2)} = \left\{ \begin{array}{l} (P_1, 0) \rightarrow (P_2, 38) \quad , \quad (P_1, 0) \rightarrow (P_2, 48) \quad , \\ (P_1, 0) \rightarrow (P_2, 58) \quad , \quad (P_3, 62) \rightarrow (P_2, 98) \quad \} \quad . \end{array} \right.$$

The cost of each of these edges is the sum of the bunker cost from sailing with the necessary speed between the ports and the cost of calling Port 2. The cost of using leg $(P_1, 0) \rightarrow (P_2, 38)$ is higher than the cost of using leg $(P_1, 0) \rightarrow (P_2, 58)$ as the sailing time is smaller ($38 < 58$) resulting in a higher sailing speed and consequently an increased bunker fuel burn.

The problem has characteristics that are not directly reflected in the graph. These are the flow of containers, extended port stays due to omissions, limits on the capacity of a port, and port closure in a period of time. The extended port stay due to an omission can readily be handled in the graph construction by adjusting the duration of the set of sailing edges in E_s , that represent the

omission. This has not been done to simplify modeling, as the effect will small. The port capacity issue can be modeled by constraining the number of vessels arriving (or used legs) at each port in each given time interval. Port closures are included by removing all edges corresponding to arriving at a port while it is closed.

6.4.2 Transportation scenarios - the impact of a recovery on the affected cargo

In order to evaluate which container groups will suffer from missed onward connections and delays we define a transportation scenario for each container group in terms of their origin, destination and planned transshipment points. $B_c \in H_v$ is defined as the origin port for a container group $c \in C$ and the port call where vessel v picks up the container group. Similarly, we define $T_c \in H_w$ as the destination port for container group $c \in C$ and the port call where vessel w delivers the container group. Intermediate planned transshipment points for each container group $c \in C$ are defined by the ordered set $I_c = (I_c^1, \dots, I_c^m)$. Here $I_c^i = (h_v^i, h_w^i) \in (H_v, H_w)$ is a pair of calls for different vessels ($v, w \in V | v \neq w$) constituting a transshipment. Each container group c has m^c transshipments. M_c^e is the set of all non-connecting edges of $e \in L_h$ that result in miss-connection of container group $c \in C$. $c_c^d \in \mathbb{R}_+$ is the cost of a delay to container group $c \in C$ exceeding a day of the planned arrival and $c_c^m \in \mathbb{R}_+$ is the cost of one or several misconnections to container group $c \in C$, which is added to the delay penalty in the model.

The cost of delaying the arrival of a container at its destination is to a large extent related to the loss of goodwill from the affected customers. This may vary by the type of container and the importance of the customer to the liner shipping company. In general refrigerated containers are more costly to delay than non-refrigerated, but more detailed classification by container type and customer value may be applied. The cost classifications used in the case-studies in this paper have been supplied by Maersk Line and are based on their internal approximations of these costs.

6.4.3 Mathematical model

The mathematical model is inspired by the work within aircraft recovery with speed-changes by Marla et al. [22]. Like others before Marla et al. (e.g. Marla et al. [22] and Dienst et al. [10]) we use a time space graph as the underlying network, but reformulate the model to address the set of available recovery techniques, which are applicable to the VSRP.

Define binary variables (x_e) for each edge $e \in E_s$ set to 1 iff the edge is sailed in the solution. Define binary variables (z_h) for each port call $h \in H_v \quad \forall v \in V$ set to 1 iff call h is omitted. For each container group c we define binary variables $o_c \in \{0, 1\}$ indicating whether the container group is delayed or not and y_c to account for container groups misconnecting. $O_e^c \in \{0, 1\}$ is a constant set to 1 iff container group $c \in C$ is delayed when arriving by edge $e \in L_{T_c}$. $M_c \in \mathbb{Z}_+$ is an upper bound on the number of transshipments for container group $c \in C$.

$$S_v^n = \begin{cases} -1 & , n = n_s^v \\ 1 & , n = n_t^v \\ 0 & \text{Otherwise} \end{cases}$$

is applied to the flow conservation constraints.

Minimize:

$$\sum_{v \in V} \sum_{h \in H_v} \sum_{e \in L_h} c_e^v x_e + \sum_{c \in C} [c_c^m y_c + c_c^d o_c] \quad (6.1)$$

Subject To:

$$\sum_{e \in L_h} x_e + z_h = 1 \quad \forall v \in V, h \in H_v \quad (6.2)$$

$$\sum_{e \in n^-} x_e - \sum_{e \in n^+} x_e = S_v^n \quad \forall v \in V, n \in N_v \quad (6.3)$$

$$y_c \leq o_c \quad \forall c \in C \quad (6.4)$$

$$\sum_{e \in L_{T_c}} O_e^c x_e \leq o_c \quad \forall c \in C \quad (6.5)$$

$$z_h \leq y_c \quad \forall c \in C, \forall h \in B_c \cup I_c \cup T_c \quad (6.6)$$

$$x_e + \sum_{\lambda \in M_e^e} x_\lambda \leq 1 + y_c \quad \forall c \in C, e \in \{L_h | h \in B_c \cup I_c \cup T_c\} \quad (6.7)$$

$$x_e \in \{0, 1\} \quad \forall e \in E_s \quad (6.8)$$

$$z_h \in \mathbb{R}_+ \quad \forall v \in V, h \in H_v \quad (6.9)$$

$$y_c, o_c \in \mathbb{R}_+ \quad \forall c \in C \quad (6.10)$$

The objective function (6.1) minimizes the cost of operating vessels at the given speeds, the port calls performed along with the penalties incurred from delaying or misconnecting cargo. The weighted sum scalarization [13], the ϵ -constraint method [13], and variable fixing has been implemented for the VSRP with promising results in the thesis by Dirksen [11].

Constraints (6.2) are *Set-Partitioning* constraints ensuring that each scheduled port call for each vessel is either called by some sailing or omitted. (6.3) are *Flow-Conservation* constraints. Combined with the binary domain of variables x_e and z_h they define feasible vessel flows through the time-space network. A misconnection is by definition also a delay of a container group and hence the misconnection penalty is added to the delay penalty. This is expressed in (6.4).

Each container group has a planned arrival time upon which it can be decided whether or not a given sailing to the destination will cause the containers to be delayed. Constraints (6.5) ensure that o_c takes the value 1 iff container group c is delayed when arriving via the sailing represented by edge $e \in E_s$. The right hand side does not have to be multiplied despite the number of summed variables may be larger than one due to the cover constraint (6.2) as this constraint ensures that only one incoming edge $x_e, e \in L_{T_c}$ can have flow. Constraints (6.6) ensure that if a port call is omitted, which had a planned (un)load of container group $c \in C$, the container group is misconnected. Constraints (6.7) are coherence constraints ensuring the detection of container groups' miss-connections due to late arrivals in transshipment ports. For each of the possible inbound sailings of a container transshipment a constraint is generated. On the left-hand side the decision variable corresponding to a given sailing, x_e , is added to the sum of all decision variables corresponding to having onward sailing resulting in miss-connections, $\lambda \in M_e^e$.

The constraint is illustrated in Figure 6.5. When implementing the constraint the variables corresponding to inbound sailings are summed.

The variable x_e is required to be binary, whereas the remaining variables are only required to be non-negative. Binary x_e combined with constraints (6.2) implies z_h to be binary. Given the binary domains of x_e and z_h combined with constraints (6.6), (6.7) and a minimization implies y_c to be binary. Finally, Minimization, binary domains of x_e and y_c combined with constraints (6.4) and (6.5) imply that o_c is binary.

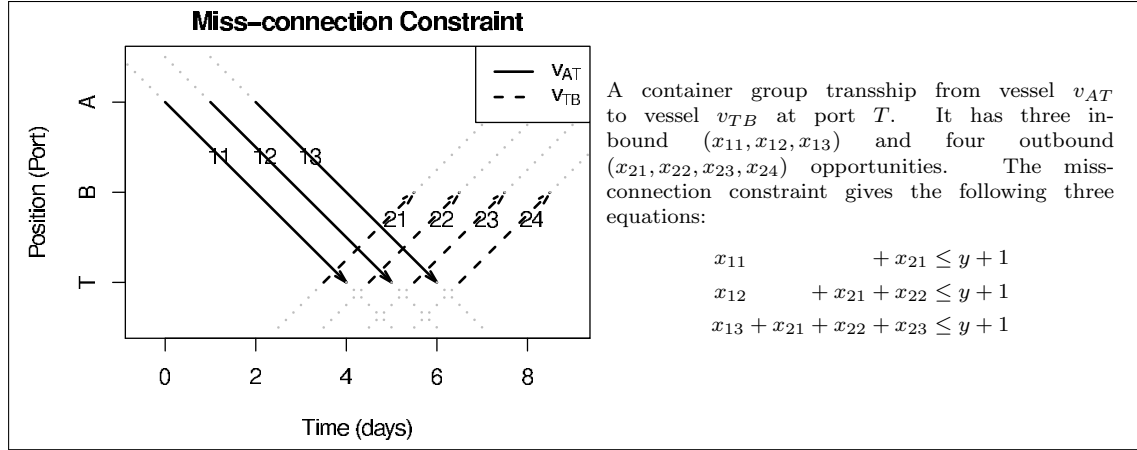


Figure 6.5: Example of the miss-connection constraint (6.7).

6.4.4 Model extensions

The model can be extended to incorporate additional features of a given problem instance such as the berth occupation constraint.

$$\sum_{v \in V} \sum_{h \in H_v} \sum_{e \in L_h} U_e^{pt} x_e \leq 1 \quad \forall p \in P, t \in T \quad (6.11)$$

$U_e^{pt} \in \{0, 1\}$ is a constant set to 1 iff edge $e \in L_h$ occupy a berth in port $p \in P$ in time slot $t \in T$. The constraint ensures that only a single vessel can enter and use a berth at a given time. This constraint will not handle berth allocation in general, which specified methods exist for, as mentioned in literature review. But when several vessels have to compete for a single berth available at a terminal, this constraint can be used to model the liner shipping company's choice of prioritization, irrespective of the terminal's options.

6.4.5 Complexity

The VSRP is NP-hard if omissions of ports is allowed, or if port swaps are allowed. Even if only one of the recovery actions is allowed, the problem is NP-hard as shown in the following: If omissions of ports are allowed in VSRP, the NP-hardness can be proved by reduction from the *0-1 Knapsack Problem (KP)*. Given an instance of the KP with a knapsack of capacity c , and n items having profit p_i and weight w_i , we transform it to an instance of the VSRP by using a single vessel and n ports which can be omitted. The cost of omitting a port is set to $-p_i$ and the duration of a port call is set to w_i . Sail times between ports are set to zero, and the recovery horizon is set to c , ensuring that a maximum profit subset of the items is chosen satisfying the capacity of the knapsack.

If port swaps are allowed in VSRP, the NP-hardness is shown by reduction from the *Traveling Salesman Problem (TSP)*. Given an instance of TSP with n nodes and edge costs c_{ij} , we construct an instance of the VSRP by introducing n ports which can be visited in arbitrary order. Port calls and travel times are set to zero, while the sail cost between ports is c_{ij} . The cost of omitting a port is set to infinity ensuring that all ports are visited following the shortest Hamiltonian cycle.

The above reductions prove that the VSRP with allowed omissions is weakly \mathcal{NP} -hard and the VSRP with multiple omissions to be strongly \mathcal{NP} -hard. Extended proofs for the \mathcal{NP} -completeness of the VSRP may be found in Dirksen [11].

6.5 Computational results

The program has been run on a MacBook Pro with 2.26 GHz processor and 2 GB of memory running Mac OSX using IBM ILOG CPLEX 12.2.0.0 as MIP solver. To test the performance and applicability of the developed model, it has been run on four real instances and a number of auto generated instances.

6.5.1 Real-life Cases from Maersk Line

The cases used to evaluate the VSRP are based on historical events at Maersk Line (ML). They are selected to represent the most common disruption scenarios and recovery options. Each case includes information about vessel schedules, port distances, container movements, recovery options, vessel speeds, and costs. ML handles these types of disruptions on a daily basis. The purpose of the cases is to test the suggested model, but also to clarify typical disruptions and how they are currently handled. An overview of the cases is given followed by a detailed presentation. The cases are

1. **A Delayed Vessel**

The vessel Maersk Sarnia is delayed out of Asia due to bad weather. The vessel is, filled with cargo, about to cross the Pacific Ocean and unload in Mexico and Panama.

2. **A Port Closure**

The port Le Havre in France is closed due to a strike. The vessel Maersk Eindhoven arriving with cargo from Asia can either wait for the port to open (giving an expected 48 hour delay) or omit the call in Le Havre.

3. **A Berth Prioritization**

The port in Jawaharlal Nehru (India) does not have the capacity for a ME3-service vessel and a MECL1-service vessel to port at the same time. As the MECL1-service vessel is delayed and the vessels will arrive at the port simultaneously, it is necessary to decide which vessel to handle first.

4. **Expected Congestion**

The feeder vessel Maersk Ravenna is planned to call three Colombian ports. Due to port maintenance at the last port to call, a delaying congestion is expected if arriving as planned. ML has to decide if the plan should be changed to avoid the congestion.

6.5.2 Case results

The computational results for the cases are promising. Good recovery strategies have been generated within 5 seconds, which proves the model applicable as a real-time decision support tool for liner shipping companies. The optimization based recovery strategies are generated with a strategic penalty for delaying and misconnecting containers. The two penalties are given the same value, i.e. $c_c^m = c_c^d$. For each of the cases discretization of the time horizon is $\delta = 3$ hours. Table 6.1 shows different size measures for the four cases. The results from the optimized runs (**OPT**) have been compared to the real life solution (**RS**). **RS** is the realized sailings for the affected vessels and the realized impact on containers. All presented costs are relative to the real cost to preserve the relativeness of bunker, port fees and container impact of a solution. However, the costs have no relation to real life costs.

An overview is given in Table 6.2. The results clearly show potential in the mathematical model. The experts at ML have indicated that in two out of the four cases they would prefer **OPT**, in one **OPT** is the same as **RS**, and in the last **RS** is preferred. However, in the case where **RS** is preferred, the recovery strategy is based on re-flowing cargo, which is not considered by the VSRP. The tendency is clearly that the model generates competitive solutions and would be a substantial support to the operator resulting in better recovery solutions using significantly less time. However, based on just four cases it is not reasonable to conclude that the optimized

solutions are generally superior. The computational times are less than 5 seconds with CPLEX consuming roughly half the execution time, while graph generation consumes the rest. Please note that Case 1 (A Delayed Vessel) has a much longer planning horizon than the remaining cases, which accounts for the increase in running times. Even for Case 1 the solution time is indeed acceptable for a operational application.

6.5.3 Case 1 (A Delayed Vessel)

Within the planning horizon of Case 1 Maersk Sarnia delivers containers to a single ML vessel in Lazaro Cardenas and seven ML vessels in Balboa. Each vessel may be delayed to the originally planned arrival time. The vessel Maersk Sarnia is allowed to omit Yokohama and either Lazaro Cardenas or Balboa. The **OPT** is structurally different to **RS**. Both are plotted in Figure 6.6. ML has chosen to call all ports with a speed increase (**RS**). However, the speed-up is not sufficient to reach the head haul ports in time. The optimized solutions (**OPT**) is to omit the call in Yokohama resulting in 400 misconnected containers while the remaining ports are called in time. The combined costs and penalties of **RS** are 24% higher than the costs and penalties of **OPT**. The experts at ML confirm that omitting Yokohama was a superior solution and note, that they were unable to convince a single important stakeholder of the superiority of this solution. It is very clear that the generalized mathematical assessment provided by a decision support tool would have been a strong argument in the discussion.

6.5.4 Case 2 (A Port Closure)

In Case 2 (A Port Closure) either Le Havre is called 48 hours delayed, or Rotterdam is called at the planned time. In Le Havre 649 containers need to be loaded and 1911 need to be unloaded. The time-space network of the case is presented in Figure 6.7.

Again **OPT** is different in structure compared to **RS** (Figure 6.8). However, as noted earlier **RS** is based on re-flowing containers not considered by the VSRP. Surprisingly, the data for the suggested solutions show that **OPT** is a better alternative with respect to cost. In real life the delay turned out to be 72 hours and a solution was obtained by allowing to merge two port calls. This option was not available to the model and hence the results are not comparable.

6.5.5 Case 3 (A Berth Prioritization)

In the third case, the additional berth occupation constraint (6.11) is added to ensure that the vessels call the port in India one at a time. The berth prioritization case is interesting as four of the connecting ML vessels may be delayed significantly and still reach their next port to call. **OPT** and **RS** result in the same solution presented in Figure 6.9. The runs confirm the decision of **RS** and verify the applicability of a decision support system in an operational setting, providing fast solutions. In this case the decision would have been reached in a matter of seconds as opposed to hours.

Case	V	PC	CG	C	RH	N	E	x_e	z_h	y_c / o_c	Constraints
1	8	26	23	5145	961	301	7073	7073	10	23	1706
2	6	22	19	12358	969	118	290	290	10	19	122
3	10	33	24	5671	548	171	411	411	13	24	221
4	1	5	6	838	166	103	416	416	3	6	300

Table 6.1: An overview of the relative sizes of the cases in terms of the number of vessels (**V**), the number of port calls in the scenario (**PC**), the number of container groups included (**CG**), the total number of containers (**C**), the recovery horizon in hours (**RH**), the size of the graph (N, E), and the number of variables (x_e, z_h, y_c, o_c).

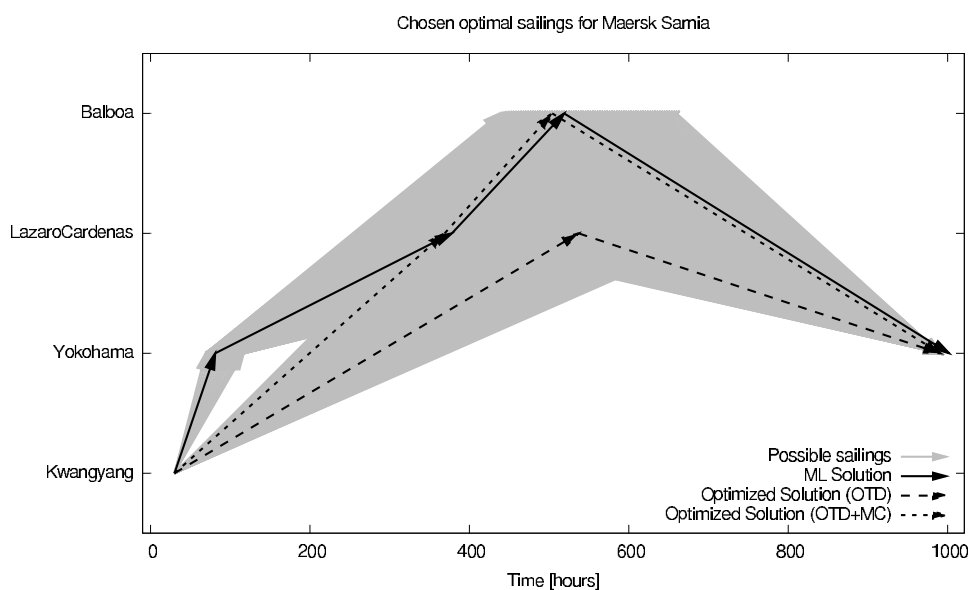


Figure 6.6: Case 1: Suggested recovery solutions for Case 1 (A Delayed Vessel).

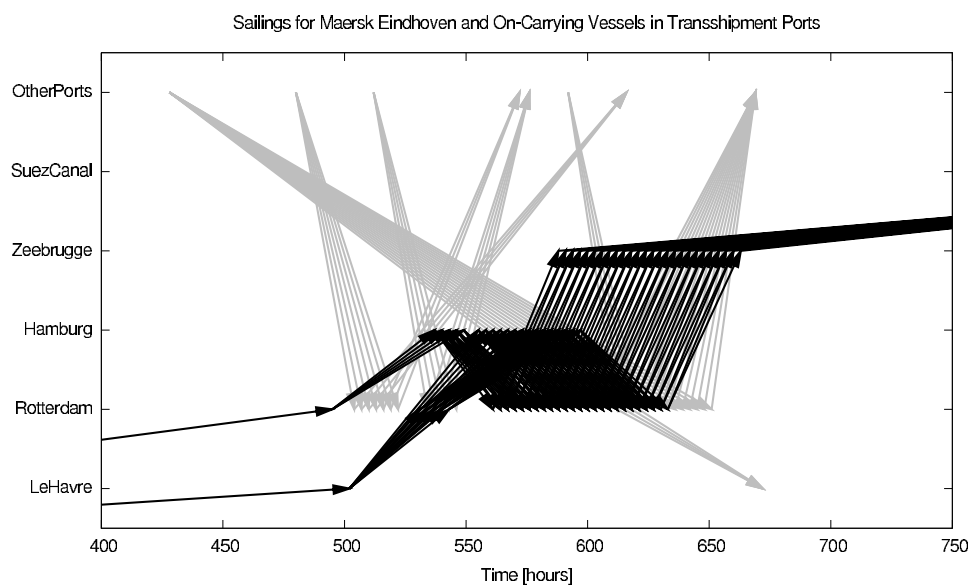


Figure 6.7: Time-space network for Case 2.

	Case	Sailing Cost		Delays		Misconnections		Solve Time	Best
		RS	OPT	RS	OPT	RS	OPT		
	1	1,000,000	914,063	(2449)	(0)	(26)	(400)	4.529	OPT
	2	1,000,000	977,392	(3111)	(3111)	(58)	(58)	0.718	RS
	3	1,000,000	1,000,000	(687)	(687)	(0)	(0)	0.681	Equal
	4	1,000,000	1,033,334	(222)	(0)	(0)	(0)	0.518	OPT

Table 6.2: Overview of results for the cases. The costs are relative, the container impact in units, and the time to solve in seconds. The best-column shows which solution the ML experts would prefer today.

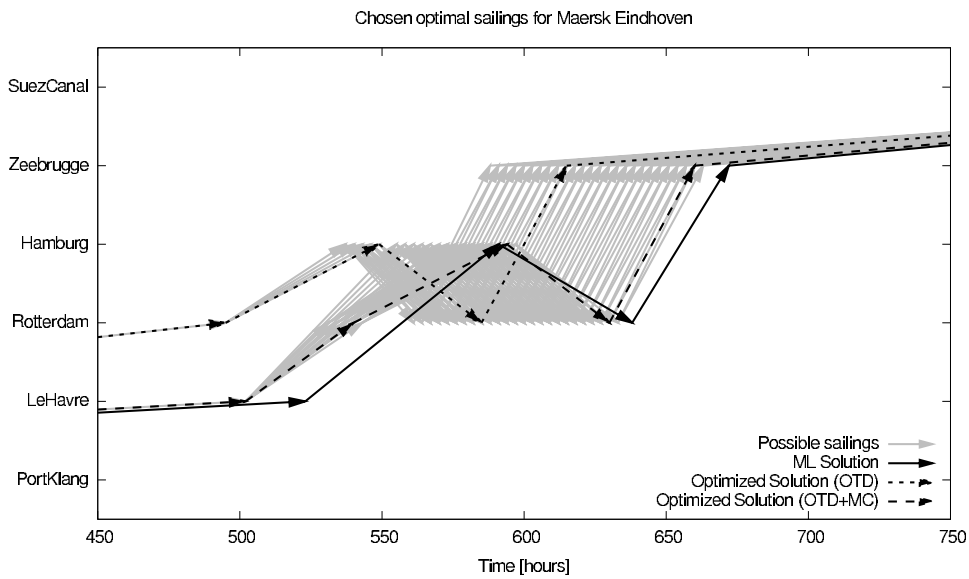


Figure 6.8: Suggested recovery solutions for Case 2.

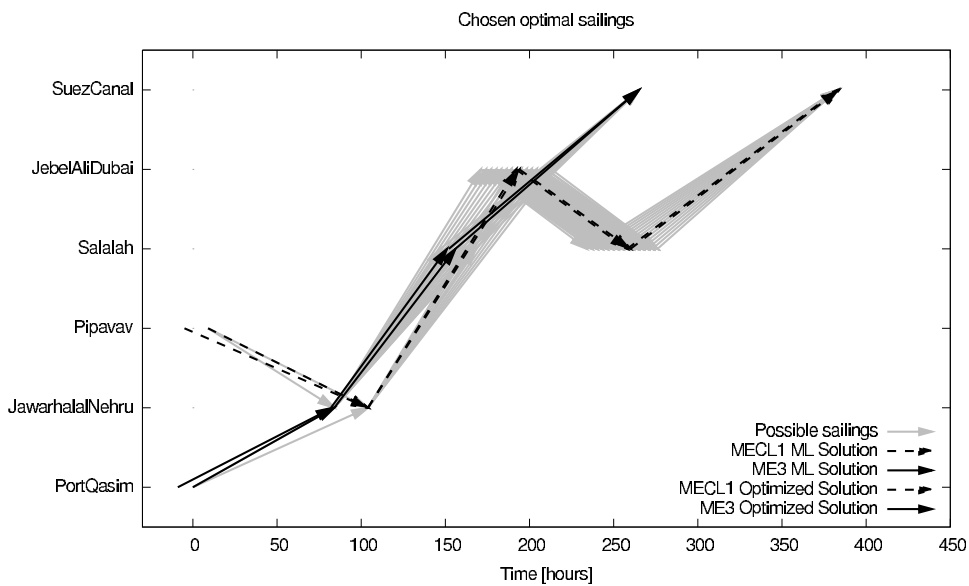


Figure 6.9: Time-space network and solution for Case 3. The ME3-vessel (full line) or the MECL1-vessel (dashed line) calls Jawaharlal Nehru in India first.

6.5.6 Case 4 (Expected Congestion)

The last case where a feeder vessel is expecting port congestions in the last port differs completely from the former cases. The feeder only carries direct import and export cargo to and from Colombia, meaning that no additional vessels need to be taken into account and that a single run is generated as misconnections are not possible. The expected port delay (of 24 hours if Santa Marta is called after $t = 100$) combined with the possibility of calling the three ports in Colombia in any order defines the problem. The time-space network of possible sailings along with the solutions is given in Figure 6.10. **RS** was to alter the order of the port calls to ensure that Santa Marta was visited long before the expected congestion. This resulted in a delay to the cargo in Cartagena. Contrary to **RS**, **OPT** suggests continuing as planned, but speeding up to

arrive at Santa Marta before the expected congestion. This solution displays slightly increased bunker cost but ensures that all containers are delivered on time. According to the experts at ML, the optimized solutions should have been implemented. The costs and penalties reveal a saving amounting to a stunning 58%.

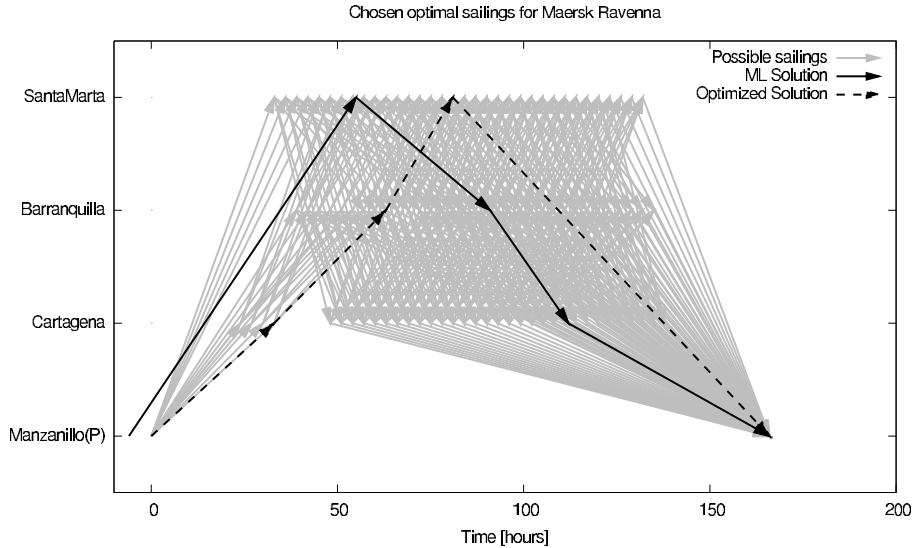


Figure 6.10: Suggested recovery solutions for Case 4 (Expected Congestion) in the time-space network.

6.5.7 Auto generated test instances

The four cases utilize different parts of the solution space satisfactorily, but lack in size and are thus relatively fast to solve. To test the scalability of the model, a set of random instances have been generated, refer to Figure 6.11 for an example. η^2 ports are placed in a squared grid, where distances and sailing times are proportional to Euclidean distances. Vessels are generated with a random schedule of $\kappa < \eta^2$ ports to call. Container itineraries are generated such that each intermediate call for each vessel and arriving container group is added with some probability. For these instances the computational time grow with increasing number of calls per vessel and number of vessels in instance as seen in Figure 6.12. It can be seen that the computational time handles an increased number of vessels well, but is impacted harder by an increased number of port calls. It seems viable that the model will solve in minutes for instances with up to 10 vessels and port calls making it viable for use in a wide range of real world problems. For more details on how the instances are generated and details on computational time please refer to the thesis by Dirksen [11].

6.6 Conclusion and future work

To the best of our knowledge this paper is the first literature on decision support for disruption management in a liner shipping network. We have presented a novel mathematical model for the Vessel Schedule Recovery Problem (VSRP). The model addresses frequently occurring disruption scenarios in the liner shipping industry. The model is based on disruption management work from airline industry and adapted to liner shipping. We show the VSRP to be NP-complete. The model is solved using a MIP solver and computational experiments indicate that the model can be solved within ten seconds for instances corresponding to a standard disruption scenario in a global liner shipping network. Computational results for four real-life cases show similar or improved solutions to historic data. The solutions have been verified by experienced planners. A

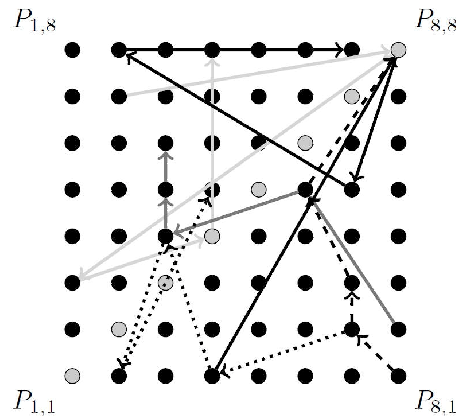
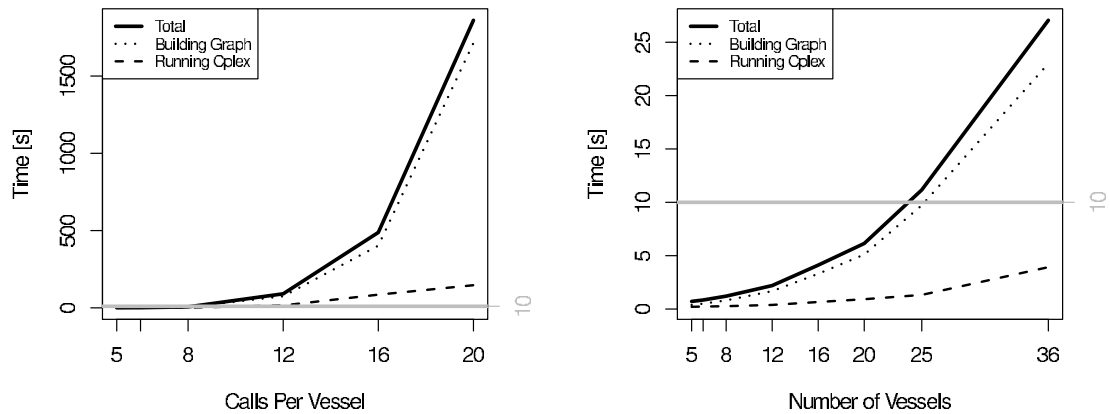


Figure 6.11: Graphical explanation of the standard way random instances of the VSRP are generated.



(a) Computational times increasing the number of port calls per vessel, with 5 vessels

(b) Computational times increasing the number of vessels, with 5 port calls

Figure 6.12: Computational times for generic generated problems with varying number of ports and vessels respectively. The times are average values based on 5 repeated runs.

set of generic test instances have been provided and computational results indicate that the model is capable of handling larger disruption scenarios than the real-life cases in seconds. However, with an increasing number of vessels, the computational time show exponential growth and can no longer reach an optimal solution within ten seconds, for larger instances. An analysis of the four real life cases, show that a disruption allowing to omit a port call or swap port calls may ensure timely delivery of cargo without having to increase speed and hence, a decision support tool based on the VSRP may aid in decreasing the number of delays in a liner shipping network, while maintaining a *slow steaming* policy. This initial work on disruption management in liner shipping show potential for interesting extensions. Other recovery modes than the three considered (speed adjustment, port call omission and port call swap) could be investigated, e.g. reducing the time spent at port by unloading but not loading, merging port calls or adding protection arcs. Another extension would be to reroute the non-satisfied cargo on the remaining, or even third party network. The connection with berth scheduling problems with disruption of fixed scheduled services as considered here could also be explored further.

Acknowledgements

The authors gratefully acknowledge the support from The Danish Council for Strategic Research to the ENERPLAN project. This work would not have been possible without the adept guidance and support from experienced Maersk Line staff, in particular we would like to thank Mikkel M. Sigurd, Daniel Bruun and Steffen Conradsen, who with their knowledge of disruption management and liner shipping, have given invaluable comments for this project.

Bibliography

- [1] Patrick M Alderton. *Sea Transport: Operation and Economics*. Reeds, 5 edition, September 2004.
- [2] José Fernando Álvarez. Joint routing and deployment of a fleet of container vessels. *Maritime Economics & Logistics*, 11(2):186–208, June 2009.
- [3] Michael Ball, Cynthia Barnhart, Martin Dresner, Mark Hansen, Kevin Neels, Amedeo Odoni, Everett Peterson, Lance Sherry, Antonio Trani, and Bo Zou. Total delay impact study - a comprehensive assessment of the costs and impacts of flight delay in the united states. Technical report, National Center of Excellence for Aviation Operations Research (NEXTOR), October 2010.
- [4] Cynthia Barnhart. *The Global Airline Industry*, chapter 7: Airline Schedule Optimization. Peter Belobaba, Amedeo Odoni, and Cynthia Barnhart for John Wiley and Sons, 2009.
- [5] Stephane Bratu and Cynthia Barnhart. Flight operations recovery: New approaches considering passenger recovery. *Journal of Scheduling*, 9:279–298, June 2006.
- [6] Berit D. Brouer, Jakob Dirksen, David Pisinger, Christian E.M. Plum, and Bo Vaaben. The vessel schedule recovery problem (vsrp) - a mip model for handling disruptions in liner shipping. *European Journal of Operational Research*, 224(2):362 – 374, 2013.
- [7] Marielle Christiansen, Kjetil Fagerholt, and David Ronen. Ship routing and scheduling: Status and perspectives. *Transportation Science*, 38(1):1–18, 2004.
- [8] Marielle Christiansen, Kjetil Fagerholt, Bjørn Nygreen, and David Ronen. *Handbooks in Operations Research & Management Science: Transportation*, volume 14, chapter 4: Maritime Transport. Elsevier, 2007.
- [9] Jens Clausen, Allan Larsen, Jesper Larsen, and Natalia J. Rezanova. Disruption management in the airline industry - concepts, models and methods. Technical report, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, 2009.
- [10] Dominik Dienst, Stefan Røpke, and Bo Vaaben. Realistic models and computational results for disruption management in the airline industry. Technical report, Technical University of Denmark, 2012.
- [11] Jakob Dirksen. Disruption management in liner shipping. Master’s thesis, Technical University of Denmark, June 2011.
- [12] Yuquan Du, Qiushuang Chen, Xiongwen Quan, Lei Long, and Richard Y.K. Fung. Berth allocation considering fuel consumption and vessel emissions. *Transportation Research Part E: Logistics and Transportation Review*, 47(6):1021 – 1037, 2011. ISSN 1366-5545.
- [13] Matthias Ehrgott. *Multicriteria Optimization*. Springer, second edition, 2005.
- [14] Antonio Frangioni and Bernard Gendron. 0 - 1 reformulations of the multicommodity capacitated network design problem. *Discrete Applied Mathematics*, 157(6):1229–1241, 3/28 2009.
- [15] Ahmad I. Z. Jarrah, Gang Yu, Nirup Krishnamurthy, and Ananda Rakshit. A decision support framework for airline flight cancellations and delays. *Transportation Science*, 27(3):266–280, August 1993.
- [16] JEC. Your flight has been delayed again - flight delays cost passenger, airlines, and the u.s. economy billions. Technical report, Joint Economic Committee (JEC), Maj 2008. Chairman, Senator Charles E. Schumer. Vice Chair, Rep. Carolyn B. Maloney.

- [17] Karina H. Kjeldsen. *Routing and Scheduling in Liner Shipping*. PhD thesis, Aarhus University, April 2012.
- [18] Niklas Kohl, Allan Larsen, Jesper Larsen, Alex Ross, and Sergey Tiourine. Airline disruption management—perspectives, experiences and outlook. *Journal of Air Transport Management*, 13(3):149–162, 2007.
- [19] Qiang Li, Shiqi Tong, Chunxia Yang, and Nuo Wang. Optimization of operation scheme of container terminal based on disruption management. *International Conference on Transportation Engineering 2009 (Conference Proceeding Paper)*, pages 3477–3482, 2009.
- [20] Berit Løfstedt, J. Fernando Alvarez, Christian Edinger Munk Plum, David Pisinger, and Mikkel Mühldorf Sigurd. An integer programming model and benchmark suite for liner shipping network design. Technical report, DTU Management Engineering, Technical University of Denmark and Maersk Line, Copenhagen, 2011.
- [21] Maersk. Slow steaming here to stay. News on Maersk Homepage, September 1, 2010. Link: <http://www.maersk.com/AboutMaersk/News/Pages/20100901-145240.aspx> on June 1, 2011.
- [22] Lavanya Marla, Bo Vaaben, and Cynthia Barnhart. Integrated disruption management and flight planning to trade off delays and fuel burn. Technical report, Technical University of Denmark, 2011.
- [23] Theo E. Notteboom. The time factor in liner shipping services. *Maritime Economics and Logistics*, 8(1):19–39, 2006.
- [24] Theo E. Notteboom and Bert Vernimmen. The effect of high fuel costs on liner service configuration in container shipping. *Journal of Transport Geography*, 17:325–337, 2009.
- [25] Ananda Rakshit, Nirup Krishnamurthy, and Gang Yu. System operations adviser: A real-time decision support system for managing airline operations at united airlines. *Interfaces*, 26:50–58, March-April 1996.
- [26] Elisabeth Rosenthal. Slow trip across sea aids profit and environment. The New York Times, February 16, 2010. Link: <http://www.nytimes.com/2010/02/17/business/energy-environment/17speed.html> on March 17, 2011.
- [27] Martin Stopford. *Maritime Economics*. Routledge, 3 edition, February 2009.
- [28] Dusan Teodorović and Slobodan Guberinić. Optimal dispatching strategy on an airline network after a schedule perturbation. *European Journal of Operational Research*, 15(2):178–182, February 1984.
- [29] Benjamin G. Thengvall, Jonathan F. Bard, and Gang Yu. Balancing user preferences for aircraft schedule recovery during irregular operations. *IEE Transactions*, 32:181–192, 2000.
- [30] Benjamin G. Thengvall, Gang Yu, and Jonathan F. Bard. Multiple fleet aircraft schedule recovery following hub closures. *Transportation Research Part A: Policy and Practice*, 35(4): 289–308, 2001.
- [31] Benjamin G. Thengvall, Jonathan F. Bard, and Gang Yu. A bundle algorithm approach for the aircraft schedule recovery problem during hub closures. *Transportation Science*, 37: 392–407, November 2003.
- [32] UNCTAD. Review of maritime transport 2010. http://www.unctad.org/en/docs/rmt2010_en.pdf, 2010. United Nations Conference on Trade And Development.

-
- [33] Bo Vaaben and Antonio Alves. Disruption management - passenger re-accommodation strategies. Presented at AGIFORS Airline Operations, 2009. URL <http://www.agifors.org/studygrp/opsct1/2009>.
 - [34] Shangyao Yan and Dah-Hwei Yang. A decision support framework for handling schedule perturbation. *Transportation Research Part B: Methodological*, 30(6):405–419, 1996.
 - [35] Chunxia Yang, Nuo Wang, and Bing Zhao. A berth scheduling problem under sailing schedule disruption. *ICCTP 2010: Integrated Transportation Systems (Conference Proceeding Paper)*, pages 1529–1536, 2010.
 - [36] Gang Yu and Xiangtong Qi. *Disruption Management: Framework, Models and Applications*. World Scientific Publishing Company, October 2004.
 - [37] Gang Yu, Michael Argüello, Gao Song, Sandra M. McCowan, and Anna White. A new era for crew recovery at continental airlines. *Interfaces*, 33:5–22, January 2003.

Chapter 7

Bunker Purchasing with Contracts

Christian E. M. Plum^{*, ‡} Peter N. Jensen[†] David Pisinger^{*}

^{*}Department of Management Engineering, Technical University of Denmark,
Produktionstorvet, Building 426, DK-2800 Kgs. Lyngby, Denmark
cemp@man.dtu.dk, pisinger@man.dtu.dk

[‡]Maersk Line
Esplanaden 50, DK-1098 København K.
mikkel.muhldorff.sigurd@maersk.com

[†]Maersk Oil Trading
Esplanaden 50, 1098 København K., Denmark
peter.neergaard.jensen@maersk.com

1

Abstract The cost for bunker fuel represents a major part of the daily running costs of liner shipping vessels. The vessels, sailing on a fixed roundtrip of ports, can lift bunker at these ports, having differing and fluctuating prices. The stock of bunker on a vessel is subject to a number of operational constraints as capacity limits, reserve requirements and sulphur content. Contracts are often used for bunker purchasing, ensuring supply and often giving a discounted price. A contract can supply any vessel in a period and port, and is thus a shared resource between vessels, which must be distributed optimally to reduce overall costs. The Bunker Purchasing with Contracts Problem (BPCP) has been formulated as a mixed integer program, which has been Dantzig-Wolfe decomposed. To solve it, a novel column generation algorithm has been developed. The algorithm has been run on a series of real world instances with up to 500+ vessels and 500+ contracts, and provide near optimal solutions. A MIP model cannot solve these instances due to memory requirements.

Keywords Bunker purchasing, Liner shipping, Mathematical programming, Maritime optimization, Decomposition methods.

7.1 Introduction

Liner shipping companies are at the core of the worlds supply chains providing cheap, compared to any other transport mode, and reliable transport from any corner of the world. This industry has grown massively in the last decades, often with two digit percentage growth rates. But lately the supply of vessels have exceeded the demand for container transport, resulting in many liner

¹Submitted for *Maritime Economics and Logistics*

carriers being loss giving. The profit margins in liner shipping are very slim, with marginal changes resulting in a company loss instead of profit.

This has shifted the shipping industry from a revenue optimizing focus, to use more resources on controlling and minimizing their costs. An example is the spend on bunker fuel, as this constitutes a very large part of the variable operating cost for the vessels.

For liner shipping companies in particular, the purchasing of bunkers can be planned some months ahead, as the vessels are sailing on a fixed schedule allowing for planning, as opposed to other types of shipping. Bunker prices are fluctuating and generally correlated with the crude oil price. But there are significant price difference between ports of up to 100 \$/mt (of a 600 \$/mt price). The price differences between ports are not stable, and the cheapest port on a roundtrip today, may not be the cheapest tomorrow. This creates the need for a frequent (daily) reoptimization of the bunker plan for a vessel, to ensure the lowest bunker costs. An example of a bunkering plan can be seen in Table 7.1.

Bunker Contracts The market for bunker trading is commoditized and liquid, the use of contracts for a specified amount, port and price (or discount to some price-index) is widespread. This is done to reduce both delivery and price risk and to leverage the strength of being a large player on this market.

Liner shipping companies engage in contracts for the purchase of bunkers at ports where they have a large and regular demand. This is done both to gain a discount compared to the spot market, by leveraging on the large volumes involved, and to increase supply certainty. An example of a liner shipping service can be seen in Figure 7.1. Bunker contracts will usually concern total lifted volumes within a calendar month, with specified minimum and maximal quantities.

The price can be agreed on in different manners, usually by using a fixed discount below the monthly average of a bunker index (Bunkerwire [5]) of the port in question. A contract is for one or more bunker grades and one or more ports, which will be close geographically and considered as the same market. Many contracts can be available in a port for a bunkering vessel, and it must then be chosen which, if any, to purchase bunker from. *Spot* bunker is assumed freely available at all ports with published prices.

Bunker is available in many different variations, grouped in different grades defined by ISO specifications. Two main parameters for a bunker's quality is its *viscosity* and its *sulphur* content. Lower grade (more viscous) bunker is considered better as it places less restrictions on the engine burning the bunker, and lower sulphur content is considered better, because of its lighter environmental footprint. In practice you will always buy the highest viscosity bunker, burnable by your ships engine and available in the port, hence the bunkers viscosity is not considered in this work.

The sulphur content must be considered as an increasing number of regions of the world have SO_x Emission Control Areas *SECA*, for details refer to DNV [10]. To model this, two bunker types are considered High Sulphur Fuel Oil, *HSFO* or *H* and Low Sulphur Fuel Oil, *LSFO* or *L*. The price difference between these varies, but often HSFO is 30 \$/mt cheaper than LSFO.

The model considered in this paper uses a *crystal ball* approach, i.e. using data not known at decision time, to benchmark the quality of already executed decisions. As the actual price of the contract is not known before a month has passed, the model will use after-the-fact prices for calculations.

The problem is to satisfy the vessels consumption by purchasing bunkers at the minimum overall cost, while considering reserve requirements, other operational constraints and adhering to a number of bunker contracts, the *Bunker Purchasing with Contracts Problem*. A novel decomposition method is presented for BPCP, and the first results, solving BPCP to near optimality are presented for large real world instances.

7.1.1 Literature

In this section we relate to literature relevant for this study within maritime optimization, liner shipping and bunker usage and purchasing in particular. For a broad introduction to shipping and the importance of bunker spend refer to Stopford [20] and for an introduction to operations

research within the maritime industry Christiansen et al. [7] and Christiansen et al. [8] provides excellent overviews. A detailed description of Liner Shipping Network Design and, the impact of bunker usage and other relevant factors appears in Brouer et al. [4]. They also introduce **LINER-LIB 2012**, a benchmark data suite, consisting of liner shipping relevant data and benchmarks specifically for liner shipping network design problems. Details on the bunkering industry in relation to shipping can be found in Boutsikas [3].

For a vessel sailing on a given port to port voyage at a given speed, the bunker consumption can be fairly accurately predicted. This gives an advantage in bunker purchasing, when a vessel has a stable schedule known for some months ahead. The regularity in the vessel schedules in liner shipping allows for detailed planning of a specific vessel, as considered in the works of Plum and Jensen [18], Besbes and Savin [2] and Yao et al. [23]. These papers consider variants of a bunker optimization problem considering a single vessel. The work of Plum and Jensen [18] considers multiple tanks in the vessel and stochasticity of both prices and consumption. Yao et al. [23] does not consider stochastic elements nor tanks, but has vessel speed as a variable of the model. Bunker contracts are not considered in these.

Besbes and Savin [2] consider different refueling policies for liner vessels and has some good considerations on the modeling of stochastic bunker prices using markov processes. This is used to show that the bunkering problem in liner shipping can be seen as a stochastic capacitated inventory management problem. Bunker contracts are not considered, neither are other operational constraints than capacity.

The work of Farina [12] is an extension of Plum and Jensen [18] with the additional consideration of bunker contracts, where a MIP model is presented capable of solving a 50 vessel instance for a 6 month period, falling short of solving real world instances of hundreds of vessels. An outline of a similar model and decomposition of this was presented at the ICCL 2012 conference, Farina et al. [13], but without computational results. The effect of the bunker price on Liner Shipping Network Design has been studied in a number of recent papers as Wang and Meng [21] and Meng et al. [16].

The effect of bunker usage by the maritime industry in relation to the bunker price is investigated by Corbett et al. [9] with the aim of reducing CO_2 emissions by imposing tax on bunkers. The work of Acosta et al. [1] considers factors impacting the choice of bunker port. Fagerholt et al. [11] considers the optimal speed and route for a ship with respect to bunker costs. Other work on bunker costs and its impact on maritime transportation includes Notteboom and Vernimmen [17], who consider how slow steaming and the cost structure of liner shipping networks are affected by changes in bunker costs, and Ronen [19], who considers the bunker price's effect on speed and fleet size. The recent work of Wang et al. [22] provides an overview of available bunker optimization methods in shipping.

Contribution The contribution of this paper is a model considering contracts and other operational constraints, as reserve requirements and minimal lift quantities, relevant when purchasing bunker for a liner shipping company. This model has been Dantzig-Wolfe decomposed and a novel column generation algorithm created. It is discussed, how the decomposition allows for the dual values of the contracts to be investigated. Finally, an implementation of the algorithm shows it can solve very large problems, met in real world problem instances. As stated in the *Summary and Future Work* of Besbes and Savin [2]:

The single-vessel approach developed in this paper can serve as a good initial step to optimizing fleet profits. At the same time, real business settings are often characterized by volume refueling discounts, which can only be fully exploited using more than one vessel. Thus, the development of multiple-vessel profit management models represent a challenging research direction that will be of immediate interest to practitioners.

The structure of the paper is as follows. Section 7.2 explains the mathematical notation used throughout the paper. A Mixed Integer Programming model for BPCP is formulated in Section

7.2.1. The model is Dantzig-Wolfe (DW) decomposed and *Column Generation* algorithm is devised in Section 7.3. Computational results and comparison of the MIP and the DW algorithm is presented in Section 7.4, where it can be seen that DW can solve many instances, which can not be solved by the MIP model. Conclusion and outlook is given in section 7.5.

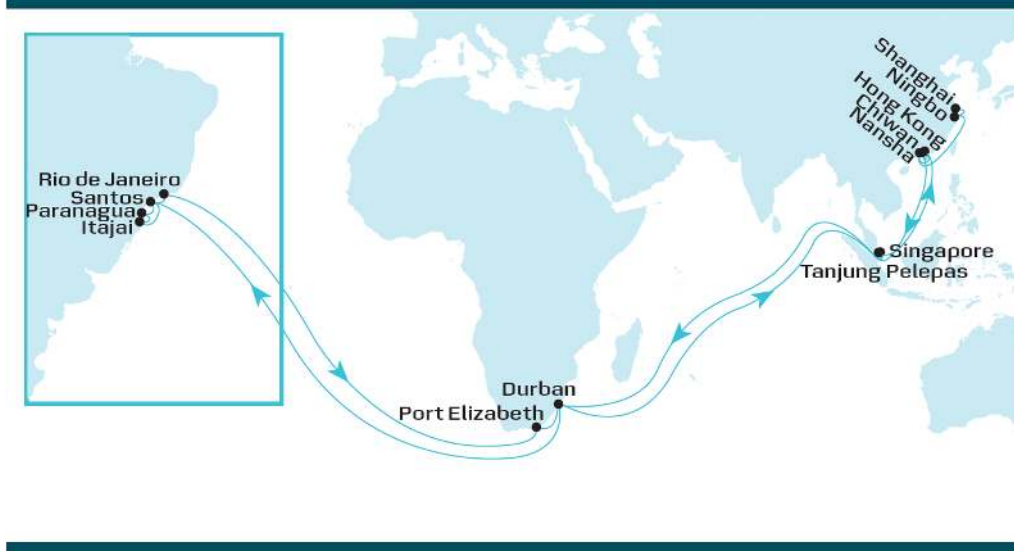


Figure 7.1: The ASAS 2 Service, transporting containers between East Coast South America, South Africa and the Far East. This service allows for bunkering in three distinct markets.

7.2 Bunker Purchasing with Contracts

We introduce the mathematical notation used throughout the paper. Let $v \in V$ be the set of vessels. Let $i \in I$ be an ordered set of *port calls*, the vessel's schedule. A port call i will be uniquely defined by a port, a vessel, $v(i)$ and a date. Let $init(v)$ and $term(v)$ be the first and last considered port call of vessel v . Let $b \in B = \{L, H\}$ be the two considered bunker types. The startup cost for bunkering at a port call i , is $startcost_i$. Each vessel, v has a capacity $D_{v,b}$ for each bunker type, b . For each leg i of the schedule, the vessel consumes $F_{i,b}$ bunker, between port call i and $i + 1$.

Contract bunker must be purchased according to details given by a number of contracts $c \in C$, minimal and maximal quantities are given by \underline{q}_c and \bar{q}_c . The specified quantities are soft constraints, which can be violated by paying a high cost, \underline{w} , for violating the minimum volume and a lower cost for breaking the maximal constraint, \bar{w} . Contract c may cover several ports and multiple vessels can call at these ports in the duration of the contract. Each contract will give rise to a number of purchase options, $m \in M$, i.e. discrete events where a specific port call i , and thus vessel v , calls within a time period, allowing it to purchase bunker from a contract c . Purchases on a purchase option m will be done at a price p_m , specified by the contract c . To simplify modelling and to increase the density of the derived model, the sets of port calls, $i \in I$ and purchase options, $m \in M$ will be used instead of their underlying sets: ports, vessels and contracts, which could give an equivalent but much larger model.

The possibility of purchasing on the spot market, is considered as a special type of contract. The minimal and maximal volumes are relaxed as $\underline{q}_c = 0$ and $\bar{q}_c = \infty$. All port calls i have two spot purchase options m for LSFO and HSFO, with prices set at the corresponding spot price of the day and port. For ports where bunker prices are not published, we assume a high cost.

The variables of the model are: l_m the purchase of bunker for each purchase option m . The binary variable $\delta_{i,b}$ is set iff a purchase of a bunker type b is made at a port call i . The volume

of bunker after a vessel leaves port, $h_{i,b}$ is a continuous variable, as is the consumption of each bunker type on vessel between port i and $i + 1$, $f_{i,b}$. The contract violation or slack variables are \underline{s}_c and \bar{s}_c .

7.2.1 Model

The BPCP can be formulated as a Mixed Integer Program:

$$\min \sum_{i \in I} \sum_{b \in B} (\delta_{i,b} \cdot \text{startcost}_i) + \sum_{m \in M} (p_m \cdot l_m) + \sum_{c \in C} (\underline{s}_c \cdot \underline{w} + \bar{s}_c \cdot \bar{w})$$

Subject to

$$h_{i,b} = h_{i-1,b} + \sum_{m \in M(i,b)} l_m - f_{i-1,b} \quad \forall i, b \quad (7.1)$$

$$f_{i,b} \leq h_{i,b} \quad \forall i, b \quad (7.2)$$

$$\sum_{b \in B} f_{i,b} = F_{i,H} + F_{i,L} \quad \forall i \quad (7.3)$$

$$f_{i,L} \geq F_{i,L} \quad \forall i \quad (7.4)$$

$$h_{i,b} \leq D_{v(i),b} \quad \forall i, b \quad (7.5)$$

$$\underline{q}_c - \underline{s}_c \leq \sum_{m \in M(c)} l_m \leq \bar{q}_c + \bar{s}_c \quad \forall c \quad (7.6)$$

$$\sum_{m \in M(i,b)} l_m \leq \delta_{i,b} \cdot D_{v,b} \quad \forall i, b \quad (7.7)$$

The objective minimizes startup costs, bunker cost and contract violation penalties. The constraints (7.1) ensures flow conservation at each port, vessel and bunker type. Constraints (7.2) ensures that no more bunker than available is used between port i and $i + 1$. Constraints (7.3) and (7.4) maintains the consumption of bunker, allowing LSFO to substitute HSFO, but not opposite. The bunker capacity of the vessels are enforced by constraints (7.5). The minimal and maximal quantity required by the contracts are ensured by double sided constraints (7.6), allowing for violations. The decision variables $\delta_{i,b}$ are set by constraints (7.7).

To facilitate use in a benchmark setting, initialization and termination criteria for start and end bunker volumes must also be set:

$$h_{i(0,v),b} = h_b^{\text{init}(v)} \quad \forall v, b \quad (7.8)$$

$$\sum_{b \in B} h_{i(t,v),b} \geq \sum_{b \in B} h_b^{\text{term}(v)} \quad \forall v \quad (7.9)$$

$$h_{i(t,v),L} \geq h_L^{\text{term}(v)} \quad \forall v \quad (7.10)$$

Variable domains:

$$h_{i,b}, l_m, f_{i,b}, \underline{s}_c, \bar{s}_c \in \mathbb{R}^+ \quad \forall i, b, m, c \quad (7.11)$$

$$\delta_{i,b} \in \{0, 1\} \quad \forall i, b \quad (7.12)$$

7.2.2 Operational Constraints

In practice bunker purchasing in liner shipping is influenced by a wide range of operational, commercial and financial factors, which dictates the properties of a good bunker plan. Some of these are described here and a few are formulated as constraints, refer to earlier mentioned literature for an elaborate discussion of other factors.

As the consumption of bunker on a leg is an uncertain parameter due to factors as changed schedule (and thus speed), wind, current, waves and hull roughness, a good bunker plan will allow for variation in the bunker consumption. A way to handle this is to enforce a minimum reserve requirement of bunker at port arrival. This can be modeled as in (7.13), where F_i is the minimal reserve requirement at port arrival.

Besides the startup cost for bunkering, $startcost_i$, bunker suppliers will usually require a minimum quantity to be purchased at each bunkering, this can be handled with constraints (7.14), where $\underline{L}_{i,b}$ is the minimal quantity.

$$F_i \leq \sum_{b \in B} (h_{i,b} - \sum_{m \in M} l_m) \quad \forall i \quad (7.13)$$

$$\delta_{i,b} \cdot \underline{L}_{i,b} \leq \sum_{m \in M} l_m \quad \forall i, b \quad (7.14)$$

Capital and carriage cost The capital costs of bunker is extensive, due to the large volumes and high prices. A model could consider this by adding this cost (or lacking interest) to the objective, proportional to the average load of bunker on the vessels. Similarly a vessel carrying a large volume of bunker will, all things equal, have a larger draft. This will in general (but not always, due to specifics in the vessels design as the bulb) imply an increased bunker consumption proportional to an increased load. With realistic values of this relation, this term could be considered in the objective in the same manner as the capital costs.

California sales tax The California bunker sales tax, as described by California Legislative Analyst's Office [6], imposes a tax on bunker bought in California, which necessarily must be used enjourney to the first out of state port. I.e. if a vessel arrives with 1000 mt at an Californian port and requires 2000 mt to reach the first non-californian port on its schedule, it must pay a tax for the first 1000 mt purchased. With additional decision variables this can be modelled and included in the objective.

Contract min/max volumes & port call max volumes Contracts may have minimum and maximal volumes that must be lifted per purchase. This can be modelled similarly to the minimum lift constraints. As can spot purchases at port calls have maximal lift restrictions due to short port stays or limited supply.

Quarantine A sample is usually taken from purchased bunker, to be analyzed for its specific content of carbohydrates, sulphur, water, ashes, etc. the ample must be within the ISO specifications of the purchased bunker grade. Until the result of the laboratory test are received, the bunker may not be used. This test can take 3-5 days. This constraint can be considered by increasing the reserve requirements at port calls with bunker purchased within the last 5 days.

Constraints for capital and carriage cost, California sales tax, contract min/max volumes, port call max volumes and quarantine have not been implemented, to limit the required implementation effort. All of them can be formulated linearly and only relate to a single vessel at a time, allowing them to be considered in a vessel specific subproblem.

7.2.3 Complexity

The problem is at least as hard as weakly NP-hard problems, seen by reduction from the knapsack problem, as described in Kellerer et al. [14]. From the knapsack problem in minimization form: Given a set N of items having profit p_i and weight w_i and a knapsack of capacity c , the problem is to fill the knapsack at minimum overall profit, such that the overall weight is at least c . Given an instance of the knapsack problem, we construct an instance of the bunker purchasing problem by having one vessel, visiting N ports. The fuel consumption between each pair of ports is 0, except

the leg after the last port visit, where the consumption is c . In each port, we have a contract of maximum w_i , and the minimum limit for lifting bunker is also w_i . The cost of buying the quantity w_i is p_i . It is easily seen that solving the bunker purchasing problem also solves the knapsack problem.

7.3 Algorithm

The fleet of a global liner shipping company may consist of hundreds of vessels, with many of these having overlapping schedules visiting the same hub ports. This means that the full problem can be of a very large size, making the MIP model impossible to solve for large instances as seen in Section (7.4). This makes it interesting to consider a decomposition of the MIP model, to solve these large problem instances.

The arc flow model given by (7.1) - (7.14) is Dantzig-Wolfe decomposed on the variables l_m . Let R_v be the set of all feasible bunkering patterns for a vessel v , satisfying constraints (7.1) - (7.14), except (7.6). This set has an exponential number of elements. Each pattern $r \in R_v$ is denoted as a set of bunkerings. Let $u_r = \sum_{m \in M} (p_m \cdot l_m) + \sum_{i \in I} \sum_{v \in V} \sum_{b \in B} (\delta_{i,b} \cdot \text{startcost}_i)$ be the cost for pattern $r \in R_v$. Let λ_r be a binary variable, set to 1 iff the bunkering pattern r is used. Let $o_{r,c}$ be the quantity purchased of contract c by pattern r . The BPCP can then be formulated as:

$$\min \sum_{v \in V} \sum_{r \in R_v} \lambda_r \cdot u_r + \sum_{c \in C} (\underline{s}_c \cdot \underline{w} + \bar{s}_c \cdot \bar{w}) \quad (7.15)$$

Subject to

$$\underline{q}_c - \underline{s}_c \leq \sum_{v \in V} \sum_{r \in R_v} \lambda_r \cdot o_{r,c} \leq \bar{q}_c + \bar{s}_c \quad \forall c \quad (7.16)$$

$$\sum_{r \in R_v} \lambda_r = 1 \quad \forall v \quad (7.17)$$

$$\lambda_r \in \{0, 1\} \quad \forall r \quad (7.18)$$

The objective minimizes the costs of purchased bunker, startup costs and slack costs. Constraints (7.16) ensures that all contracts are fulfilled. Convexity constraints (7.17) ensure that exactly one bunker pattern is chosen for each vessel.

7.3.1 Pricing Problem

Let $\underline{\pi}_c \leq 0$ and $\bar{\pi}_c \leq 0$ be the dual variables for the upper and lower contract constraints (7.16), due to the structure of these constraints at least one of these will be 0 for each contract c . Let $\theta_v \in \mathbb{R}$ be dual variables for the convexity constraints (7.17). Then the pricing problem becomes:

$$\text{Min: } u_r + \sum_{c \in C} (\underline{\pi}_c - \bar{\pi}_c) - \theta_v \quad (7.19)$$

Subject to constraints (7.1) - (7.14), except (7.6).

This pricing problem is a Mixed Integer Program, considering a single vessel. This size of problem can be solved in reasonable time by a standard MIP solver, as done in Plum and Jensen [18]. Columns λ_r with negative reduced cost will then be added to the master problem, also solved as a MIP.

7.3.2 Column Generation Algorithm

Due to the large number of columns the problem is solved by a *Column Generation* algorithm, where the root node is solved to LP optimality. The root node is then solved with integral property on all columns by a MIP solver.

Initially all dual variables are set to zero, a subproblem is constructed for each vessel and solved as a MIP problem. The first master problem is then constructed with one solution for each vessel as columns. This master is solved and the first dual values are found. The subproblems are resolved for all vessels (only the objective coefficients for the contracts needs updating) and new columns are generated for the master. This continues until no negative reduced cost columns can be generated, and the LP optimal solution is achieved.

Following, the problem is solved as a MIP, providing an integral solution. The subproblems only need to find a negative reduced costs column, to ensure progress of the algorithm. This means that initially they are allowed to return solutions with considerable subproblem gaps. As the algorithm progresses, the allowable subproblem gap is reduced, until it reaches the tolerance level.

7.3.3 Dual stabilization

A simple form of dual stabilization has been used in the implementation to speed up convergence. The Boxstep method described in Marsten et al. [15] imposes a box around the dual variables, which are limited from changing more than π_{max} per iteration. This has been motivated by the dual variables only taking on values $\{-\underline{w}, \bar{w}, 0\}$ in the first iteration, these then stabilize at smaller numerical values in subsequent iterations.

7.3.4 Interpretation of dual values

Besides using the developed method for benchmarking the historical performance of bunker purchases, it can be used in a context of evaluating the gain of a considered contract.

Using best estimates for bunker consumption and prices (current prices for instance) together with known or expected contracts a baseline bunker purchasing plan could be run. A new scenario could then be constructed with the addition of the considered contract and by analyzing the output, it could be seen whether the overall costs of the scenario increased or decreased as compared with the baseline.

Another investigation could be to solely consider the baseline's final dual variables, π_c and $\bar{\pi}_c$, and depending on the magnitude of these evaluate the contracts effect. As these dual values are the same for all subproblems, they can be interpreted as balancing out the price of the contract, increasing the price if it is a popular contract or decreasing it otherwise, converging when they are in balance. The magnitude of this will be proportional to the contracts gain.

7.4 Computational Results

The proposed algorithm has been implemented in **ILOG OPL** as modelling language and **Cplex 12.2** as LP/MIP solver, this implementation is referred to as **DW**. To evaluate the performance of **DW** the MIP model of Section 7.2.1 has been implemented in **Cplex 12.2**.

Real life data for a large number of liner vessels describing their schedules, consumptions, tank capacities and other relevant data has been made available by **Maersk Oil Trading**, who have also supplied data on a large number of actual bunker contracts and spot prices available in a range of ports. Based on these data a number of instances have been constructed to test the scalability and performance of the implementations. Due to confidentiality reasons the price's have been distorted by $\pm 10\%$, in order to maintain the structure of the problem. The penalty \underline{w} for violating minimum volume is set at 200 \$/mt, and the penalty \bar{w} for breaking the maximal constraint at 50 \$/mt. If a bunker price is not available at a port, the price is set at 1000 \$/mt. Details about the instances can be seen in Table 7.2.

7.4.1 Parameter Tuning

Three parameters have been tested to improve the running time and performance of the DW algorithm, all test have been performed on 4 medium and large problem instances and the average

changes in running time and objective value, as compared with the best are shown.

Table 7.5 shows the effect of different values of the maximal Box Step size of the Boxstep method. A value of 1000 indicates that the duals are free, as these are bounded by $\{-\underline{w}, \bar{w}\}$. It can be seen, that the algorithm performs better in terms of time and quality using the boxstep method. A value of 75 is chosen for further runs, as it gives the best tradeoff of running time and objective value.

As described in section 7.3.2 the subproblems are initially not solved to optimality. Table 7.6 investigates the initial gap the subproblems terminates with. The improvement effect is small, but a benefit arises with an initial gap of 0.01. Table 7.7 investigates by which factor this gap should decrease until reaching $10e-6$. A value of 25 is chosen.

An overview of the performance and results can be found in Tables 7.3 and 7.4. It can be seen that the **DW** model is able to solve the problem for all instances. For larger instances **MIP** runs out of memory and finds no solution, due to the size of the instances and their resulting MIPs. Both models find solutions with very small gaps, but still considerable absolute gap's to the optimal solution. **MIP** only finds optimal solutions for the smallest instances, for all medium and large instances the solver runs out of memory before it has closed the gap. **DW** is able to find solutions with relatively small gaps for even the largest problem instances covering all vessels and all contracts on a global level. In practice the resulting gaps of the algorithms, can be much less in reality as they are based on a lower bound.

7.5 Conclusion and Further Work

We have presented a MIP model for the Bunker Purchasing with Contracts Problem. This model has been Dantzig-Wolfe decomposed and a novel *Column Generation* algorithm was presented. The MIP model and the DW algorithm have been implemented and run on very large instances. The DW algorithm is able to find good solutions for all instances within the timelimit, where the MIP model is unable to find solutions before running out of memory. The advantage of the DW algorithm is that many of the constraints can be dealt with in the pricing problem. Additionally the dual information is provided which can be used to evaluate a contract. It has been certified by Maersk Oil Trading that the produced bunker plans are operationally feasible and that contracts are considered by the model in a adequate manner. Maersk Oil Trading will investigate the method in depth, with the aim of implementing its results in their operations, to better utilize their bunker contracts. The algorithm could also be used to evaluate the potential of a new contract, or what the right upper and lower volumes of a contracts should be, in a negotiating setting.

Numerous additional operational constraints can be included in the modeling, as reserve requirements, multiple bunker tanks, mixing penalties and others mentioned in related literature for single vessel bunkering problems. These could be included in the subproblem as they all deal with a single vessel. Further work could be done on the DW to close the optimality gaps, by branching on fractional variables of the root node, devising a *Branch-and-Price* algorithm.

Acknowledgements

The authors gratefully acknowledge the support from The Danish Council for Strategic Research to the ENERPLAN project. This work would not have been possible without the adept guidance and support from experienced Maersk Oil Trading and Maersk Line staff, in particular we would like to thank Claus Anker and Mikkel M. Sigurd who have given invaluable support for this project. We also thank Stephen Hall for good support with OPL/CPLEX and Federico Farina for considerable assistance in preparing the data for the investigated instance.

Bibliography

- [1] Manuel Acosta, Daniel Coronado, and Ma Del Mar Cerban. Bunkering competition and competitiveness at the ports of the gibraltar strait. *Journal of Transport Geography*, 19(4): 911 – 916, 2011.
- [2] Omar Besbes and Sergei Savin. Going bunkers: The joint route selection and refueling problem. *Manufacturing and Service Operations Management*, 11(4):694–711, 2009.
- [3] Angelos Boutsikas. The bunkering industry and its effect on shipping tanker operations. Master’s thesis, Massachusetts Institute of Technology. Dept. of Ocean Engineering., 2004.
- [4] Berit D. Brouer, J. Fernando Alvarez, Christian E. M. Plum, David Pisinger, and Mikkel M. Sigurd. A base integer programming model and benchmark suite for liner shipping network design. *Transportation Science*, 2013. Accepted by Transportation Science.
- [5] Platts Bunkerwire. Platts bunkerwire. <http://www.platts.com/IM.Platts.Content/ProductsServices/Products/b> 2013.
- [6] California Legislative Analyst’s Office. California bunker sales tax. http://www.lao.ca.gov/2001/bunker_fuel/012501_bunker_fuel.html.
- [7] Marielle Christiansen, Kjetil Fagerholt, and David Ronen. Ship routing and scheduling: status and perspectives. *Transportation Science*, 38(1):1–18, 2004.
- [8] Marielle Christiansen, David Ronen, Bo Nygreen, and Kjetil Fagerholt. Maritime transportation. In *Handbooks in Operations Research and Management Sciences*, pages 189–284. North Holland, 14 edition, 2007.
- [9] James J. Corbett, Haifeng Wang, and James J. Winebrake. The effectiveness and costs of speed reductions on emissions from international shipping. *Transportation Research Part D: Transport and Environment*, 14(8):593 – 598, 2009.
- [10] Det Norske Veritas DNV. Sulphur emission control area. http://www.dnv.com.ar/binaries/flyer_marpol_seca_tcm158-278353.pdf.
- [11] Kjetil Fagerholt, Gilbert Laporte, and Inge Norstad. Reducing fuel emissions by optimizing speed on shipping routes. *The Journal of the Operational Research Society*, 61(3):523–529, 2009.
- [12] Federico Farina. Bunker optimization with contracts. Master’s thesis, Universita Di Bologna, 2012.
- [13] Federico Farina, Peter N. Jensen, Christian E. M. Plum, and David Pisinger. Bunker purchasing with contracts. 2012. Presented at <http://iwi.econ.uni-hamburg.de/iccl2012/index.htm>.
- [14] Hans Kellerer, Ulrich Pferschy, and David Pisinger. *Knapsack Problems*. Springer, 2004.
- [15] Roy E. Marsten, William W. Hogan, and Jacob W. Blankenship. The boxstep method for large-scale optimization. *Operations Research*, 23(3):389–405, May/June 1975.
- [16] Qiang Meng, Shuaian Wang, and Zhiyuan Liu. Network design for shipping service of large-scale intermodal liners. *Transportation Research Record: Journal of the Transportation Research Board*, 2269:42–50, 2012.
- [17] Theo E. Notteboom and Bert Vernimmen. The effect of high fuel costs on liner service configuration in container shipping. *Journal of Transport Geography*, 17(5):325–337, 2009.
- [18] Christian E. M. Plum and Peter Neergaard Jensen. Minimization of bunker costs. Master’s thesis, 2007.

-
- [19] David Ronen. The effect of oil price on containership speed and fleet size. *J Oper Res Soc*, 62:211–216, 2010.
- [20] Martin Stopford. *Maritime Economics*. Routledge, third edition, 2009.
- [21] Shuaian Wang and Qiang Meng. Robust schedule design for liner shipping services. *Transportation Research Part E: Logistics and Transportation Review*, 48(6):1093 – 1106, 2012.
- [22] Shuaian Wang, Qiang Meng, and Zhiyuan Liu. Bunker consumption optimization methods in shipping: A critical review and extensions. *Transportation Research Part E: Logistics and Transportation Review*, 53(0):49 – 62, 2013.
- [23] Zhishuang Yao, Szu Hui Ng, and Loo Hay Lee. A study on bunker fuel management for the shipping liner services. *Computers and Operations Research*, 39(5):1160 – 1172, 2012.

Port Id	Departure stock (mt)		Consumption (mt)		Purchase (mt)		Spot Price (\$mt)	
	LSFO	HSFO	LSFO	HSFO	LSFO	HSFO	LSFO	HSFO
NLROT	462	648	5	0	0	0	481	455
DEBRV	457	648	100	0	0	0	500	490
GBFXS	357	648	97	648	0	0	1000	675
USNWK	260	1004	0	134	0	186+818	491	465
USCHS	260	870	0	425	0	0	490	477
USSAV	260	445	0	74	0	0	493	457
USMIA	260	371	0	211	0	0	1000	1000
USHOU	260	1456	0	122	0	1296	1000	442
USMOB	260	1334	183	201	0	0	1000	1000
USNFK	77	1133	24	555	0	0	484	469
GBFXS	53	578	8	0	0	0	1000	641
NLROT	1053	4314	0	2340	1008	3737	442	421
DEBRV	1053	1974	0	447	0	0	457	447
USNWK	1053	1527	2	110	0	0	490	466
USCHS	1051	1417	0	25	0	0	495	482
USSAV	1051	1392	0	82	0	0	502	471
USMIA	1051	1310	0	211	0	0	1000	1000
USHOU	1051	1099	0	128	0	0	1000	451
USMOB	1051	971	0	365	0	0	1000	1000
USNFK	1051	606	221	606	0	0	510	495
NLROT	830	4021	21	0	0	4021	436	415
GBFXS	809	4021	65	0	0	0	1000	652
DEBRV	744	4021	98	511	0	0	467	456
USNWK	646	3510	0	161	0	0	485	464
USCHS	646	3349	0	19	0	0	496	483
USSAV	646	3330	0	98	0	0	499	468
USMIA	646	3232	0	183	0	0	1000	1000
USHOU	646	3049	0	135	0	0	1000	465
USMOB	646	2914	0	388	0	0	1000	1000

Table 7.1: An example of a bunker plan. *Departurestock* is the stock of bunker at departure of the port, as calculated by the model. *Consumption* is, a model input, the consumption of bunker from this port to the next. *Purchase* is the quantity of bunker purchased at the port and *SpotPrice* is the market price of bunker at the spot market. Possible bunker contracts are not shown. At the fourth port call 186 mt HSFO is bought at the spot market and 818 mt HSFO through a contract.

Instance	Size	V	P	C
RULED	Small	6	1048	29
FRFSM	Small	8	2128	10
ZADUR	Small	49	5973	35
US_WC	Small	32	6022	68
USNWK	Medium	49	9048	69
USSAV	Medium	50	9194	23
PALBL	Medium	65	9817	27
AEJAL	Medium	80	15442	9
09_H2	Large	408	16214	307
11_H2	Large	572	18426	254
10_H1	Large	469	18704	332
10_H2	Large	534	21907	424
11_H1	Large	609	23453	376
HKHKG	Large	158	29177	20
10_FY	Large	535	40611	756

Table 7.2: Instances of varying sizes for the BPCP. **Instance** is the name, **Size** is a grouping of the instances. V the number of vessels, P the number of port calls, C the number of Contracts.

Instance	Obj_{MIP}	LB_{MIP}	Gap_{MIP}	t_{MIP}
RULED	5.404 e+7	5.404 e+7	0.00 %	1083
FRFSM	1.319 e+8	1.319 e+8	0.00 %	21
ZADUR	7.064 e+8	7.063 e+8	0.02 %	609
US_WC	6.628 e+8	6.626 e+8	0.03 %	481
USNWK	9.067 e+8	9.063 e+8	0.03%	834
USSAV	9,830 e+8	9.826 e+8	0.04 %	775
PALBL	1.108 e+9	1.107 e+9	0.06%	906
AEJAL	1.490 e+9	1.489 e+9	0.03%	686
09_H2	2.115 e+9	2.113 e+9	0.10%	1160
11_H2	2.478 e+9	2.475 e+9	0.09%	1107
10_H1	2.255 e+9	2.253 e+9	0.09%	1181
10_H2	Out of Mem			
11_H1	Out of Mem			
HKHKG	Out of Mem			
10_FY	Out of Mem			

Table 7.3: Results and performance of MIP implementation. **Instance** is the name. Obj_{MIP} is the best found solution for algorithm, LB_{MIP} is the best found lower bound. Gap_{MIP} is the resulting gap between upper and lower bound and t_{MIP} is the timed used in seconds.

Instance	Obj_{DW}	LB_{DW}	Gap_{DW}	t_{DW}
RULED	5.408 e+7	5.404 e+7	0.08 %	118
FRFSM	1.321 e+8	1.319 e+8	0.20 %	86
ZADUR	7.071 e+8	7.064 e+8	0.10 %	653
US_WC	6.654 e+8	6.627 e+8	0.41 %	1142
USNWK	9.077 e+8	9.066 e+8	0.11 %	1114
USSAV	9.830 e+8	9.829 e+8	0.00 %	399
PALBL	1.108 e+9	1.108 e+9	0.01 %	672
AEJAL	1.490 e+9	1.490 e+9	0.00 %	415
09_H2	2.120 e+9	2.115 e+9	0.22 %	8642
11_H2	2.479 e+9	2.477 e+9	0.07 %	9411
10_H1	2.259 e+9	2.255 e+9	0.19 %	7267
10_H2	2.529 e+9	2.526 e+9	0.12 %	10649
11_H1	3.217 e+9	3.214 e+9	0.09 %	10075
HKHKG	3.427 e+9	3.427 e+9	0.00 %	4344
10_FY	4.835 e+9	4.807 e+9	0.59 %	28922

Table 7.4: Results and performance of DW implementation. **Instance** is the name. Obj_{DW} is the best found solution for algorithm, LB_{DW} is the best found lower bound. Gap_{DW} is the resulting gap between upper and lower bound and t_{DW} is the timed used in seconds.

$DualBoxSize$	$\Delta Time$	$\Delta Objective$	$Iterations$
10	85.3%	0.026%	39
50	34.0%	0.061%	26
75	8.4%	0.052%	20
100	15.9%	0.056%	20
150	3.8%	0.095%	20
1000	8.2%	0.088%	20

Table 7.5: Parameter tuning for the Box step size used for Dual variable stabilization. $\Delta Time$ is the average percentage improvement in running time as compared with the fastest overall, taken over 5 medium and large problem instances. Likewise $\Delta Objective$ is the average percentage improvement in objective value. $Iterations$ gives the average number of iterations to solve the root node.

$InitSubGap$	$\Delta Time$	$\Delta Objective$	$Iterations$
0.00001	10.8%	0.078%	19
0.0001	11.9%	0.064%	19
0.001	9.3%	0.085%	19
0.01	10.3%	0.054%	19
0.1	15.4%	0.052%	19
0.5	30.2%	0.081%	21

Table 7.6: Parameter tuning for the initial gap the subproblems are allowed to terminate with. $\Delta Time$ is the average percentage improvement in running time as compared with the fastest overall, taken over 5 medium and large problem instances. Likewise $\Delta Objective$ is the average percentage improvement in objective value. $Iterations$ gives the average number of iterations to solve the root node.

<i>GapIncrement</i>	$\Delta Time$	$\Delta Objective$	<i>Iterations</i>
2	20.2%	0.044%	18
5	6.4%	0.088%	17
10	9.0%	0.052%	17
25	6.1%	0.041%	17
100	12.2%	0.044%	17

Table 7.7: Parameter tuning for the factor the sub problem gap is reduced with. $\Delta Time$ is the average percentage improvement in running time as compared with the fastest overall, taken over 5 medium and large problem instances. Likewise $\Delta Objective$ is the average percentage improvement in objective value. *Iterations* gives the average number of iterations to solve the root node.

Part IV

Conclusion

Chapter 8

Conclusion

Two overall research paths have been followed within liner shipping in this thesis: How to construct methods to assist with the design of liner shipping networks and how to model operational problems met in liner shipping. It has aimed at opening up research in the important area of liner shipping network design in a number of ways. By giving a thorough introduction to the domain, presenting a number of benchmark instances and proposing several models for liner shipping network design, which highlights important and not previously studied aspects of the problem. The mathematical models aid future research into LSNDP and the design of algorithms to solve significant instances of the problem. LINER-LIB 2012 data is already used by researchers from a number of countries around the world (Germany, Norway, Denmark, China, Singapore), for network design and other liner shipping projects. With the new research project, the *Competitive Liner Shipping Network Design* project ¹ an important foundation for research and industry collaboration is laid, which can be used by the world's liner shipping companies in the future, to optimize their networks. The project extends, among others, on the work of this thesis.

Operational liner shipping problems on bunker purchasing and disruption management, have been considered with good results, showing the breadth of research existing in liner shipping. From an industry point of view both operational models are applicable for implementation in actual decision support systems. These can help overcome some of the complex problems faced in liner shipping, showing that OR techniques can be applied to real liner shipping problems. This chapter highlights the main findings of the thesis, summarizing the conclusions of the works herein. It is followed by some thoughts on current trends in liner shipping network design research, and research in other liner shipping problems.

The paper in Chapter 2 highlights the potential for making cost effective and energy efficient liner shipping networks using OR. It is argued that a reason for lacking research has been, that access to domain knowledge and data is a barrier for researchers to approach the important liner shipping network design problem. The purpose of the benchmark suite and the paper is to provide easy access to the domain and the data sources of liner shipping for OR researchers in general. The liner shipping domain is analyzed and applied to network design and present a rich integer programming model based on services. It is proven that the liner shipping network design problem is strongly \mathcal{NP} -hard. A benchmark suite of data instances to reflect the business structure of a global liner shipping network is presented. The design of the benchmark suite is discussed in relation to industry standards, business rules and mathematical programming. Computational results yielding the first solutions for 6 of the 7 benchmark instances is provided using a heuristic combining tabu search and heuristic column generation.

¹A recently started research project at DTU Management, in cooperation with Maersk Line and funded by the Danish Maritime Fund.

The paper in Chapter 3 develops a method to guide the optimal deployment of vessels. A single vessel round trip is considered by minimizing operational costs and flowing the best paying demand under commercially driven constraints. The Single Service Design Problem have been introduced and arc-flow and path-flow models are presented. A Branch-and-Cut-and-Price algorithm is proposed and implemented. The algorithm can solve instances of up to 25 ports to optimality - a very promising result as real-world vessel roundtrips seldom involve more than 20 ports.

Chapter 4 presents a new path based **MIP** model for the Liner Shipping Network Design Problem. The proposed model reduces problem size using a novel aggregation of the demands. A decomposition method enabling delayed column generation is presented. The subproblems have similar structure to **Vehicle Routing Problems**, which can be solved using dynamic programming. An algorithm has been implemented for this model, unfortunately with discouraging results due to the structure of the subproblem and the lack of proper dominance criteria in the labeling algorithm.

In Chapter 5 a novel compact formulation of the liner shipping network design problem is presented based on service flows. The formulation alleviates issues faced by arc flow formulations with regards to handling multiple calls to the same ports, butterfly ports, which previously has not been fully considered by liner shipping network design problem formulations. The method introduces service nodes, together with port nodes in a graph of the problem. Arcs from a port node to a service node represents whether a service is calling a port, and the demand load / unloads, at the port call. This representation allows any number of butterfly ports, which other models have not handled, while generating multiple interconnected services. The model is solved as a Mixed Integer Program. Results are presented for the two smallest instances of the benchmark suite LINER-LIB 2012 presented in Chapter 2.

Chapter 6 investigates the huge impact of operational disruptions in liner shipping on costs and delayed cargo. The Vessel Schedule Recovery Problem evaluates a given disruption scenario and selects a set of recovery actions balancing the trade off between increased bunker consumption and the impact on cargo in the remaining network. It is proven that the Vessel Schedule Recovery Problem is \mathcal{NP} -hard. The model is applied to four real life cases from Maersk Line and solutions are found in less than 5 seconds with results comparable or superior to those chosen by operations managers in real life. Cost savings of up to 58% may be achieved by the suggested solutions compared to realized recoveries of the real life cases.

Chapter 7 considers bunker purchasing for a liner shipping company. The cost for bunker fuel represents a major part of the daily running cost of liner shipping vessels. The vessels, sailing on a fixed roundtrip of ports, can lift bunker at these ports. The ports have differing and fluctuating prices. Contracts are often used to purchase bunker, ensuring supply and often at a discounted price. A contract can supply any vessel in a period and port, and is a shared resource between vessels, which must be distributed optimally to reduce overall costs. The Bunker Purchasing with Contracts Problem has been formulated as a mixed integer program, which has been Dantzig-Wolfe decomposed. To solve it, a column generation algorithm has been developed. The algorithm has been run on a series of real world instances with up to 500+ vessels and 500+ contracts, and provide near optimal solutions. A MIP model cannot solve these instances due to memory requirements.

8.1 Contribution

The contributions of this thesis' research papers are:

- The liner shipping network design benchmark data set of LINER-LIB 2012. This detailed real life data set allows OR researchers to investigate the LSNDP and compare developed methods. A detailed domain description scopes liner shipping network design in terms of OR.
- A basic model for the LSNDP, which captures its core structures. An algorithm has been constructed and implemented, reporting the first results for the LINER-LIB 2012 instances.
- A path based model for a liner shipping problem is presented together with a novel aggregation scheme for the demands, which greatly reduces the number of demands that must be considered.
- A model and algorithm is developed for constructing a single liner shipping service, which can solve instances of real world size, of up to 25 ports. While considering vessel capacity and demand path duration limits. The path durations limits are a key commercial factor in designing services to guarantee customer service levels.
- A novel service flow formulation for a liner shipping network design problem is presented, which allows for multiple interconnected services with any number of recurrent port calls to a port, a problem not solved by other liner shipping models. The implementation finds solutions for two of the LINER-LIB 2012 instances.
- A model and experimental results are presented for the *Vessel Schedule Recovery Problem*, which can be used for disruption management in liner shipping. Savings of up to 58 % are reported compared with actual chosen recoveries.
- A model, algorithm and implementation considering the problem of purchasing bunkers for a fleet of scheduled vessels, while considering a number of bunker contracts, is presented. Managing this huge cost effectively seen over a fleet of vessels, can have significant economical impact.

8.2 Trends in Liner Shipping Optimization Research

Current trends for research in liner shipping network design and other optimization problems in liner shipping are discussed in this section.

8.2.1 Liner Shipping Network Design

Most of the investigated algorithms for liner shipping network design in this thesis are based on optimal methods. The outcome of these studies has been an elaborated understanding of LSNDP problems, their structure and their complexity. But it has also highlighted that optimal methods are unlikely to scale to the problems sizes of real world instances with hundreds of ports and vessels and thousands of demands. To tackle these large problems, a focused effort on developing heuristic methods, handling the complex landscape of operational and commercial constraints found in LSNDP problems, is needed. Such steps have already been taken in the works of Agarwal and Ergun [1], Alvarez [3], Wang et al. [11] and Chapter 2, but further work is needed to include important constraints as transit time limits, empty container repositioning and others. This could be a closer study of state of the art heuristic methods such as Pisinger and Røpke [8]. This approach could lead to good solutions for the largest instances of LINER-LIB 2012.

Should an optimal method be pursued for LSNDP, the method of Stålhane et al. [10] could be investigated, where a column generation method, generates new rows associated with new columns, but still achieves optimality. A decomposition of the service flow model of Plum et al. [9] could be based on this idea.

An even more challenging problem will be to develop methods that can improve on the networks of actual liner shipping network carriers. Due to the complexity of the networks, the uncertainty of

the demand and the scale of the business, it is unlikely that such networks will ever be optimized using a *blank sheet* approach. When the best methods in the literature consider significant parts of the complexity of the problem, tools might be developed that can optimize on real liner shipping networks, in an incremental approach. E.g. by fixing large parts of the network design problem, while keeping the flow of demand open, a tool can optimize effectively on the small free part of the network. This has the advantage in an actual implementation phase, that network changes can be managed, using company practice for network design, and the size of the changes can be limited to what the organization can control and have faith in.

At least two projects are currently following this path. A matheuristic was presented in Brouer and Desaulniers [6], which is run on the LINER-LIB 2012 instances. The column generation inspired local search method, flows the demand in a master problem and creates, destroys and modifies new services in a local search algorithm. The LINER-LIB 2012 was solved using an initial construction heuristic, the solution was then refined by the matheuristic. The results are comparable with those of Chapter 2. This method was further developed in Brouer [5] to take offset in an existing liner shipping network, resembling a global carriers network. The approach would then fix parts of this network and optimize on the free part. The algorithm will find improved solutions in the local search neighborhood of this restricted network design problem. The results are encouraging and validated by the global carrier as being interesting, but still lacking constraints as transit time limits on demands.

A similar approach is being pursued in Wang et al. [11], where a heuristic network design algorithm takes offset in the existing network of a global liner shipping carrier. Parts of the network is fixed and new services are generated and evaluated for entry in the network. The method is reported to improve the network measured on key KPI's.

On a longer term a great challenge for the research in liner shipping network design, is how to face the complex world of competition and collaboration existing in liner shipping, as the number of alliances and vessel sharing agreements between competing carriers increase. This problem opens even greater depths of complexity. For the next years focusing on deterministic one carrier versions of the problem will be challenging, but further in the horizon, even greater tasks await.

Stochastic versions of the liner shipping network design problem could be considered. For instance the demand or parts of the costs could be viewed as uncertain, which would pose a considerable challenge.

This thesis will aid the quest for efficient heuristics in several ways. With suggestions on how to model and solve problem aspects as transit time limits, butterfly services, etc. Good heuristic often have an inspiration from optimal methods, guiding towards problem aspects that can be managed easily with efficient methods, but also highlighting which aspect are hard computational and burdened by degeneration and symmetry. A lesson, from this thesis could be that the combined solving of the flow problem and network design problem in one model, greatly impacts the scalability of the method. Splitting the flow and design problem in two phases as master and sub problems seems needed.

To evaluate the quality of a heuristic, it is powerful to have optimal solutions to compare against, to get bounds. For instance as done in Chapter 5 providing initial bounds on the heuristic solutions of Chapter 2.

Heuristic methods show the way for solving large liner shipping network design problems, by fixing significant parts of the network, while flowing the demand freely and optimizing on the remaining free network.

The aim of Chapter 2, to unite the research in liner shipping network design by formulating a full model of the problem, is probably not reached yet. To illustrate, this thesis proposes a number of different models for the problem highlighting the need for alternate models. Furthermore the formulation of Chapter 2 has shortcomings, as being exponential and not including constraints as time limits. The quest for the ultimate liner shipping network design model is still open.

8.2.2 Other Liner Shipping Problems

Interesting challenges of great practical relevance exist in other problems met in liner shipping. An initial draft of a project for a feeder network design problem, considering ports served *first come, first serve* is presented in Appendix Chapter 10. Other interesting research projects could extend on the fleet asset management approach in Alvarez et al. [4] investigated for bulk shipping. This could be extended to consider liner shipping. Little research has been done on liner shipping yield management, which could be interesting to investigate. As could the game theoretic approaches of Agarwal and Ergun [2] and their effect on important Vessel Sharing Agreements and carrier alliances. Eefsen and Cerup-Simonsen [7] considers the effects of inventory costs of goods transported in containers, and the economic impact of speed reduction, an interesting approach that could be investigated further in liner shipping network design.

8.3 Final Remarks

This thesis has investigated methods for the liner shipping network design problem. The domain has been described in the terms of an OR professional. The LINER-LIB 2012 benchmark instances have been presented, and are already showing the first signs of opening the research field to new researchers. Four different network design models have been formulated and implemented, investigating important aspects of the problem as scalability, decomposition, butterfly ports and transit time on demands. These have provided results for the LINER-LIB 2012 instances. Two decision support methods for bunker purchasing with contracts and vessel schedule recovery, have been proposed, implemented and tested on real problems with promising results.

These works highlight the relevance and importance of developing optimization methods for planning and execution problems met in liner shipping.

Bibliography

- [1] Richa Agarwal and Ozlem Ergun. Ship scheduling and network design for cargo routing in liner shipping. *Transportation Science*, 42(2):175–196, 2008.
- [2] Richa Agarwal and Ozlem Ergun. Mechanism design for a multicommodity flow game in service network alliances. *Operations Research Letters*, 36(5):520 – 524, 2008. doi: 10.1016/j.orl.2008.04.007.
- [3] J. Fernando Alvarez. Joint routing and deployment of a fleet of container vessels. *Maritime Economics and Logistics*, 11:186–208, 2009.
- [4] J. Fernando Alvarez, Panagiotis Tsilingiris, Erna Engebretsen, and Nikolaos M. P. Kakalis. Robust fleet sizing and deployment for industrial and independent bulk ocean shipping companies. *INFOR*, Special Issue on Maritime Transportation.
- [5] Berit D. Brouer. *Liner Service Network Design*. PhD thesis, The Technical University of Denmark, Management Engineering, 2012.
- [6] Berit D. Brouer and Guy Desaulniers. A matheuristic for the liner shipping network design problem. In *Proceedings of the second International Conference on Logistics and Maritime Systems, Bremen*, 2012.
- [7] Thomas Eefsen and Bo Cerup-Simonsen. Speed carbon emissions and supply chain in container shipping. In *Proceedings of the Internatioanl Association of Maritime Economists Conference, Lisbon*, July 2010.
- [8] David Pisinger and Stefan Røpke. Large neighborhood search. In *Handbook of metaheuristics*, pages 399–419. Springer, 2010.
- [9] Christian E. M. Plum, David Pisinger, and Mikkel M. Sigurd. A service flow model for the liner shipping network design problem. Submitted for special issue of European Journal of Operational Research on *Maritime Logistics*, 2012.
- [10] Magnus Stålhane, Henrik Andersson, Marielle Christiansen, Jean-François Cordeau, and Guy Desaulniers. A branch-price-and-cut method for a ship routing and scheduling problem with split loads. *Computers & Operations Research*, 2012.
- [11] Shuaian Wang, Zhiyuan Liu, and Qiang Meng. Systematic network design for liner shipping services. To appear in *Transportation Research Record*.

Part V
Appendix

Chapter 9

Bachelor and Master Thesis Projects

Following bachelor and master thesis projects has been supervised as part of the project.

- Jakob Dirksen, Master Thesis, *Disruption Management in Liner Shipping*, Supervisor: David Pisinger, Co-supervisors: Christian E. M. Plum, Berit Løftstedt, Bo Vaaben, June 27, 2011.
- Hanne Reck, Master Thesis, *Cargo Routing at a Major Liner Shipping Company*, Supervisor David Pisinger, Co-supervisor Christian E. M. Plum, December 1, 2011.
- Federico Farina, Master Thesis / Tesi di Laurea in Ricerca Operativa, *Bunker Optimization with Contracts*, Supervisor Prof. Silvano Martello, Prof. David Pisinger, Co-supervisors: Christian E. M. Plum, Peter N. Jensen, 2012.
- Christian Lous, Bachelor Project, *Optimizing the construction of strings in liner shipping* Supervisor's David Pisinger, Line B. Reinhardt, Christian E. M. Plum, Mikkel M. Sigurd January 25, 2013.
- Anders Østergaard, Bachelor Project, *Bunker Reserve Management in Liner Shipping* Supervisor's Kouros M. Rasmussen, Co-supervisors: Christian E. M. Plum and Peter N. Jensen, March 12, 2013.

Chapter 10

Models for Feeder Vessel Scheduling with Split Deliveries and Time Windows

Christian E. M. Plum^{*, †} Line B. Reinhardt^{*} David Pisinger^{*} Mikkel M. Sigurd[†]

^{*}Department of Management Engineering, Technical University of Denmark,
Produktionstorvet, Building 426, DK-2800 Kgs. Lyngby, Denmark
cemp@man.dtu.dk, lbre@man.dtu.dk pisinger@man.dtu.dk,

[†]Maersk Line
Esplanaden 50, DK-1098 København K.
mikkel.muhldorff.sigurd@maersk.com

1

Introduction

A global liner shipping network is made up of intercontinental routes, served by very large vessels carrying cargo between the major hubs of the world. In hubs the containers are often transhipped to smaller vessels delivering the cargo to the end destination. This paper considers how to construct such a feeder network, from a single hub under a range of operational constraints. We denote the non-hub ports: *Outports*. Demand between outports is assumed negligible.

We consider a feeder network which contains multiple time windows for berthing at some of the ports and which allows split deliveries and pickups. The problem presented is inspired by a real-life routing problem originating in West Africa, where all transports to and from the region arrives through a hub in Spain and the demands between the different ports in West Africa is insignificant. Additionally vessels will have a start time, from when they can be used, as will the demand have an earliest and latest time it can be picked up. Port draft limits must be respected when entering and leaving a port. The Feeder vessel scheduling with split deliveries and time windows, in short FVSP, is a version of the vehicle routing problem with divisible simultaneous pickup and delivery and time windows where liner shipping specific constraints are included.

The vehicle routing problem with simultaneous pickup and deliveries (VRPSPD) was introduced by [4]. In [6] a short description of the VRPSPD is given, as a variation of the vehicle routing problem with backhauls which they solve using a unified heuristic. Another version of the vehicle

¹Accepted for *Tristan 8 Conference*

routing problem, the split delivery problem (SDVRP), has simultaneously been investigated by [1] and [2]. In [1] the SDVRP with time windows (SDVRPTW) has been solved using an exact branch, price and cut method for problems with up to 100 locations.

A combination of the two problems where the VRPSPDP allows split deliveries (VRPSPDSL) seems to only have been considered recently. [7] solves the pickup and delivery problem with time windows and split loads with a branch and price and cut method for the tramp vessel routing problem.

The feeder network for West Africa modeled here is a vessel routing problem with divisible deliveries and pickups (VRPDPD). This is similar to the VRPSPDSL, however, in VRPDPD all demands originate or terminate at a single depot, see [5]. The model also includes a set of available berthing time windows for some ports and a time windows for the when the demand is to be picked up to ensure some regularity for the customers. This has a connection to the Vehicle Routing Problem with Multiple Time Windows, which was considered in Jong et al. [3].

10.1 Mathematical Model Of FVSP

For now we assume a homogeneous fleet of vessels. We have the parameters:

- K is the set of commodities, where K^m, K^x is the set of import, respectively export commodities to / from the outport, giving $K = K^m \cup K^x$. The demand has a demand volume p_k / d_k for pickup / delivery. An earliest and latest time a demand must be picked up at port $p(k)$ is t_e^k, t_l^k . M_T is the latest time at which any port can be visited. Let $f(k) = i$ be the outport of the commodity.
- V is the set of vessels available. Each vessel has a capacity Q_v and an earliest time it can be used from the hub, t^v .
- N is the set of ports and 0 is the Hub. $N^+ = N \cup \{0\}$. D_i is the maximal load of containers allowed when visiting port i . c_i , is the cost of calling port i . A port can have a set, $g \in i_G$ of time windows in which a visit can be started, between a_g^i and b_g^i . If $i_G = \emptyset$ the port can freely be served. Each port has a service time t_s^i . A is the set of arcs, $t_{v,ij}$ is the time and $c_{v,ij}$ is the cost of edge ij when sailed by vessel v . Let $t_{v,ij}^l = t_{v,ij} + t_s^j$.

Variables: Let $T_{v,i}$ be the time vessel v arrives at port i . The binary variable $x_{v,ij}$ is 1 if the arc ij is used by vessel v . The binary variable $y_{v,i}$ is 1 if the vessel v visits port i . The binary variable $z_{v,i,g}$ is 1 if the vessel v visits port i and uses time window g . $\delta_{v,k}$ is the amount of demand $k \in K^m$ on vessel v . $\eta_{v,k}$ is the amount of demand $k \in K^x$ on vessel v . $L_{v,ij}$ is the load along edge ij on vessel v . $s_{v,k}$ is 1 if some commodity k is transported on vessel v .

The problem can be formulated as the following mixed-integer linear program. The objective (10.1) minimizes arc cost and port call cost. Constraints (10.2) and (10.3) ensures that all demand is satisfied and constraint (10.4) links variables x and y . A vessel can only service demand at called ports, and no more than available is handled by constraint (10.5) and (10.6). Vessel tours are balanced by constraint (10.7), and vessel arrival times by constraint (10.8). Berth time windows are managed by constraints (10.9) - (10.11). Vessel load is considered by constraints (10.12) - (10.14).

$$\min \sum_{v \in V} \sum_{ij \in A} (c_{v,ij} + c_i) x_{v,ij} \quad (10.1)$$

$$s.t. \quad \sum_{v \in V} \delta_{v,k} \geq d_k \quad \forall k \in K^m \quad (10.2)$$

$$\sum_{v \in V} \eta_{v,k} \geq p_k \quad \forall k \in K^x \quad (10.3)$$

$$\sum_{j \in N^+} x_{v,ij} = y_{v,i} \quad \forall i \in N, v \in V \quad (10.4)$$

$$\sum_{k \in K^m, f(k)=i} \eta_{v,k} \leq d_k y_{v,i} \quad \forall i \in N^+, v \in V \quad (10.5)$$

$$\sum_{k \in K^x, f(k)=i} \delta_{v,k} \leq p_k y_{v,i} \quad \forall i \in N^+, v \in V \quad (10.6)$$

$$\sum_{j \in N^+} (x_{v,ij} - x_{v,ji}) = 0 \quad \forall i \in N, v \in V \quad (10.7)$$

$$T_{v,j} + M(1 - x_{v,ij}) - T_{v,i} - t_{v,ij}^l > 0 \quad \forall i, j \in A, j \neq 0, v \in V \quad (10.8)$$

$$\sum_{g \in i_G} a_g^i z_{v,i,g} \leq T_{v,i} \leq \sum_{g \in i_G} b_g^i z_{v,i,g} \quad \forall i \in N, v \in V \quad (10.9)$$

$$\sum_{g \in i_G} z_{v,i,g} = y_{v,i} \quad \forall i \in N^+, v \in V \quad (10.10)$$

$$\sum_{v \in V} z_{v,i,g} \leq 1 \quad \forall i \in N^+, g \in i_G \quad (10.11)$$

$$\sum_{j \in N} (L_{v,ij} - L_{v,ji}) = \sum_{k \in K, f(k)=i} (\eta_{v,k} - \delta_{v,k}) \quad \forall i \in N, v \in V \quad (10.12)$$

$$\sum_{j \in N} L_{v,0j} = \sum_{k \in K} \delta_{v,k} \quad \forall v \in V \quad (10.13)$$

$$\sum_{j \in N} L_{v,j0} = \sum_{k \in K} \eta_{v,k} \quad \forall v \in V \quad (10.14)$$

$$L_{v,ij} \leq \text{Min}(D_i, D_j, Q_v) \quad \forall i, j \in A, v \in V \quad (10.15)$$

$$T_{v,0} \geq t^v \quad \forall v \in V \quad (10.16)$$

$$\delta_{v,k} + \eta_{v,k} \leq (d_k + p_k) s_{v,k} \quad \forall k \in K, v \in V \quad (10.17)$$

$$t_e^k \leq T_{v,p(k)} + t_e^k (1 - s_{v,k}) \quad \forall v \in V, k \in K \quad (10.18)$$

$$T_{v,p(k)} \leq t_l^k + M_T (1 - s_{v,k}) \quad \forall v \in V, k \in K \quad (10.19)$$

$$z_{v,i,g}, x_{v,ij}, y_{v,i}, s_{v,k} \in \{0, 1\} \quad \forall i, j \in A \quad (10.20)$$

$$L_{v,ij}, T_{v,i}, \delta_{v,k}, \eta_{v,k} \in \mathbb{R}^+ \quad \forall v, k, i, j \quad (10.21)$$

Draft and capacity by constraint (10.15). Constraints (10.16) - (10.19) handles vessel and demand start and end times. Variable bounds are specified by constraints (10.20) and (10.21).

10.2 Outlook

Real world problem instances of the FVSP holds up to 30 outports. In order to solve problem instances of this size we expect it necessary to use state of the art Branch and Cut and Price methods used for earlier mentioned related rich VRP problems. It would seem interesting to decompose model (10.1) on variables $x_{v,ij}$ generating columns representing routes for a vessel. The

commodity constraints (10.2) and (10.3) as well as the berth constraint (10.11) will remain in the master problem, the remaining constraints will enter the subproblems. Independent subproblems will exist for each vessel as they will have different start times in constraint (10.16), it will be investigated if these can be aggregated. Each sub problem will be a special case of the Elementary Shortest Path Problem with Resource Constraints (ESPPRC), where the resources must consider, capacity, draft limits, multiple port time windows and vessel and demand start and end times.

Bibliography

- [1] C. Archetti, M. Bouchard, and G. Desaulniers. Enhanced branch and price and cut for vehicle routing with split deliveries and time windows. *Transportation Science*, 45:285–298, 2011.
- [2] M. Dror and P. Trudeau. Split delivery routing. *Naval Res. Logist.*, 37:383–402, 1990.
- [3] C. de Jong, G. Kant, and A. van Vliet. On Finding Minimal Route Duration in the Vehicle Routing Problem with Multiple Time Windows. Technical Report, Department of Computer Science, Utrecht University, The Netherlands 1996.
- [4] H. Min. The multiple vehicle routing problem with simultaneous delivery and pick-up points. *Transportation Research Part A: General*, 23:377–386, 1989.
- [5] S. Mitra. A parallel clustering technique for the vehicle routing problem with split deliveries and pickups. *Journal of the Operational Research Society*, 59:1532–1546, 2008.
- [6] P. D. Ropke, S. A unified heuristic for a large class of vehicle routing problems with backhauls. *European Journal of Operational Research*, 171:750–775, 2006.
- [7] M. Stålhane, H. Andersson, M. Christiansen, J.-F. Cordeau, and G. Desaulniers. A branch-price-and-cut method for a ship routing and scheduling problem with split loads. *Computers and Operations Research*, 39:3361–3375, 2012.