

Shipment Scheduling In Hub Location Problems

by

Mobina Masaeli

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Science
in
Management Sciences

Waterloo, Ontario, Canada, 2017

© Mobina Masaeli 2017

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In this thesis, we incorporate shipment scheduling decisions into hub location problems. Our aim is to determine optimal locations of hubs, hub network structure, and the number of vehicles to operate on the hub network as well as the time period of dispatching each vehicle from a hub. We develop mathematical models for different versions of this problem. We initially propose a mixed-integer shipment scheduling hub network design model where the costs of holding freight are negligible. We then expand the model to keep track of the holding decisions where the holding costs are not negligible. We further analyze the shipment scheduling model with holding costs when different types of vehicles are available to operate on the inter-hub links. We investigate the impact of shipment scheduling decisions and holding costs on hub network configurations, routing decisions, and total cost of the network. We solve the models on instances from a new USAF dataset with real data.

Keywords: Hub location, Shipment scheduling, Economies of scale, Holding costs.

Acknowledgements

Foremost, I would like to express my deepest gratitude to my supervisors, Professors Sibel Alumur Alev and James Bookbinder, for their continuous support and guidance throughout my Master's journey. Professor Alumur Alev, I am very grateful to you for doing research on hub location area. I truly enjoyed and learned a lot from you. I appreciate all insightful discussions and constructive suggestions. Professor Bookbinder, thank you for your invaluable advice, and sharing your broad knowledge and experience. I also really appreciate all your feedbacks to improve my writing. Thank you for placing your confidence in my abilities and talents, your warm encouragement, and providing me a Teaching Assistant opportunity.

Professors Bookbinder and Alumur Alev, I truly appreciate all your continuous help and great understanding specially when obstacles came up during my journey.

My sincere thanks go to Professor Abouee Mehrizi and Professor Koo. Professor Abouee, I am very thankful for all your kind help, advice, and providing me with Teaching Assistant opportunities. Professor Koo, I do appreciate all your help and support when I severely needed it.

Moreover, many thanks to Professor Elhedli and Professor Gzara for accepting to act as the readers of this thesis and providing useful feedbacks and comments. Professor Gzara, I also genuinely appreciate all your knowledge that you patiently shared with us. I am grateful to have you as my teacher.

I also would like to express my gratitude towards my friends for all their support. Nasrin, thanks for your continuous support and encouragement. Over the years, our friendship has continued to thrive even though we are thousands of miles apart. My sweetest and truest friends; Chen, Dani, Niloofar, and Yidan, I feel lucky to have your support through thick and thin. Thanks for all the great time we have had together. Forough, Gita, Ata, and Akram thanks for all your help and happy moments we have had. I warmly thank

my adorable officemates, Aliaa and Xioyang, for lifting me up with their wise words and comforting me with their hugs. Gwahram and Daren, thank you so much for all insightful discussions and your help. My special thank goes to Nasser for being there for me, cheering me up, and his endless support throughout this journey in spite of the long distance between us.

Last but not least, my warmest thanks go to my loving and caring parents for their endless amounts of love, unconditional support, sacrifices, and prayers. You have always been there for me and words cannot express my gratitude and appreciation to you.

To my professors and family.

Table of Contents

List of Tables	ix
List of Figures	x
1 Introduction	1
2 Literature Review	4
2.1 Classical Hub Location Problems	4
2.2 Economies of Scale in Hub Location Problems	6
2.3 Multi-Period Hub Location Problems	9
2.4 Service Network Design	11
2.5 Hub Location and Scheduling	14
3 Problem Definition and Mathematical Formulations	16
3.1 Hub Location Model with stepwise constant cost function	16
3.2 Hub Location and Shipment Scheduling Problem	19
3.2.1 Problem formulation with no holding costs	21
3.2.2 Problem formulation with holding costs	23

3.3	Size of the Models	25
4	Computational Analysis	28
4.1	Parameter Estimation and Data Generation	28
4.2	SHLPSCF Model Results	32
4.3	SSHLP Model Results	33
4.4	Results for SSHLPH Model	36
4.5	Sensitivity Analysis of Demand and Length of Planning Horizon	44
4.6	Results of SSHLPHV Model	46
5	Conclusion	49
	References	52
	APPENDIX	59

List of Tables

3.1	Number of variables and constraints in the formulations.	27
4.1	The capacity of airplanes and corresponding fixed costs.	30
4.2	Cost comparison between airplane and truck.	30
4.3	Holding cost classes.	31
4.4	Results of SHLPSCF model on the 15-node network	32
4.5	Results of SHLPES model on the 20-node network	33
4.6	Results of SSHLP model on the 15-node network.	34
4.7	Results of SSHLP model on the 20-node network.	34
4.8	Results for SSHLPH model on the 15-node network.	37
4.9	The results of the SSHLPH model on the 20-node network.	41
4.10	Results of SSHLPH model with B747-F when demands are increased by 50%.	44
4.11	Results of operating B747-F within a planning horizon of four time periods.	45
4.12	Results of the SSHLPHV model with two aircraft types.	46
1	20 cities with corresponding node number	59

List of Figures

2.1	Piecewise linear approximation function (O’Kelly and Bryan (1998))	6
2.2	The cost as a function of flow through link (k, m) (Podnar et al. (2002))	7
2.3	Cost structure with fixed costs (Kimms (2006))	8
2.4	Stepwise constant cost function	9
4.1	Location and allocation decisions on 15-node USAF dataset by operating B747-100 airplane under $f_k = 5,000$	39
4.2	Resulting hub configurations with B747-F airplane and $f_k = 10000$	42
4.3	Resulting hub network of 15-node USAF dataset under $f_k = 5,000$, second time period [1,2).	48

Chapter 1

Introduction

A special type of networks, so called *hub-and-spoke* network, is designed to serve flows of commodities, people, or information between a large set of origins and destinations. Hubs are central transshipment facilities that provide switching, sorting or connecting, and consolidation/break-bulk functions (O’Kelly and Bryan, 1998). This configuration allows carriers to exploit economies of scale on inter-hub transportation costs through bundling flows. Hub location models are widely used for many applications such as air transportation, ground freight transportation, postal, and parcel delivery. These models provide an optimal structure for the hub network and routing decisions with minimum total costs.

Operational characteristics of parcel delivery companies stimulate the primary motivation of this study. Parcel delivery networks use a hub-and-spoke network structure where branch offices are typically referred to as origins/destinations or spoke nodes. Parcels are routed from the originating branch office through at least one hub, known as an operation center, to the consignee branch office. Every branch office is allocated to a single operation center. This is referred to as *single allocation* in the literature. Parcels dropped at branch offices customarily travel relatively short distances to arrive at the origin’s operation center.

They are consolidated with other parcels, loaded onto vehicles with large capacities, and destined toward consignees' operation centers. Each parcel delivery company ordinarily uses its own private fleet of vehicles to serve the inter-hub links.

Growing competition leads delivery companies, such as Fedex, UPS, and DHL, to provide express services such as same-day or next day-delivery. Such services are provided in response to urgent freight and just-in-time production schedules in which cost efficiency is also important. Parcel delivery networks adopt a hub-and-spoke network structure to reduce transportation costs. Such networks take advantage of economies of scale through consolidation of freight shipments at hub facilities. To do so, delivery companies need to schedule the shipments in between hubs. Individual shipments might be held at hub facilities to dispatch a consolidated load at a later time. In this way, by efficiently scheduling the shipments, economies of scale can be exploited. Introducing shipment scheduling into hub location problems is essential to efficiently design hub networks, as consolidation decisions may have an effect on the optimal hub locations. Nevertheless, this issue has not been addressed in the literature to date.

In this thesis, we incorporate shipment scheduling decisions into hub location problems. Our aim is to determine optimal locations of hubs, optimal hub network structure, and the optimal number of vehicles to operate on the hub network as well as the time of dispatching each vehicle from a hub. We develop mathematical models for different versions of this problem. We initially propose a mixed integer shipment scheduling hub network design model where the costs of holding parcels are negligible. We expand the model to keep track of the holding decisions where the holding costs are not negligible. We aim to investigate the impact of shipment scheduling decisions and holding costs on optimal hub configurations, routing decisions, and total cost of the network. We further analyze the shipment scheduling model with holding costs when different types of vehicles are available to operate on the inter-hub links. We solve the models on instances from a new USAF dataset and interpret the results.

The remainder of the thesis is organized as follows. Chapter 2 addresses the relevant literature. In Chapter 3, we provide the formal problem definition, notations and mathematical formulations for the hub location and shipment scheduling problem. Chapter 4 contains the results of the computational experiments with the proposed models. We present our concluding remarks in Chapter 5.

Chapter 2

Literature Review

In this chapter, we review hub location problems in five main sections. Section 2.1 addresses the hub location problems in general contexts. In four following sections, we specifically address the bodies of literature most relevant to this research. We also point out how this thesis can be distinguished from the literature. In Section 2.2, we focus on approaches to modeling economies of scale in hub location problems. Multi-period hub location problems are studied in Section 2.3. We review service network design problems in Section 2.4. Section 2.5 provides the studies on shipment scheduling in hub location problems.

2.1 Classical Hub Location Problems

The work of O’Kelly (1986a,b) pioneered the research on hub location problems. O’Kelly (1986a) proposed location models for two cases; continuous and discrete single hub networks, and a continuous two-hub network, in order to serve a set of interacting nodes. In addition, he showed that the single-hub location problem is equivalent to a Weberian least cost location model. O’Kelly (1986b) developed a heuristic method to find locally optimal locations of hubs, and realized that applying a discount factor on transportation costs of

inter-hub links furthers a wider hub facility configuration. The publication of the seminal paper by O’Kelly (1987), who proposed the first mathematical formulation for a general p -hub median problem, has triggered several papers that addressed hub location problem with a fixed or variable number of hubs (see e.g., Campbell, 1996; Skorin-Kapov et al., 1996; Ernst and Krishnamoorthy, 1999; Ebery, 2001; Marìn et al., 2006). p -hub location problems aim at serving the demands through a predetermined number of hub facilities (p) at minimum total transportation cost. The p -hub median models usually ignore the fixed costs of opening hub facilities. O’Kelly (1992) introduced the single allocation hub location problem with fixed costs, where the number of hubs is a decision variables. He formulated the problem as a quadratic integer program and proposed a two-step procedure to find a good upper and a tight lower bound.

Early surveys on hub location problems go back to the works of Campbell (1994) and O’Kelly and Miller (1994). The interested reader may refer to more recent surveys (see e.g., Campbell et al., 2002; Alumur and Kara, 2008; Campbell and O’Kelly, 2012; Farahani et al., 2013; Contreras, 2015) for thorough reviews on hub location problems.

The literature contains a few data sets that have been used extensively with hub location models. O’Kelly (1987) introduced the Civil Aeronautics Board (CAB) data set that consists of the airline passenger interactions between 25 cities in United States in 1970. The unit transportation costs in the CAB dataset correspond to the air distances between cities. Three decades later, Ernst and Krishnamoorthy (1996) contributed Australia Post (AP) data that includes 200 postal districts for the postal (ground) delivery network in Sydney, Australia. Tan and Kara (2007) introduced the Turkish dataset that contains 81 demand centers. This dataset includes travel distances, travel times, flow values, and fixed link costs between 81 cities in Turkey. All these three datasets are readily accessible in the OR Library (Beasley, 1990). Additionally, Karimi and Setak (2016) generated the Iranian Roads Network (IRN) dataset that includes 20 cities. They calculated the fixed cost of hubs by the procedure proposed by Tan and Kara (2007). They also used a population-

based function to find origin-destination demands. The IRN dataset is available in Karimi (2016).

2.2 Economies of Scale in Hub Location Problems

The economies of scale is known as one of the primary motivations for designing hub networks. Hub networks allow carriers to take advantage of economies of scale through consolidation of human, commodity or information flow at hubs. Most of the hub location models in the literature apply a constant discount factor α , where $0 \leq \alpha < 1$, to the unit transportation costs between hubs. However, modeling economies of scale by a discount factor α which is independent of flows contradicts the underlying motivation of hub networks. Some studies attempted to resolve this drawback by modeling economies of scale in different ways.

O’Kelly and Bryan (1998) replaced the constant factor α by a concave function, $1 - \theta X_{km}^\beta$ where $0 < \theta \leq 1$, $\beta > 0$ and X_{km} indicates $\frac{\text{flow passing through interhub link (m,k)}}{\text{total flow}}$. They approximated the nonlinear function by a piecewise linear function as shown in Figure 2.1.

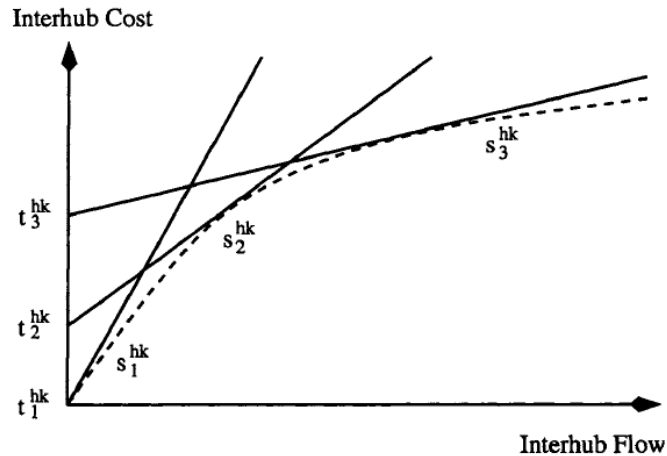


Figure 2.1: Piecewise linear approximation function (O’Kelly and Bryan (1998))

Bryan (1998) further extended the model discussed in O’Kelly and Bryan (1998) by considering economies of scale on all links. He used the same piecewise linear approximation method to model economies of scale.

Horner and O’Kelly (2001) also applied a nonlinear cost function dependent on link flow that rewards economies of scale on all links in the network. Unlike O’Kelly and Bryan (1998), however, they analyzed this nonlinear concave function instead of its piecewise linear approximation function.

Podnar et al. (2002) used a threshold-based discounting factor to formulate economies of scale. They assumed that if the amount of traffic on a given link is large enough (or exceeds a predetermined value), then the cost of this flow is discounted by a factor α . Figure 2.2 demonstrates the cost function of flow routed through a given link (k, m) .

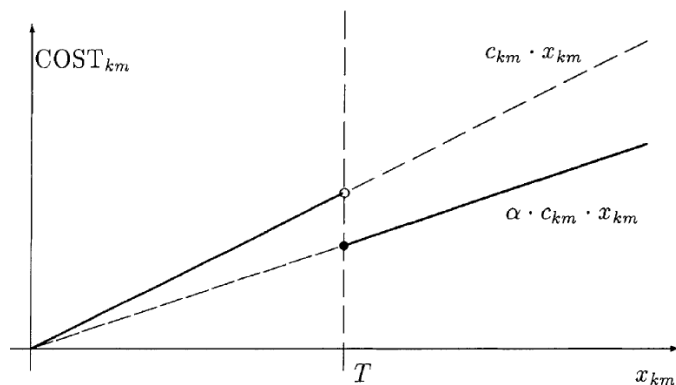


Figure 2.2: The cost as a function of flow through link (k, m) (Podnar et al. (2002))

Kimms (2006) proposed an alternative way to model economies of scale by considering flow-independent fixed costs and a unit total cost. The economies of scale is simply incorporated into the cost model, such that an increase in flow reduces the unit handling cost. He considered the economies of scale for all links in a multiple allocation p-hub median problem with direct services. He further took capacity constraints of trucks into account. Operating each truck on a given link incurs a fixed cost. Figure 2.3 illustrates the cost

structure proposed by Kimms (2006), where c_{ih}^f , c_{ih}^t , and c_{ih}^v denote a fixed cost, a truck fixed cost, and a unit handling cost, respectively.

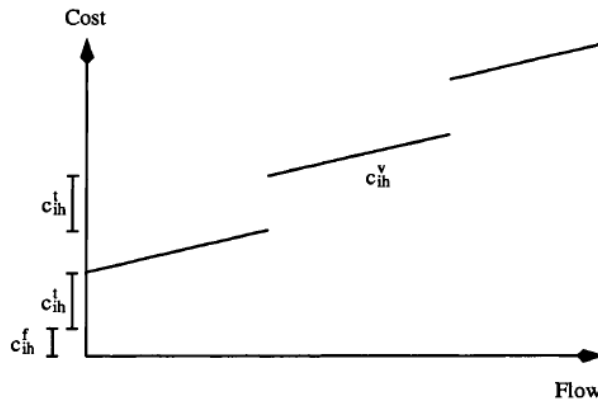


Figure 2.3: Cost structure with fixed costs (Kimms (2006))

Kimms (2006) also considered a more general setting of the model where M transportation modes are available to operate on each link. He decided on the required number of vehicles of each type. They extended the model by adding a set of constraints on the number of modes serving the links, as well.

Cunha and Silva (2007) introduced the discount factor as a function of the total amount of freight between hubs. The function might take any form including continuous, discrete, nonlinear, linear, symmetric, and asymmetric. They designed a hub network for a less-than-truckload company in Brazil. They proposed a genetic algorithm (GA) with an efficient local search to solve the model, since GA allows high degree of flexibility to define the discount-factor function.

Chen et al. (2014) analyzed a hybrid hub-and-spoke network for less-than-truckload freight by minimizing the total operation costs: transportation costs and transfer costs at hubs. They proposed an incremental quantity discount factor α . They further formulated economies of scale by the weighted average unit discount rate, based on the discount rates of intervals and the amount of flows that adopt the discounts of the intervals. They also

expanded the model by imposing capacity constraints on links.

In this thesis, we introduce a stepwise constant cost function to formulate transportation costs on inter-hub links. This function includes two costs: i) fixed cost of dispatching a vehicle and ii) operational cost per vehicle that varies with distance. This cost model, together with the vehicle capacities, can simply overcome the weakness of using a constant flow-independent discount factor. We illustrate our fixed cost-oriented formulation in Figure 2.4, where η indicates the vehicle capacity.

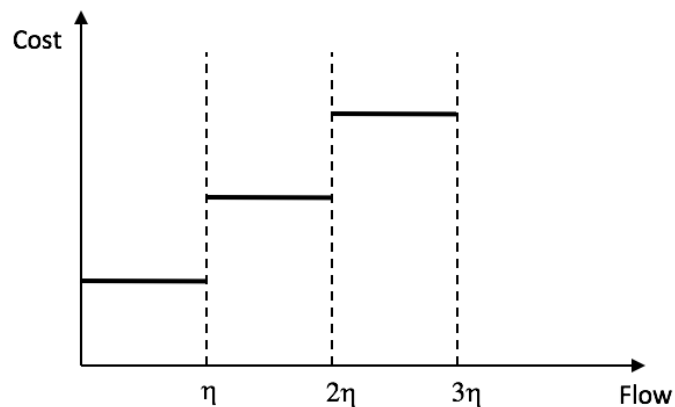


Figure 2.4: Stepwise constant cost function

Since benefiting from economies of scale requires freight to be consolidated at hub facilities, we also introduce the dimension of time into our hub location problem. The next section addresses dynamic hub location problems.

2.3 Multi-Period Hub Location Problems

Research on dynamic hub location problems began with the pioneering work of Campbell (1990). He proposed a continuous-approximation model to locate/relocate freight carrier terminals in response to increasing demand. The shipment origins and destinations are

randomly dispersed over a fixed service region. The model aimed to keep low cost terminal configurations through the trade-off between transportation costs, location and relocation costs of the terminals. He examined the performance of myopic strategies which are based on current demands and terminal locations. The results showed that a myopic strategy is nearly optimal when relocation is not expensive, and provides lower and upper bounds on the transportation cost for the optimal strategy.

Contreras et al. (2011) analyzed a dynamic (multi-period) uncapacitated multiple allocation hub location problem over a finite discrete time planning horizon. For given sets of commodities in each time period, all single-period demand of a particular commodity is fully routed through a distinct set of hubs toward the destinations. They developed a branch-and-bound (BB) algorithm that yielded optimal solutions for a set of benchmark instances with up to 100 nodes and 10 time periods. A Lagrangian relaxation approach was employed to provide tight lower and upper bounds at each node of the BB tree. To reduce the number of nodes of the tree, they used an effective partial enumeration phase to explore the solution space.

Classical hub location models primarily address long-term strategic location and allocation decisions. Gelareh et al. (2015) illustrated certain applications such as maritime and land transport dealing with seasonal decisions. Since hub facilities are leased, the locations of the hubs can change when the current contracts conclude. They developed a mathematical model for a multi-period uncapacitated multiple allocation hub location problem with a budget constraint. The mixed-integer linear model is solved by CPLEX for small-sized instances. They proposed a meta-heuristic and a Bender's decomposition algorithm to solve large scale instances.

Alumur et al. (2016) studied other multi-period versions of hub location problems. Taking construction time into consideration, they presumed that the hub network is progressively constructed over time and its operating capacity may increase, gradually. They proposed mixed-integer linear programming formulations for multi-period modular-

capacitated hub location problems for both single and multiple allocation versions. Reducing hubs' capacities and closing existing hubs are not allowed. Operating inter-hub links incurs fixed costs and might change over time, even if the set of hubs remains the same. They improved the models through several sets of valid inequalities. Using CPLEX software, they solved the models for moderately-sized instances. They developed a measure for the value of the multi-period solution that compares the optimal value of the multi-period formulation with feasible one achieved by solution to static (single period) counterpart of the problem. Computational tests revealed the importance of integrating the time dimension into hub location problems.

Analogous to multi-period hub location models, in this study we assume that the planning horizon is discretized into a set of time periods. However, as the planning horizon is assumed to be short (e.g., a couple of days), the resulting structure of the hub network remains steady over all time periods. Moreover, it is assumed that every parcel is delivered to a hub facility (consignee operation centers) by the end of the planning horizon. To take advantage of economies of scale through shipment scheduling, any parcel might be held at a hub facility for one or more time periods. Hence, as opposed to the static hub location model, here the flows on inter-hub links might change over different time periods.

2.4 Service Network Design

Another relevant subject area might be the service network design problem (SNDP). SNDP addresses operational and tactical decisions associated with vehicle routing, service schedules, consolidation, repositioning of empty vehicles, and vehicle stops as well as waiting times using preexisting terminals. SNDP problems provide consolidation-based transportation systems with a service plan for all demands. SNDP problems are NP-Hard, and are generally formulated as mixed-integer multi-commodity, capacitated network models (see e.g., Dall'Orto et al., 2006; Wang and Lo, 2008).

Early research on SNDP goes back to the work of Crainic and Rousseau (1986). They analyzed the multi-mode, multi-commodity freight transportation problem when the same authority regulates or plans the supply of transportation services, and also makes the routing decisions. They defined a set of possible itineraries (the sequence of terminals for a given traffic class) and efficiently solved the model by a decomposition-and-column generation-based algorithm.

Armacost et al. (2002) proposed a new formulation for the express shipment service network design problem using composite variables to solve realistic size instances. They combined aircraft route and package flow variables as a single composite variable. The LP relaxation of the composite-based formulation provides stronger lower bounds. The model was applied to the Next Day Air network of the United Parcel Service (UPS).

Sung and Song (2003) introduced cross-docking (CD) centers into a service network design problem. They simultaneously decided on the locations of CD centers and the number of vehicles for each origin-CD and CD-destination direct service. Two types of vehicles are considered to serve the links. Vehicle assignments are subject to restrictions on capacity and service time. The authors formulated the problem by using a path-based model. They proposed tabu-search-based algorithm along with tightened lower bounds that are derived by a set of strong valid inequalities.

Lai and Lo (2004) proposed a single ferry type service network design problem with ferry capacity constraints. They considered two types of time-space networks: i.e. a ferry and passenger flow network, to formulate the problem. Origin-destination demands and their arrival patterns at the piers are assumed to be known. Their mixed-integer model determined the optimal fleet size, ferry routes, and departure schedules. They developed a two-phase heuristic algorithm that exploited the polynomial-time performance of the shortest path algorithm. They applied the heuristic to efficiently solve instances of ferry services in Hong Kong under two scenarios: morning peak period and entire day.

Andersen et al. (2009) studied multiple fleet coordination within service network design. They aimed at synchronizing both new services and existing connected services to determine the departure times of each. A service is identified by its origin, possible intermediate stops, and its destination. Characteristics such as capacity, speed, and priority can distinguish particular services. They solved the mixed-integer programming model using a commercial solver, applying the model to real-life data for a network of 17 nodes within a planning horizon of 84 time periods of 2 hours. The authors obtained solutions within an optimality gap of 10% after 24 hours of running time, as the LP-relaxation of the model is weak. To speed up the solution process, they used additional strong forcing constraints and also eliminated some specific arcs. The model eventually was optimally solved within 12 hours.

Pedersen et al. (2009) addressed asset management considerations in a generic service network design problem. They developed both arc- and cycle-based formulations for a designed-balanced service network design problem. The aim was to transport origin-destination demands within their time windows. The model picked services, determined their schedules and routed assets to be available at terminals. Here, the term “asset” may refer to crews, tractors, ships, trucks, etc. The authors proposed a tabu search algorithm that includes two local-search phases to solve the arc-based formulation. They solved several network design problem instances from the literature by implementing the tabu search algorithm. Comparing the results with exact solutions obtained by a commercial solver showed the efficiency of that algorithm.

Teypez et al. (2010) proposed a new formulation for service network design with fleet assignments that take the carrier’s profit into consideration. They presumed that any demand can be completely or partially rejected by the carrier. In addition, different types of vehicles are available to be operated during predetermined time windows. A three-phase heuristic decomposition method was developed to solve large-scale instances.

Ng and Lo (2016) proposed robust models for service network design under demand

uncertainties. The authors basically assumed that the mean and upper bound are the only known information on passenger demands. This information is further supplemented by a lower bound, and the variance of demand, as two separate cases. They compared and analyzed two models using a ferry transportation service data in Hong Kong.

An interested reader may refer to Crainic (2000) and Wieberneit (2008) for thorough reviews on this area. Our study and service network design problem come close in spirit to *network design* problem. SNDP derives an optimal transportation plan on a network of terminals where freight is to be delivered within a delivery deadline. However, our study mainly focuses on an optimal design of the network including locations of the hub facilities and allocation of the demand nodes to hubs subject to shipment scheduling decisions.

2.5 Hub Location and Scheduling

The literature on scheduling in hub location problems is scarce. Yaman et al. (2012) proposed release-time scheduling and hub location problems for next-day delivery. They investigated a single allocation p -hub median problem under release-time decisions for trucks from each demand node, with a predetermined service level. They presumed a given closing time at all centers for receiving packages from their customers. In addition, all trucks must leave a demand center at the same time. The authors aimed at improving service quality at minimum total routing cost. They extended their model by considering uniform and piecewise linear arrival patterns, release-time dependent cargo arrival patterns, and multiple service levels. Applying a set of valid inequalities, they solved their models using CPLEX for CAB and Turkish data instances in a reasonable computation time.

Karimi and Setak (2016) analyzed an incomplete hub location problem subject to service time constraints. They proposed an integer programming model to decide the locations of hub facilities, allocation of non-hub nodes to hub nodes, and departure times at the nodes. It was assumed that direct connections are allowed and there are no capacity

restrictions on hubs. The authors reduced the problem size by carrying out a set of preprocessing steps. They further solved their model using CPLEX on instances from the CAB and AP data sets, as well as a new data set called Iranian Roads Network (IRN).

To the best of the author's knowledge, the present research is the first attempt at modeling simultaneously the decisions on hub location with those on shipment scheduling. Economies of scale are achieved through consolidating freight at hubs by the scheduling of shipments between those hubs. A consolidated shipment results from the accumulation of freight that is held at hubs over one or more time periods. We take the capacity of vehicles into consideration, and decide on the number of dispatches on each inter-hub link during each time period, in addition to the locations of hubs.

We develop mathematical models for different versions of this hub location and shipment scheduling problem. These variations are described at the start of the next chapter.

Chapter 3

Problem Definition and Mathematical Formulations

In this chapter, we first incorporate the stepwise constant cost function into a classical hub location problem. Then, we extend the model by considering the impact of shipment scheduling decisions on hub configurations. We initially define the hub location and shipment scheduling problem and further propose three different mathematical formulations for it. The first model is applicable when holding costs at hub facilities are negligible, whereas the second incorporates holding costs at hub facilities. In the third model, we include the possibility of using different types of vehicles.

3.1 Hub Location Model with stepwise constant cost function

In this section, we use the multi-commodity flow-based formulation proposed by Ernst and Krishnamoorthy (1996). We modify the inter-hub cost expression by introducing our

proposed cost function (see Section 2.2) into the model. The fleet is considered to be homogeneous. The sets, parameters, and variables required for the formulations are defined as below.

Sets

- N Set of demand nodes.
 H Set of candidate hub nodes.
 T Set of time periods.

Parameters

- w_{ij} The flow originating from node $i \in N$ destined to $j \in N$, at the beginning of the planning horizon.
 η Capacity of a vehicle serving an inter-hub link.
 f_k The fixed setup cost associated with locating a hub facility at node $k \in H$.
 \bar{f} Fixed cost to dispatch a vehicle.
 \bar{c}_{kl} Variable transportation cost from hub $k \in H$ to hub $l \in H$ per vehicle including operational costs.
 c_{ij} The cost of sending one unit of flow from node $i \in N$ to $j \in N$ on an allocation connection.

To simplify the notation, let O_i denote the total flow originating at node $i \in N$ that is available in the beginning of the time horizon. Under the assumption of single allocation, $O_i = \sum_{j \in N} w_{ij}$. Likewise, let $D_i = \sum_{j \in N} w_{ji}$, where D_i denotes the total flow destined to node $i \in N$.

Variables

$$x_{ik} = \begin{cases} 1 & \text{If node } i \in N \text{ is assigned to hub } k \in H \\ 0 & \text{o.w.} \end{cases}$$

($x_{kk} = 1$ denotes that node $k \in H$ is a hub node.)

y_{kl}^i	The total amount of flow of node $i \in N$ (i.e., traffic originating from node $i \in N$) routed through the hub link from $k \in H$ to $l \in H$.
z_{kl}	The number of dispatches required from hub $k \in H$ to hub $l \in H$.

The hub location problem with the stepwise constant cost function (SHLPSCF) is formulated as follows:

$$\begin{aligned}
[SHLPSCF] \quad \text{Minimize} \quad & \sum_{k \in H} f_k x_{kk} + \sum_{i \in N} \sum_{k \in H} c_{ik} O_i x_{ik} + \sum_{k \in H} \sum_{l \in H} (\bar{f} + \bar{c}_{kl}) z_{kl} \\
& + \sum_{i \in N} \sum_{k \in H} c_{ki} D_i x_{ik}
\end{aligned} \tag{3.1}$$

Subject to

$$(3.7), (3.8), (3.13)$$

$$\sum_{l \in H} y_{kl}^i - \sum_{l \in H} y_{lk}^i = O_i x_{ik} - \sum_{j \in N} w_{ij} x_{jk} \quad \forall i \in N, \forall k \in H \tag{3.2}$$

$$\sum_{i \in N} y_{kl}^i \leq \eta z_{kl} \quad \forall l \neq k \in H, \tag{3.3}$$

$$y_{kl}^i \geq 0 \quad \forall i \in N, \forall k, l \in H \tag{3.4}$$

$$z_{kl} \in \mathbb{Z}^+ \quad \forall l \neq k \in H \tag{3.5}$$

The third term in objective function (3.1) accounts for inter-hub shipping cost. Equations (3.2) guarantee flow balance. Constraints (3.3) ensure sufficient capacity on hand to ship flows between hubs. Constraints (3.4) and (3.5) are non-negative and integer requirements on their corresponding decision variables.

3.2 Hub Location and Shipment Scheduling Problem

The problem is to design a parcel delivery network to serve a given set of demand nodes denoted by N . Several hub location studies assume that a number of potential hub locations are identified (see e.g., Yaman et al., 2007; Alumur et al., 2012; Peker et al., 2016). Likewise, we define H to indicate a set of candidate hubs. A hub network is to be designed by determining at which of the candidate locations a hub will be operated, and allocating the demand nodes to those hubs. Shipments are then to be scheduled within a given planning horizon on this hub network that is designed. The planning horizon, T , is finite and discretized into $|T|$ time periods, i.e., $t = 1, \dots, |T|$. The first time period corresponds to the time interval $[0,1)$, the second one to $[1,2)$, etc. It is assumed that the planning horizon is relatively short (e.g., two days considering a next-day delivery service, or several days), hence the location and allocation decisions do not change over time.

Each demand node can send flow to any other node through this network, which is referred to as the full-cross-traffic assumption. A quantity of freight, w_{ij} , originating from node $i \in N$ and destined to node $j \in N$ becomes available at its origin at the beginning of the time horizon. All nodes send or receive flow through a single hub node which is referred as single allocation. Freight is to be shipped from the originating demand node to its hub during the first time period $[0,1)$. Due to short distances between demand nodes and their respective hubs, freight is assumed, without loss of generality, to become available at the originating hubs at time $t = 1$. Shipment schedules are to be determined only within the inter-hub network. Initially, we assume that a set of homogeneous vehicles V operate on the inter-hub links, each with limited capacity η .

It is important to note that, in this thesis, the term *vehicle* denotes a particular entity for transportation. Between any two hubs in the network, freight is said to be carried by a vehicle. In our numerical examples below, inter-hub movements occur by plane. In this context, a “vehicle” thus denotes an aircraft.

The problem is to decide on the number of vehicle dispatches. Each vehicle traverses an inter-hub link from $k \in H$ to $l \in H$ in transit time τ_{kl} . Hence, flow entering an inter-hub link (k, l) at time $t \in T$, arrives at $l \in H$ at time $t + \tau_{kl}$. The literature on dynamic network flow assumes that the flow over time has time horizon T . Analogously, we presume that the latest time that flow can enter the inter-hub link (k, l) is $T - \tau_{kl}$.

Upon arrival at hubs, the freight gets unloaded, sorted, consolidated either immediately or during a later period, and loaded onto vehicles. If freight stays at a hub facility to be shipped in a later time period, then inventory holding costs may be incurred. For each time period, we need to determine the quantity of freight to be shipped from each hub and the number of vehicles to dispatch on each inter-hub link.

The problem is thus to design the hub network and schedule the shipments of a single parcel-delivery company, at minimum total cost. It is assumed that the company employs its own private fleet to serve the inter-hub links. There is a fixed cost to dispatch a vehicle, and a variable transportation cost per unit distance consisting of fuel, long-term, and regular maintenance costs. Given the distances between nodes, variable transportation costs can be computed for each vehicle trip. Introducing predefined fixed costs and vehicle variable costs as the sources for economies of scale, total cost incurred on inter-hub links is calculated by the number of vehicles serving the inter-hub links times the summation of vehicle fixed costs and variable costs. On the other hand, routing the flow on allocation links, i.e. links in which either origin or destination is non-hub, incurs a cost proportional to the amount of flow.

Retaining all other sets and parameters, we need to define a new set and parameter to formulate the problem as follows.

Sets

T Set of time periods.

Parameters

τ_{kl} Transit time between hub $k \in H$ and hub $l \in H$.

Note that all cost parameters should be scaled considering the length of the planning horizon. In the next two sections, we introduce different mathematical formulations of this problem.

3.2.1 Problem formulation with no holding costs

We initially assume that freight holding costs at hub facilities are negligible. Since the planning horizon is assumed to be short, it may be reasonable to assume that holding freight for an even shorter period of time (e.g., for a half day) at a hub facility incurs no cost. To this end, retaining x_{ik} , where $i \in N$ and $k \in H$, we define the following variables to formulate the problem.

y_{kl}^{it}	The total amount of flow of node $i \in N$ (i.e., traffic originating from node $i \in N$) entering the hub link from $k \in H$ to $l \in H$ at the beginning of the time period $t \in T$.
z_{kl}^t	The number of dispatches required from hub $k \in H$ to hub $l \in H$ in time period $t \in T$.
r_k^{it}	Total amount of flow originating at node $i \in N$, and remaining at hub node $k \in H$ by the end of the time period $t \in T$.

The shipment scheduling hub location problem with no holding costs is formulated as

$$\begin{aligned}
 [SSHLP] \quad \text{Minimize} \quad & \sum_{k \in H} f_k x_{kk} + \sum_{i \in N} \sum_{k \in H} c_{ik} O_i x_{ik} + \sum_{t \in T} \sum_{k \in H} \sum_{l \in H} (\bar{f} + \bar{c}_{kl}) z_{kl}^t \\
 & + \sum_{i \in N} \sum_{k \in H} c_{ki} D_i x_{ik}
 \end{aligned} \tag{3.6}$$

Subject to

$$\sum_{k \in H} x_{ik} = 1 \quad \forall i \in N \tag{3.7}$$

$$x_{ik} \leq x_{kk} \quad \forall i \in N, \forall k \in H \quad (3.8)$$

$$r_k^{i1} = \sum_{j \in N} w_{ij} x_{ik} - \sum_{\substack{l \in H \\ l \neq k}} y_{kl}^{i1} \quad \forall i \in N, \forall k \in H \quad (3.9)$$

$$r_k^{i,t-1} + \sum_{\substack{\{l: t-\tau_{lk} > 0\} \\ l \neq k}} y_{lk}^{i,t-\tau_{lk}} = r_k^{it} + \sum_{\substack{l \in H \\ l \neq k}} y_{kl}^{it} \quad \forall i \in N, \forall k \in H, \forall t = 2, \dots, T \quad (3.10)$$

$$r_k^{iT} = \sum_{j \in N} w_{ij} x_{jk} \quad \forall i \in N, \forall k \in H \quad (3.11)$$

$$\sum_{i \in N} y_{kl}^{it} \leq \eta z_{kl}^t \quad \forall l \neq k \in H, \forall t \in T \quad (3.12)$$

$$x_{ik} \in \{0, 1\} \quad \forall i \in N, \forall k \in H \quad (3.13)$$

$$z_{kl}^t \in \mathbb{Z}^+ \quad \forall l \neq k \in H, \forall t \in T \quad (3.14)$$

$$y_{kl}^{it} \geq 0 \quad \forall i \in N, \forall l, k \in H, \forall t \in T \quad (3.15)$$

$$r_k^{it} \geq 0 \quad \forall i \in N, \forall k \in H, \forall t \in T \quad (3.16)$$

The objective function (3.6) represents overall cost of the network consisting of setup costs for locating the hub facilities, the fixed and variable transportation costs. Constraints (3.7) indicate that each node either is a hub or is allocated to a hub node. Relation (3.8) ensures that each node is allocated to an *operating* hub facility. Equalities (3.9)-(3.11) refer to the flow conservation. Constraints (3.9) calculate the total amount of freight available to be shipped at a hub node by the end of the first time period. Equation (3.10) represents flow divergence, written for each period excluding the first. Constraints (3.11) assure that all demand becomes available at the destination's hub by the end of the planning horizon. Inequalities (3.12) calculate the total number of vehicles required on each inter-hub link. Relations (3.13)-(3.16) represent the binary, nonnegative integer, and real variables.

3.2.2 Problem formulation with holding costs

In the previous formulation, we incorporated the dimension of time into a hub location problem, which made possible the scheduling of shipments. However, this model might result in holding freight at hub facilities for some periods of the planning horizon. In a more general application setting, the cost to hold freight at hubs might not be negligible. Therefore, for a delivery company, keeping track of carrying costs at hubs may become important. As a matter of fact, a trade-off between holding costs and transportation costs determines the optimal design of hub networks. To address the stated issue, we propose another formulation for the problem.

Let h_k indicate the holding cost per unit at hub node $k \in H$; this remains constant over all periods. Retaining the previous assumptions and variables, we need to introduce the following additional decision variables to keep track of the amount of freight being held at each hub facility during each time period:

v_{kj}^{it} The demand originating at node $i \in N$ destined to node $j \in N$
in period $t \in T$, via hub $k \in H$ as the last hub on the route.

q_k^{it} The amount of freight originating from node $i \in N$ to be held
at hub $k \in H$ in time period $t \in T$.

We propose the shipment scheduling hub location problem with holding cost (SSHLPH) as follows.

$$\begin{aligned}
 [SSHLPH] \text{ Minimize } & \sum_{k \in H} f_k x_{kk} + \sum_{i \in N} \sum_{k \in H} c_{ik} O_i x_{ik} + \sum_{t \in T} \sum_{k \in H} \sum_{l \in H} (\bar{f} + \bar{c}_{kl}) z_{kl}^t \\
 & + \sum_{i \in N} \sum_{k \in H} c_{ki} D_i x_{ik} + \sum_{t \in T} \sum_{i \in N} \sum_{k \in H} h_k q_k^{it} \tag{3.17}
 \end{aligned}$$

Subject to

$$(3.7)-(3.16)$$

$$\sum_{t \in T} v_{kj}^{it} = w_{ij} x_{jk} \quad \forall i, j \in N, \forall k \in H \tag{3.18}$$

$$q_k^{it} = r_k^{it} - \sum_{j \in N} \sum_{t'=1}^t v_{kj}^{it'} \quad \forall i \in N, \forall k \in H, \forall t \in T \quad (3.19)$$

$$v_{kj}^{it} \geq 0 \quad \forall i, j \in N, \forall k \in H, \forall t \in T \quad (3.20)$$

$$q_k^{it} \geq 0 \quad \forall i \in N, \forall k \in H, \forall t \in T \quad (3.21)$$

The last term in the objective function (3.17) accounts for the holding costs. Constraints (3.18) ensure that demands between all pairs of nodes are met. Equality (3.19) calculates the amount of freight originating from an individual node that is held at each hub in each time period. Constraints (3.20) and (3.21) ensure the non-negativity of the quantities held.

In a more general setting, a heterogeneous set of vehicles, each with different capacities, might be available to be operated on the inter-hub links. To consider this situation, we can extend the SSHLPH model by modifying corresponding parameters and variables. Let M denote the set of vehicle types, with respective capacity $\eta_m, m \in M$, that are available to serve the inter-hub links. Let \bar{f}_m indicate the fixed cost to dispatch a vehicle of type $m \in M$. The cost \bar{c}_{klm} represents the variable transportation cost from hub $k \in H$ to hub $l \in H$ for a vehicle of type $m \in M$. Retaining all other decision variables, we need to define only one new variable z_{klm}^t as follows.

z_{klm}^t The number of dispatches of vehicle type $m \in M$ required
from hub $k \in H$ to hub $l \in H$ in time period $t \in T$.

The shipment scheduling hub location model with holding costs and different types of vehicles (SSHLPHV) can be formulated as:

$$\begin{aligned} [SSHLPHV] \text{ Minimize } & \sum_{k \in H} f_k x_{kk} + \sum_{i \in N} \sum_{k \in H} c_{ik} O_i x_{ik} + \sum_{m \in M} \sum_{t \in T} \sum_{k \in H} \sum_{l \in H} (\bar{f}_m + \bar{c}_{klm}) z_{klm}^t \\ & + \sum_{i \in N} \sum_{k \in H} c_{ki} D_i x_{ik} + \sum_{t \in T} \sum_{i \in N} \sum_{k \in H} h_k q_k^{it} \end{aligned} \quad (3.22)$$

Subject to

(3.7)-(3.11),(3.13),(3.15),(3.16),(3.18)-(3.21)

$$\sum_{i \in N} y_{kl}^{it} \leq \sum_{m \in M} \eta_m z_{klm}^t \quad \forall l \neq k \in H, \forall t \in T \quad (3.23)$$

$$z_{klm}^t \in \mathbb{Z}^+ \quad \forall l \neq k \in H, \forall t \in T, \forall m \in M \quad (3.24)$$

The third term in the objective function (3.22) calculates the total cost of shipments on inter-hub links, considering all types of vehicles. Constraints (3.23) compute the number of vehicles of each type required on each inter-hub link. Relations (3.24) assure integrality of the variables.

3.3 Size of the Models

In this section, we demonstrate the number of variables and constraints in each of the formulations. Let $|N| = n$, $|H| = h$, $|T| = \bar{t}$, and $|M| = \mu$. As can be seen from Table 3.1, for all four models, the number of binary variables is the same and equals $O(nh)$. The number of integer variables follows $O(h^2)$ for single allocation hub location with the stepwise constant cost function (SHLPSCF). The number of integer variables for the shipment scheduling hub location models with negligible holding costs (SSHLP) and non-negligible holding costs (SSHLPH) increases to $O(h^2\bar{t})$. When multiple types of vehicles are available to operate between hubs (SSHLPHV), the formulation involves $O(h^2\bar{t}\mu)$ integer variables.

The SHLPSCF formulation with $O(nh^2)$ has the least number of continuous variables. The SSHLP formulation has $O(nh^2\bar{t})$ continuous variables, while SSHLPH and SSHLPHV models have $O(n^2h\bar{t})$. The number of constraints in SHLPSCF formulation follows $O(n^2)$. The SSHLP formulation has $O(nh\bar{t})$ constraints, whereas the SSHLPH and SSHLPHV formulations contain $O(n^2h + nh\bar{t})$.

As can be observed, the SSHLPH formulation consists of more continuous variables

and constraints than the SSHLP. Although the SSHLPHV model has as many constraints, binary and continuous variables as the SSHLPH model does, the SSHLPHV model has more integer variables. To sum up, we expect that the SSHLP formulation is computationally less demanding than SSHLPH and SSHLPHV formulations. Furthermore, SHLPSCF is expected to be computationally far less demanding formulation than the other.

Table 3.1: Number of variables and constraints in the formulations.

	SHLPSCF	SSHLP	SSHLPH	SSHLPHV
Number of binary variables	nh	nh	nh	nh
Number of integer variables	h^2	$h^2\bar{t}$	$h^2\bar{t}$	$h^2\bar{t}\mu$
Number of continuous variables	nh^2	$nh^2\bar{t} + nh\bar{t}$	$n^2h\bar{t} + nh^2\bar{t} + 2nh\bar{t}$	$n^2h\bar{t} + nh^2\bar{t} + 2nh\bar{t}$
Number of constraints	$n^2 + 2nh + n - h$	$nh\bar{t} + h^2\bar{t} + 2nh - h\bar{t} + n$	$2nh\bar{t} + n^2h + h^2\bar{t} + 2nh - h\bar{t} + n$	$2nh\bar{t} + n^2h + h^2\bar{t} + 2nh - h\bar{t} + n$

Chapter 4

Computational Analysis

In this section, we perform extensive computational tests of the models *SHLPSCF*, *SSHLP*, *SSHLPH*, and *SSHLPHV*. We begin by describing our procedure for estimating and generating the problem data in Section 4.1. In Section 4.2, we address the results of the *SHLPSCF* model. We report the results of the *SSHLP* and *SSHLPH* models in Sections 4.3 and 4.4, respectively. In Section 4.5, the sensitivity of the models are further discussed, in relation to demand and length of the planning horizon. Results of the *SSHLPHV* model are presented in Section 4.6.

4.1 Parameter Estimation and Data Generation

We generated a new data set to be referred as the United States Air Flow (USAF) dataset, with 15 nodes (225 O-D pairs) and 20 nodes (400 O-D pairs), to test the shipment scheduling hub location models. This dataset is available in Masaeli (2017). Locations of demand nodes in this dataset correspond to the first 15 and 20 cities in the Civil Aeronautics Board (CAB) dataset used in O’Kelly (1987) (see Appendix, Table 1, for a list of those 20 cities, with corresponding node numbers). We extracted the values of demands (w_{ij}) from the

Freight Analysis Framework (2016) developed by the Center for Transportation Analysis. That framework provides values of the commodity flows shipped between domestic origins and destinations in the United States. The flow data with value and weight are reported for 34 different commodity types and seven modes of transportation. At the time of this study, the data available were for the year 2014. Among the seven transportation modes, we considered only shipments by air. For each pair of cities, we summed the flow values over all commodities to calculate the total demand between each pair of nodes in the USAF dataset.

We generated a set of candidate hub locations from among the set of demand nodes. We adopted one of the best-performing node-importance measures, reported in Peker et al. (2016) for the CAB data set, to pick promising locations for hubs. We used the measure $(O_i + D_i)C_i$ where $C_i = \sum_{j \in N} d_{ij}$. For a given node $i \in N$, the summation of incoming and outgoing flow represents the magnitude of demand at the node. C_i indicates the node's centrality measure, calculated by summation of the Euclidean distances between node $i \in N$ and all other nodes. For instance, by applying the node importance measure on the USAF dataset with 15 nodes, the first eight cities as potential hub locations in descending order of their importance are Los Angeles, Miami, Chicago, Memphis, Dallas, Atlanta, Detroit, and Cleveland. The list of eight candidate hub locations when there are 20 nodes remains almost the same, except that the last three cities are replaced by New York, Philadelphia, and Boston.

Fixed cost per dispatch (\bar{f}) is associated with purchasing/leasing costs of aircraft. We calculated this cost by assuming that the lifespan of a plane is 20 years and that it traverses four legs per day on average. Variable transportation cost per vehicle (\bar{c}_{kl}) corresponds to operating costs that include flying operations, maintenance, and fuel costs per plane. We estimated the values of the fixed costs and variable costs per flight according to the costs and airplane operating characteristics reported by Coyle et al. (2011) for two different types of cargo planes: B747-F and B747-100. We make an inflation-adjusted estimate of

the costs by using inflation factors reported by the Bureau of Statistics (2017). Table 4.1 shows values of the fixed cost per flight and the capacity of airplanes.

Table 4.1: The capacity of airplanes and corresponding fixed costs.

Type	Capacity, η (tons)	Fixed cost per flight, \bar{f} (\$)
B747-F	72.58	9247
B747-100	46.34	8082

We determine the value of transportation costs on allocation links (c_{ij}) based on average operational costs of trucking per mile, reported by American Transportation Research Institute (2015). The capacity of a truck serving allocation links is assumed to equal the average capacity of a large straight truck. That capacity is taken as 14000 lbs, as reported by FedEx (2017). Hence, the operational cost of trucking is estimated as \$0.24 per ton per mile. Table 4.2 compares the cost per ton of using a B747-F aircraft vs a truck, for a number of sample origin-destination pairs. From Table 4.2, we observe that after a certain distance, approximately 1900 miles, it is cheaper to operate a B747-F than to use a truck.

Table 4.2: Cost comparison between airplane and truck.

Origin-Destination	Distance (mile)	Airplane cost (\$/ton)	Truck cost, c_{ij} (\$/ton)
Atlanta-Baltimore	577	201	161
Atlanta-Denver	1208	371	337
Atlanta-Los Angeles	1937	518	541
Boston-Los Angeles	2600	652	726

To test different values for holding cost, we picked four product types that cover a wide range of values from among the 34 types provided by the Freight Analysis Framework

(2016). These are pharmaceutical, electronics, plastics or rubber, and furniture. Holding costs are often computed yearly and expressed as a percentage of the value of the inventory item. Assuming the percentage holding cost is $r = 0.25\$/\$/yr$, and with 365 days/year, we calculate the holding cost per ton per day for each category in Table 4.3. The resulting values represent four holding cost classes ranging from high to low.

Table 4.3: Holding cost classes.

Class	Range	Example products	Holding cost (h_k) (\$/ton/day)
A	high	pharmaceutical	286
B	medium-high	electronics	124
C	medium-low	plastics or rubber	19
D	low	furniture	9

To analyze the effect of the fixed setup cost of hub facilities on location, routing, and holding decisions, we vary hub setup costs such that $f_k \in \{5000, 10000, 25000\}$ for all $k \in H$. These costs are estimated upon the information in Civil Aviation Forums (2002) and Ferraiuolo (2014).

The travel time between cities is calculated by assuming an average cruising speed of 550 miles/h. Moreover, it is assumed that the sorting, loading, and unloading times are embedded in the traveling time. We initially consider a planning horizon of three time periods in our computational results. Each period corresponds to a half-day. We round the travel times up to the nearest whole number of periods. We further conduct some numerical experiments with four time periods (2 days), as well.

We solve all four models: those for single allocation hub location with the stepwise constant cost function (SHLPSCF), shipment scheduling hub location (SSHLP), shipment scheduling hub location with holding costs (SSHLPH), shipment scheduling hub location

with holding costs and heterogeneous vehicles (SSHLPHV). CPLEX version 12.6.3 is employed on a server with 16xAMD Opteron 2.61 GHz and 40 GB RAM. In order to obtain acceptable results within a reasonable amount of computational time, we impose a five-hour CPU time limit for the 15-city dataset.

For the 20-city dataset, we set the CPU time limit to eight hours and also set an optimality gap of 2.00% for the last three models. The CPLEX solver terminates, once either the time or gap condition holds, for these 20-city instances.

4.2 SHLPSCF Model Results

Results of the single allocation hub location model with the stepwise constant cost function for 15-node and 20-node networks are given in Tables 4.4 and 4.5, respectively.

Table 4.4: Results of SHLPSCF model on the 15-node network

Type	Setup cost	Obj. Value	Hub locations	No. of flights	CPU time (s)
B747-F	5000	540,386	4,12,13	11	36.35
	10000	551,821	4,13	7	8.55
	25000	567,393	4	-	0.93
B747-100	5000	532,970	4,12,13	16	19.01
	10000	547,970	4,12,13	16	18.43
	25000	567,393	4	-	0.22

The model tends to open a greater number of hub facilities in Table 4.5 than in Table 4.4. This is because more demand nodes in the network are to be served. Optimal locations of hubs vary with the type of aircraft and setup-cost values. For example, when the B747-100 is employed and setup costs equal \$5,000, the network is expanded to include 6 hubs at Chicago, Dallas, Los Angeles, Memphis, Miami, and New York, and operate 38 flights.

The last column of Tables 4.4 and 4.5 show that required CPU times are sensitive to

Table 4.5: Results of SHLPES model on the 20-node network

Type	Setup cost	Obj. Value	Hub locations	No. of flights	CPU time (s)
B747-F	5000	843,220	4,12,13,14,17	24	172.21
	10000	863,270	4,12,13,17	20	60.77
	25000	923,270	4,12,13,17	20	19.56
B747-100	5000	842,722	4,7,12,13,14,17	38	2848.55
	10000	870,175	4,12,13,14,17	35	2096.25
	25000	937,270	4,12,13,17	31	100.09

the type of aircraft and also to setup costs. All instances on 15-node network are solved to optimality within at most one minute. For 20-node network, B747-F instances are solved on average within 85 seconds. However, for B747-100 instances, the average CPU time is significantly higher. Among them, those under \$5,000 and \$10,000 setup costs are around 45 minutes as the model tends to open more hubs and use flights more frequently. By increasing the setup cost values to \$25,000, the CPU time drops to about 2 minutes as fewer hubs tend to be opened.

We further compare the results for SSHLPSCF model with those for SSHLP and SSHLPH models in Sections 4.3 and 4.4.

4.3 SSHLP Model Results

Tables 4.6 and 4.7 show results of the shipment scheduling hub location model for the USAF dataset with three time periods, and 15 or 20 cities, respectively. The first columns of these tables indicate the type of plane operating on the inter-hub links. For each type, we test three different values of setup costs for hub facilities so as to investigate the impact of these values on location and routing decisions, and the CPU times in each case. For each problem instance, Tables 4.6 and 4.7 report the respective objective function value,

the locations of hub facilities, the total number of flights operating on the network over the planning horizon, and the CPU time requirements in seconds.

Table 4.6: Results of SSHLP model on the 15-node network.

Type	Setup cost	Obj. Value	Hub locations	No. of flights	CPU time (s)
B747-F	5,000	540,386	4,12,13	11	768
	10,000	551,821	4,13	7	209
	25,000	567,393	4	-	3
B747-100	5,000	532,970	4,12,13	16	959
	10,000	547,970	4,12,13	16	362
	25,000	567,393	4	-	2

Table 4.7: Results of SSHLP model on the 20-node network.

Type	Setup cost	Obj. value	Hub locations	No. of flights	CPU time (s)
B747-F	5,000	843,270	4,12,13,17	20	3,670
	10,000	863,270	4,12,13,17	20	1,185
	25,000	923,270	4,12,13,17	20	623
B747-100	5,000	849,020	4,12,13,14,17	35	28,800*
	10,000	874,020	4,12,13,14,17	35	22,852
	25,000	937,270	4,12,13,17	31	4,372

* This instance failed to achieve the optimality gap within the time limit. Reported gap is 2.28%.

The resulting hub locations in Table 4.6 are affected by both setup costs and airplane types. In the case of operating B747-F on inter-hub links, the number of hub facilities gradually decreases from three to one with increasing setup costs. However, in the case of operating the B747-100, hub configurations remain the same for the first two setup-cost values. When the setup cost exceeds a certain threshold, the model locates only a single

hub. In this case, we no longer can take advantage of economies of scale through installing a hub network. Since airplanes operate only on inter-hub links, the hyphen (-) indicates that all flow is shipped by truck when the model locates just a single hub facility. It is also noteworthy to mention that, regardless of parameter values, Chicago is always picked as an optimal hub location.

As can be seen in Table 4.1, the B747-100 has less than two-thirds the capacity of a B747-F. However, in light of the fixed cost per flight and the variable costs, the latter is roughly twice as expensive to operate. Therefore, more flights are operated on the network when B747-100 type aircraft are employed (see Tables 4.6 and 4.7).

CPU times reported in these tables show that the computational difficulty of the problem is very sensitive to parameter values. B747-100 instances are computationally more demanding than the B747-F instances, as the model tends to operate more flights in the former case. In addition, for a given airplane type, we observe that setting the fixed cost parameters to 25,000 decreases the CPU time significantly, as fewer hubs tend to be opened. In Table 4.7, the B747-100 instance with a \$5,000 setup-cost value could not be solved within eight hours; the gap reported by CPLEX at the end of that running time was 2.28%. The resulting hub networks in Table 4.6 are quite identical to those in Table 4.4. In other words, disregarding the time dimension in the 15-city USAF dataset with negligible holding costs fails to affect the optimal hub locations and total number of dispatches during the planning horizon. This is not surprising, since *travel time* between any pair of Chicago, Los Angeles, and Memphis takes only a single time period while airplanes are operated. When it comes to the network with slower transport modes or farther distances between nodes, it is expected that taking the time dimension into account would result in a different hub network.

The number of opened hub facilities is greater in Table 4.7 than in Table 4.6. For a given value of setup costs, more cities must be served within the planning horizon for the 20-node dataset, whose total demand is almost twice that of the 15-node case. Therefore,

the model tends to open additional hub facilities to serve all demands within the same given planning horizon. Comparing locations in Tables 4.6 and 4.7 for a given aircraft type, we observe that the resulting hub networks in the 20-node case are less sensitive to the setup-cost variations. Four cities: Chicago, Los Angeles, Memphis, and New York are the most frequently chosen locations of hub facilities.

4.4 Results for SSHLPH Model

Next, we solve the model for shipment scheduling hub location with non-negligible holding cost (SSHLPH) for the 15-node and 20-node USAF datasets by varying the holding costs from low to high. We report results in Tables 4.8 and 4.9 (holding-cost values and their designation are as in Table 4.3).

As can be seen in Table 4.8, the optimal hub networks (optimal locations of hub facilities and numbers of flights) are influenced by all three parameters: type of airplane, setup cost, and holding cost. Therefore, it is difficult to draw general conclusions about the effect of each parameter on the resulting networks. For the B747-F instances, as setup costs increase to \$10,000, the model tends to open fewer hub facilities. As a result, an increased amount of freight is shipped by truck and consequently, the model operates fewer flights on the network. However, in the case of B747-100 instances, the same hub facilities are operated under both setup-cost values except for the last instance listed in Table 4.8 (only a single hub, Chicago, is operated in the case of high holding cost).

We also tested instances when setup costs equal \$25,000. In this case, with either aircraft type and under all holding-cost values, the model locates a single hub in Chicago, with the objective function value of \$567,393.

The resulting hub networks in Table 4.8 are quite different in some cases from those in Table 4.4. For the B747-F instances, while the optimal network under a \$5,000 setup-cost

Table 4.8: Results for SSHLPH model on the 15-node network.

Type	Setup cost	Holding cost	Obj. value	Hub locations	No. of flights	CPU time (s)
B747-F	5,000	low	540,524	4,12,13	11	6,908
		medium-low	540,678	4,12,13	11	5,705
		medium-high	541,821	4,13	7	1,635
		high	541,821	4,13	7	1,192
	10,000	low	551,821	4,13	7	4,621
		medium-low	551,821	4,13	7	2,029
		medium-high	551,821	4,13	7	570
		high	551,821	4,13	7	324
B747-100	5,000	low	533,272	4,12,13	16	4,556
		medium-low	533,608	4,12,13	16	11,434
		medium-high	537,131	4,12,13	16	10,860
		high	540,927	4,12,13	17	8,353
	10,000	low	548,272	4,12,13	16	4,620
		medium-low	548,608	4,12,13	16	1,782
		medium-high	552,131	4,12,13	16	2,510
		high	552,393	4	-	1,090

in Table 4.4 comprises three hub facilities, the number of hubs decreases to two under medium-high and high holding costs in Table 4.8. In the case of B747-100 instances under a \$10,000 setup cost, both the SHLPSCF and SSHLPH models tend to open three hubs and to operate 16 flights, except when holding freight incurs a high cost. In that case, the SSHLPH model tends to open a single hub at Chicago.

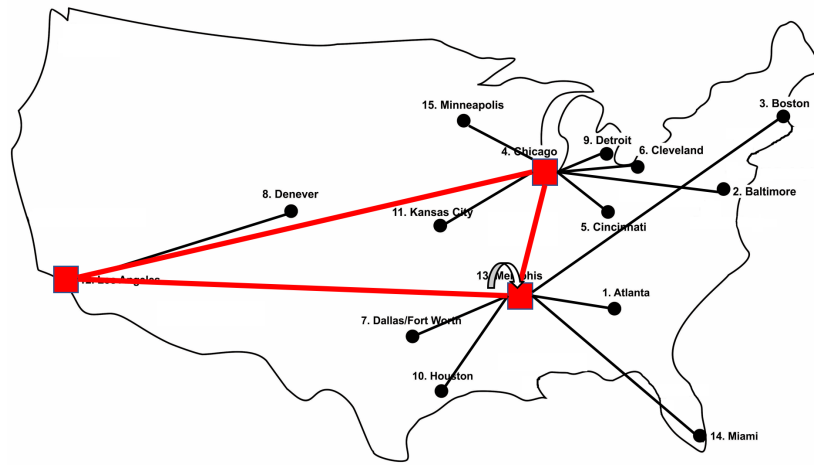
For some problem instances, we analyzed the amount of freight held at hubs. In the case of B747-F, when setup costs are \$5,000 and holding costs are either low or medium-low, the same quantity of freight is held in the SSHLPH model as in the SSHLP model: 30.76 tons at Memphis at the beginning of the second time period. All demands are met by the end of the third time period. As the holding costs increase to medium-high, the model ships all freight in the beginning of the second time period. Hence, fewer hub facilities are

needed. In addition, all demands are served by the end of the second time period. For B747-F instances when setup costs are \$10,000, the model opens two hubs, in Chicago and Memphis, under all holding costs, and does not hold freight at any hubs. For this reason, the objective function values reported in Table 4.8 for B747-F and \$10,000 are all the same.

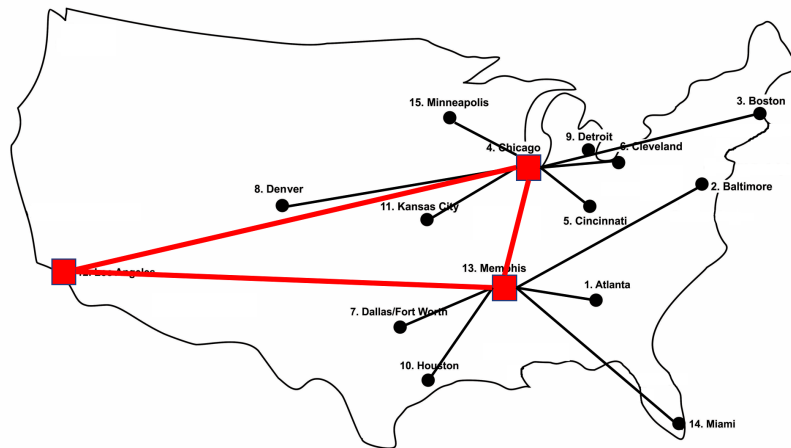
When the B747-100 are employed and the setup cost is \$5,000, the model ships freight through a network with three hubs located in Chicago, Los Angeles, and Memphis under all holding costs. However, optimal allocation decisions do change. Figure 4.1 illustrates location and allocation decisions under low and high holding costs. In this figure, a hub facility is shown by a solid square. Thin and thick lines depict allocation links and inter-hub links, respectively. A curved arrow presented on top of a hub indicates that freight was held there during a particular time period. As can be observed from this figure, the demand nodes: Baltimore, Boston, and Denver, are allocated to different hubs under low and high holding costs. Note also that no freight is held in the latter case.

The last column of Table 4.8 shows that required CPU times are sensitive to the type of airplane and also to the setup and holding costs. Increasing the holding costs to a high value results in reduction of CPU times, since the model tends to hold less inventory at hub facilities. For instances in which the SSHLPH model holds no inventory at hubs, the CPU time requirement is less than the required CPU time of instances that result in freight being held.

We calculate the average CPU times reported in Tables 4.6 and 4.8 for \$5,000 and \$10,000 setup costs. These averages are 384 and 4,291 seconds, respectively. Comparison of results for the two models reveals that the SSHLPH model is computationally more demanding than the SSHLP model. This is not surprising, as the SSHLPH model includes additional variables and constraints to keep track of freight held at hubs. We should point out that the optimal solution of the SSHLP model provides a lower bound on the optimal solution of SSHLPH as the former does not include any holding costs. In some cases, this lower bound may turn out to be an optimal solution to the SSHLPH model. For example,



(a) Low holding cost, second time period [1,2)



(b) High holding cost, second time period [1,2)

Figure 4.1: Location and allocation decisions on 15-node USAF dataset by operating B747-100 airplane under $f_k = 5,000$.

the objective function values in Tables 4.6 and 4.8 for B747-F instances with a \$10,000 setup cost value are exactly the same (\$551,821). This is because no freight is held at hubs in these instances. Note also that, in this case, both models yield the same hub locations. Hence, solutions of the SSHLP model can be used to give some intuition into the solutions

of the SSHLPH model.

Next, we solve the SSHLPH model for the 20-node network. Results in Table 4.9 show that hub structures are resilient to the variation in both setup and holding costs. For all B747-F instances, the model ships freight through a network with four hub facilities located in Chicago, Los Angeles, Memphis, and New York. Hub locations are almost the same for the case of operating the B747-100. For the latter, however, when setup costs are set to either \$5000 or \$10,000 and holding costs are low, the model establishes one more hub in Miami. This is not surprising, since operating B747-100 airplanes is less expensive than operating B747-F, hence the model has more incentive to use a greater number of hubs when operating costs are low enough. In that case, the number of required flights increases as a greater amount of freight is to be shipped through the inter-hub network.

For all B747-F instances except those under high holding costs, the model holds freight in Chicago, Memphis, and New York. In the case of high holding costs, the model tends to hold less freight at hubs: about half the preceding amount is then held at Chicago and Memphis, and none in New York. The main motivation to hold freight at hub facilities is to consolidate that flow with freight that will arrive in a later period. In the case of B747-F instances under high holding costs, the model tends to ship more freight in the second time period, since holding freight incurs a greater carrying cost. Therefore, the total number of required flights increases

The last column of Table 4.9 presents the gap reported by CPLEX at the end of eight hours. As can be observed, the SSHLPH model is very difficult to solve to optimality. However, given that we use deterministic values for all parameters, and considering that the actual data contains uncertainty, a gap of 2-3% is reasonable, in our opinion.

We demonstrate some of the solutions in Figure 4.2. We pick two instances from Table 4.9; B747-F instances with \$10,000 setup cost when holding costs are high (e.g., pharmaceutical products) and low (e.g., furniture). Figures 2a-d show the locations of

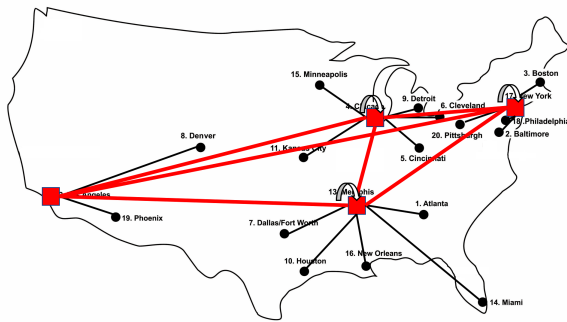
Table 4.9: The results of the SSHPH model on the 20-node network.

Type	Setup cost	Holding cost	Obj. Value	Hub locations	No. of flights	Gap (%)
B747-F	5,000	low	844,036	4,12,13,17	20	2.89
		medium-low	844,888	4,12,13,17	20	2.00
		medium-high	853,830	4,12,13,17	20	2.54
		high	861,853	4,12,13,17	21	2.00
	10,000	low	864,036	4,12,13,17	20	2.61
		medium-low	864,888	4,12,13,17	20	2.00
		medium-high	873,830	4,12,13,17	20	1.55
		high	881,853	4,12,13,17	21	2.00
	25,000	low	924,036	4,12,13,17	20	2.13
		medium-low	924,888	4,12,13,17	20	2.00
		medium-high	933,830	4,12,13,17	20	1.80
		high	941,853	4,12,13,17	21	2.00
B747-100	5,000	low	849,687	4,12,13,14,17	35	3.48
		medium-low	857,697	4,12,13,17	31	3.20
		medium-high	860,056	4,12,13,17	31	3.72
		high	863,696	4,12,13,17	31	3.58
	10,000	low	874,687	4,12,13,14,17	35	2.89
		medium-low	877,697	4,12,13,17	31	3.04
		medium-high	880,056	4,12,13,17	31	3.80
		high	883,696	4,12,13,17	31	3.60
	25,000	low	937,472	4,12,13,17	31	3.17
		medium-low	937,697	4,12,13,17	31	3.12
		medium-high	940,056	4,12,13,17	31	2.55
		high	943,696	4,12,13,17	31	2.00

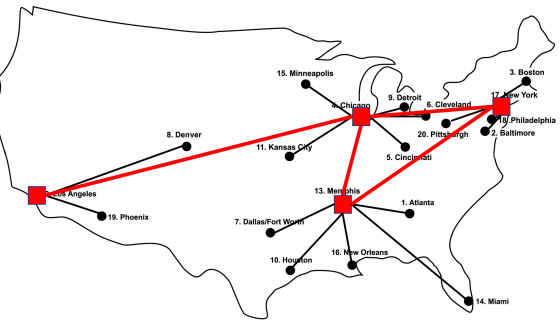
hub facilities and the holding decisions in the second and third time periods for these two instances.

During the first time period, $[0,1)$, freight is shipped from the supply nodes to hubs,

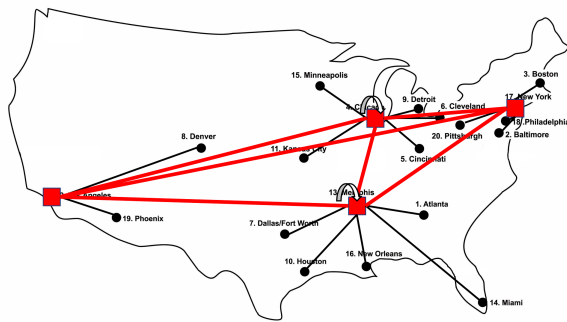
so that all freight is available at the corresponding hub facilities at the beginning of the second time period. In both instances depicted in Figure 4.2, there is a flight from each hub to every other hub during the second time period. The model establishes four hub facilities, and operates a total of 16 flights with low holding costs, and 19 flights when those costs are high. However, during the third time period, delivery services are not provided between some hubs. The number of flights decreases from 16 to 4, for low holding costs, and from 19 to 2 for high holding costs. No hub facility holds freight in the third time period (Figure 4.2-b,d).



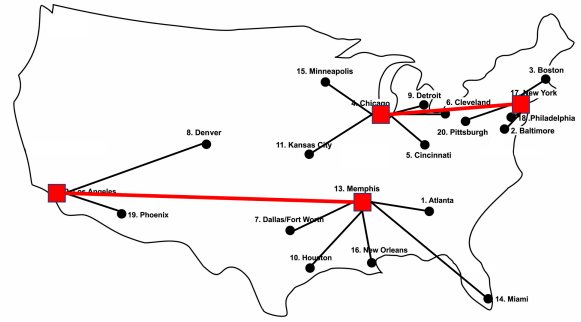
(a) Low holding cost, second time period [1,2)



(b) Low holding cost, third time period [2,3)



(c) High holding cost, second time period [1,2)



(d) High holding cost, third time period [2,3)

Figure 4.2: Resulting hub configurations with B747-F airplane and $f_k = 10000$.

Observe from Figure 4.2 that, whether holding cost is high or low, the locations of hubs are unchanged. However, that cost value does affect the holding decisions and consequently

the number of required flights. For the case of low holding cost, the model holds around 2.4% of total demand at three hub facilities: Chicago, Memphis, and New York (Figure 4.2-a) in the second time period. With high holding costs, only about 0.5% of the total demand is held in Chicago and Memphis, and no freight is held in New York (Figure 4.2-c). Thus, no flights are operated between New York and Memphis in the third time period. Compared with the low-holding-cost instance, more flights (19 vs 16) are required to ship freight in the second time period in the case of high holding cost. However, during the third time period, the number of required flights is decreased from 4 to 2.

Let us now consider some details of the routing in various periods, for the cases of both low and high holding costs, for the instances of Figure 4.2. According to the demand data, 0.22 tons of freight is to be shipped from Boston to Houston. When the holding cost is low, the entire demand is shipped to New York within the first time period. That freight arrives there at the beginning of the second time period, and is held in New York during this period. At the beginning of the third time period, freight is consolidated with the shipment from Chicago. During period 3, the consolidated freight is shipped on the same flight to Memphis. At the end of the third period, freight is shipped to Houston and demand is met. With high holding costs, on the other hand, the Boston-Houston demand is served by the end of the second time period. During period 1, freight arrives in New York. At the beginning of the second period, freight is shipped to Memphis and demand is met by the end of that period. As can be observed, the routing decisions and times of delivery for the demand from Boston to Houston are affected by the values of holding costs.

Another example is the demand to be shipped from Boston to Denver. Under both low and high holding costs, the entire demand is shipped from New York to Los Angeles at the beginning of the second time period, and demand is met by the end of period 3. So, the routes of this origin-destination pair remain the same under the two holding-cost scenarios.

4.5 Sensitivity Analysis of Demand and Length of Planning Horizon

To observe the effect of demand on the resulting hub configurations, we increased all demand values by 50%. Table 4.10 shows the optimal solutions obtained by the SSHPH model with increased demand on the 15-node USAF dataset with B747-F aircraft. Note that the model opens three hubs, in Chicago, New York, and Memphis, and operates 15 flights to ship freight for all instances listed in Table 4.10. As opposed to the B747-F instances presented in Table 4.8, the resulting hub network (hub locations and required number of flights) is quite robust against variations in setup and holding-cost values. Even when the setup costs increase to \$25,000, the model still takes advantage of economies of scale in designing a hub network. The required number of flights to ship freight through the network increases as the demands grow.

Table 4.10: Results of SSHPH model with B747-F when demands are increased by 50%.

Setup cost	Holding cost	Obj. Value	Hub locations	No. of flights	CPU time (s)
5,000	low	762,851	4,12,13	15	11,370
	medium low	763,160	4,12,13	15	13,642
	medium high	766,402	4,12,13	15	1,946
	high	771,406	4,12,13	15	529
10,000	low	777,851	4,12,13	15	2,145
	medium low	778,160	4,12,13	15	1,702
	medium high	781,402	4,12,13	15	444
	high	786,406	4,12,13	15	303
25,000	low	822,851	4,12,13	15	817
	medium low	823,160	4,12,13	15	358
	medium high	826,402	4,12,13	15	356
	high	831,406	4,12,13	15	197

Next, we investigate the sensitivity of the resulting networks and the CPU time to the length of the planning horizon. We solve both models, i.e., SSHLP and SSHLPH, on the 15-node USAF dataset over *four* time periods. Table 4.11 reports the results of operating B747-F where setup costs are set to \$5,000 and \$10,000. Observe that the resulting hub networks and the corresponding costs remain exactly the same as those presented in Tables 4.6 and 4.8, because all demand is served by the end of either the second or third time periods. So, although the models allow shipping freight over a longer time horizon with four time periods, no flow is shipped during the last period. Actually, the three-time-period solutions are all feasible for the problem with four time periods. In this case, solutions of the three-period problem are actually optimal for the 15-node USAF dataset with four time periods.

Table 4.11: Results of operating B747-F within a planning horizon of four time periods.

Model	Setup cost	Holding cost	Obj. Value	Hub locations	No. of flights	CPU time (s)
SSHLP	5,000	-	540,386	4,12,13	11	18,000*
	10,000	-	551,821	4,13	7	18,000*
SSHLPH	5,000	low	540,524	4,12,13	11	18,000*
		medium low	540,678	4,12,13	11	18,000*
		medium high	541,821	4,13	7	6,214
		high	541,821	4,13	7	2,830
	10,000	low	551,821	4,13	7	18,000*
		medium low	551,821	4,13	7	7,581
		medium high	551,821	4,13	7	773
high		551,821	4,13	7	301	

* The instance reached the time limit.

Increasing the number of time periods from three to four drastically raises the CPU times. The models are solved on average within 45 minutes when there are three time periods, whereas the average CPU time grows to three hours with four time periods. In half the instances listed in Table 4.11, optimal solutions could not be obtained within the

time limit. The average optimality gap for those instances is around 1.20%. Results for CPU time show that extending the planning horizon makes the instances computationally more demanding, especially for the case of low holding costs.

4.6 Results of SSHLPHV Model

We additionally tested the SSHLPHV model on the 15-node USAF dataset where *both* types of aircraft, B747-F and B747-100, can operate on the inter-hub links. Table 4.12 reports the objective function values, locations of hub facilities, the required number of flights for each type of plane, and required CPU times in seconds.

Table 4.12: Results of the SSHLPHV model with two aircraft types.

Setup cost	Holding Cost	Obj. Value	Hub locations	No. flights		CPU time(s)
				B747-F	B747-100	
5,000	low	513,532	4,12,13	4	9	18,000*
	medium low	513,751	4,12,13	4	9	18,000*
	medium high	514,367	4,12,13	4	9	5,743
	high	514,367	4,12,13	4	9	519
10,000	low	528,532	4,12,13	4	9	18,000*
	medium low	528,751	4,12,13	4	9	7,565
	medium high	529,367	4,12,13	4	9	1,029
	high	529,367	4,12,13	4	9	351

* The instance reached the time limit.

Observe from the hub locations given in Table 4.12 that Chicago, Los Angeles, and Memphis are always selected as hub facilities under all values of setup and holding cost. In every instance, the model ships freight by operating four B747-F flights and nine B747-100 flights on the inter-hub links. Note that if the B747-100 airplane flies at full capacity, operating that aircraft will be cheaper than the B747-F. Since full capacity is the case, the

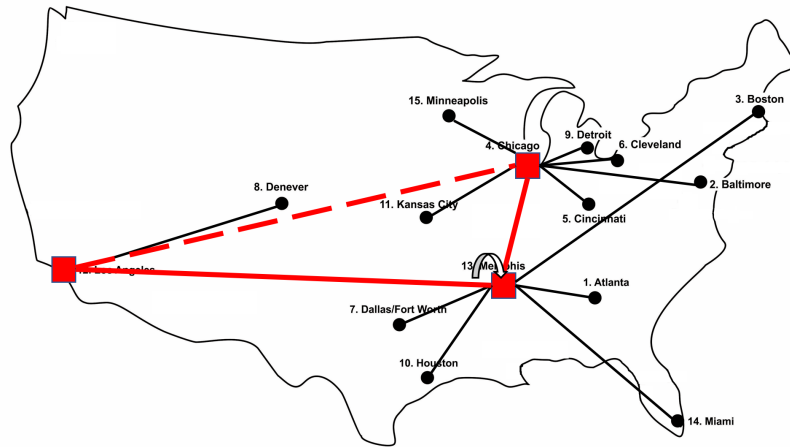
model tends to use B747-100 more frequently.

CPU time requirements in the last column show that, on average, the model is solved within 2.7 hours. Only three out of eight instances could not reach an optimal solution within the time limit. The average gap reported by CPLEX for those instances is around 0.53%.

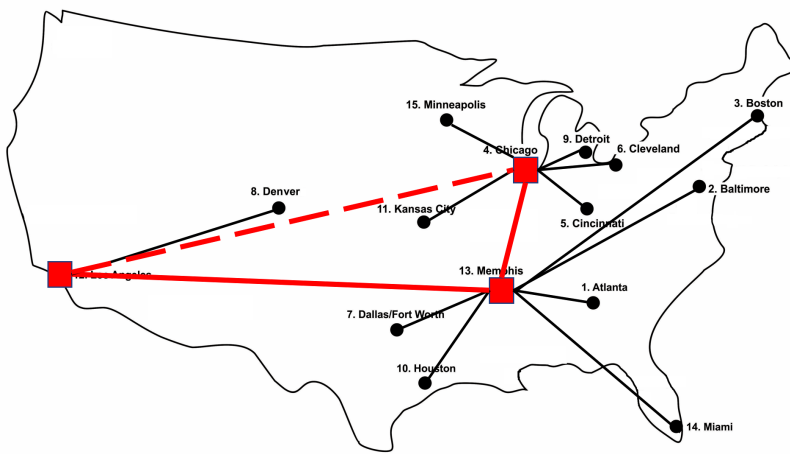
Let us compare the optimal hub locations in Tables 4.8 and 4.12. It can be inferred that the hub network which results from SSHLPHV model is most robust against variations in setup and holding costs. In addition, Chicago and Los Angeles are always picked as hubs. As observed from Tables 4.8 and 4.12, the locations of hub facilities are identical for some instances. However, it should be noted that the shipment *strategies* are different for such instances. In addition, comparison of the objective function values reveals that the SSHLPHV model is able to serve demands at a lower cost through employing two different types of planes.

Figure 4.3 illustrates the locations of hub facilities, allocation of demand nodes to hubs, and flights on inter-hub links with B747-F and B747-100 aircraft.

Figures 4.3-a and 4.3-b depict the resulting hub networks under low and high holding costs during the second time period [1,2). Solid thick lines indicate the inter-hub links served by both types of airplanes, whereas the dashed line implies operating only B747-100 on the link. The allocation decisions under both holding costs are the same except for Baltimore. When the holding costs are low, Baltimore is allocated to the hub facility at Chicago and freight is held at Memphis. However, under high holding costs, Baltimore is allocated to Memphis and freight is no longer held at any hub. From Table 4.12, it is interesting to note that the objective function value for the case of high holding costs is greater than that for low holding costs, even though no freight is held at hubs. This difference is due to the change in the shipment structure, as the same number of flights are operated under the two holding-cost cases.



(a) B747-F and B747-100, low holding costs



(b) B747-F and B747-100, high holding costs

Figure 4.3: Resulting hub network of 15-node USAF dataset under $f_k = 5,000$, second time period [1,2).

Chapter 5

Conclusion

In this thesis, we introduced hub location and shipment scheduling problem. Shipments are to be delivered within a given planning horizon on the hub network to be established. Freight might be held at hubs to be consolidated with other shipments at a later time period. Economies of scale are exploited by scheduling consolidated shipments.

We developed several mixed-integer programming models incorporating shipment scheduling decisions into hub location problems. The models aim to determine the optimal locations of hubs, design of the hub network, and the number of vehicles to operate on that network in each time period. We initially proposed a model where there is negligible cost to hold freight at hubs. We then extended this model to keep track of the holding decisions when holding costs are non-negligible. Moreover, we provided a third model allowing for a heterogeneous fleet of vehicles to operate between hubs.

We generated a new dataset, referred to as the USAF dataset, consisting of real data to test our models. As an initial effort toward solving shipment scheduling hub location models, we used CPLEX to conduct the computational experiments. All instances were solved optimally within reasonable computational times, or terminated with acceptable optimality gaps.

Results showed the computational difficulty of a problem to be very sensitive to the values of parameters. For example, higher fixed-cost values can reduce the CPU time significantly, as fewer hubs tend to be opened. The lower the holding costs, the greater the CPU time. In addition, instances with lower operational cost are computationally more demanding, as the models tend to operate more flights. CPU time is also increased when the planning horizon is extended. The numerical experiments also revealed that the SSHLPH model is computationally more demanding than the SSHLP model. This is not surprising, since the former model contains extra variables and constraints to keep track of the freight held at hubs.

We analyzed the effect of shipment scheduling decisions and holding costs on optimal hub network configurations, routing decisions, and total cost of the network. Optimal hub networks, including the locations of hub facilities and the number of dispatches, are influenced by the type of vehicle, setup cost, holding cost, and demand. When there are more cities to be served, a higher demand to be met within the planning horizon, or operational costs are low, the model has greater incentive to use a larger hub network. The resulting network is also quite robust against variations in setup and holding costs.

Results of the SSHLPHV model showed that hub networks operating a heterogeneous fleet are more resilient against changes in setup and holding costs. We also tested both the SSHLP and SSHLPH models over a longer planning horizon. Interestingly, our results showed that solutions for the shorter horizon might be feasible, and even optimal, for the longer-horizon problem.

Future research may concentrate on developing exact solution methodologies or heuristic algorithms to solve large-scale instances. In addition, one might develop the model by taking asset management considerations into account. For example, vehicle-flow conservation constraints might be added to the set of existing constraints. Another extension is to formulate the problem under a continuous planning horizon. Further investigation can focus on different transportation modes operating on allocation and inter-hub links along

with their corresponding travel times. The model also can be expanded by considering direct connections or multiple allocation assumptions.

Less-than-truckload (LTL) carriers check for freight and modify improper classes and weights to reduce shipping costs. One interesting possible future direction is formulating economies of scale by combining threshold-based discounting factor that is proposed by Podnar et al. (2002) with fixed costs that are proportional to distance, and taking *bumping clause* phenomenon into account.

References

- Alumur, S., Kara, B. Y., 2008. Network hub location problems: The state of the art. *European Journal of Operational Research* 190 (1), 1–21.
- Alumur, S. A., Kara, B. Y., Karasan, O. E., 2012. Multimodal hub location and hub network design. *Omega* 40 (6), 927–939.
- Alumur, S. A., Nickel, S., Saldanha-da Gama, F., Seçerdin, Y., 2016. Multi-period hub network design problems with modular capacities. *Annals of Operations Research* 246 (1-2), 289–312.
- American Transportation Research Institute, 2015. An analysis of the operational costs of trucking. <http://atri-online.org/wp-content/uploads/2015/09/ATRI-Operational-Costs-of-Trucking-2015-FINAL-09-2015.pdf>.
- Andersen, J., Crainic, T. G., Christiansen, M., 2009. Service network design with management and coordination of multiple fleets. *European Journal of Operational Research* 193 (2), 377–389.
- Armacost, A. P., Barnhart, C., Ware, K. A., 2002. Composite variable formulations for express shipment service network design. *Transportation Science* 36 (1), 1–20.
- Beasley, J. E., 1990. OR-library: distributing test problems by electronic mail. *Journal of the Operational Research Society*, 1069–1072.

- Bryan, D., 1998. Extensions to the hub location problem: Formulations and numerical examples. *Geographical Analysis* 30 (4), 315–330.
- Bureau of Statistics, 2017. The United States inflation calculators. <https://www.statbureau.org/en/united-states/inflation-calculators?dateBack=2007-12-1&dateTo=2015-12-1&amount=1000>.
- Campbell, J. F., 1990. Locating transportation terminals to serve an expanding demand. *Transportation Research Part B: Methodological* 24 (3), 173–192.
- Campbell, J. F., 1994. A survey of network hub location. *Studies in Locational Analysis* 6 (6), 31–49.
- Campbell, J. F., 1996. Hub location and the p-hub median problem. *Operations Research* 44 (6), 923–935.
- Campbell, J. F., Ernst, A. T., Krishnamoorthy, M., 2002. Hub location problems. *Facility location: applications and theory* 1, 373–407.
- Campbell, J. F., O’Kelly, M. E., 2012. Twenty-five years of hub location research. *Transportation Science* 46 (2), 153–169.
- Chen, W., He, K., Fang, X., 2014. Optimization of hybrid hub-and-spoke network operation for less-than-truckload freight transportation considering incremental quantity discount. *Mathematical Problems in Engineering* 2014.
- Civil Aviation Forums, 2002. Fee At Airports. <http://www.airliners.net/forum/viewtopic.php?t=425719>.
- Contreras, I., 2015. Hub location problems. In: *Location science*. Springer, pp. 311–344.
- Contreras, I., Cordeau, J.-F., Laporte, G., 2011. The dynamic uncapacitated hub location problem. *Transportation Science* 45 (1), 18–32.

- Coyle, J. J., Novack, R., Gibson, B., Bardi, E., 2011. *Transportation: A supply chain perspective*. 7. Auflage. Mason, OH: South-Western Cengage Learning.
- Crainic, T. G., 2000. Service network design in freight transportation. *European Journal of Operational Research* 122 (2), 272–288.
- Crainic, T. G., Rousseau, J.-M., 1986. Multicommodity, multimode freight transportation: A general modeling and algorithmic framework for the service network design problem. *Transportation Research Part B: Methodological* 20 (3), 225–242.
- Cunha, C. B., Silva, M. R., 2007. A genetic algorithm for the problem of configuring a hub-and-spoke network for a ltl trucking company in brazil. *European Journal of Operational Research* 179 (3), 747–758.
- DallOrto, L. C., Crainic, T. G., Leal, J. E., Powell, W. B., 2006. The single-node dynamic service scheduling and dispatching problem. *European journal of operational research* 170 (1), 1–23.
- Ebery, J., 2001. Solving large single allocation p-hub problems with two or three hubs. *European Journal of Operational Research* 128 (2), 447–458.
- Ernst, A. T., Krishnamoorthy, M., 1996. Efficient algorithms for the uncapacitated single allocation p-hub median problem. *Location Science* 4 (3), 139–154.
- Ernst, A. T., Krishnamoorthy, M., 1999. Solution algorithms for the capacitated single allocation hub location problem. *Annals of Operations Research* 86, 141–159.
- Farahani, R. Z., Hekmatfar, M., Arabani, A. B., Nikbakhsh, E., 2013. Hub location problems: A review of models, classification, solution techniques, and applications. *Computers & Industrial Engineering* 64 (4), 1096–1109.
- FedEx, 2017. Owner operators. <http://customcritical.fedex.com/us/owneroperator/quals/lgstrtrk.shtml>.

- Ferraiuolo, J., 2014. 10 most expensive airports. <http://www.therichest.com/luxury/most-expensive/10-most-expensive-airports-to-land-your-plane/>.
- Freight Analysis Framework, 2016. Domestic Flows. <http://faf.ornl.gov/fafweb/Extraction2.aspx>.
- Gelareh, S., Monemi, R. N., Nickel, S., 2015. Multi-period hub location problems in transportation. *Transportation Research Part E: Logistics and Transportation Review* 75, 67–94.
- Horner, M. W., O’Kelly, M. E., 2001. Embedding economies of scale concepts for hub network design. *Journal of Transport Geography* 9 (4), 255–265.
- Karimi, 2016. IRN dataset. <http://wp.kntu.ac.ir/hkarimi/files/IRN.rar>.
- Karimi, H., Setak, M., 2016. Flow shipment scheduling in an incomplete hub location-routing network design problem. *Computational and Applied Mathematics*, 1–33.
- Kim, D., Pardalos, P. M., 2000. A dynamic domain contraction algorithm for nonconvex piecewise linear network flow problems*. *Journal of Global Optimization* 17 (1-4), 225–234.
- Kimms, A., 2006. Economies of scale in hub & spoke network design models: we have it all wrong. In: *Perspectives on operations research: essays in honor of Klaus Neumann*. Gabler Publishing, Wiesbaden. Springer, pp. 293–317.
- Lai, M., Lo, H. K., 2004. Ferry service network design: optimal fleet size, routing, and scheduling. *Transportation Research Part A: Policy and Practice* 38 (4), 305–328.
- Marín, A., Cánovas, L., Landete, M., 2006. New formulations for the uncapacitated multiple allocation hub location problem. *European Journal of Operational Research* 172 (1), 274–292.

- Masaeli, M., 2017. USAF dataset. Mendeley Data, v1. <http://dx.doi.org/10.17632/z6xs9c5xbr.1>.
- Ng, M., Lo, H. K., 2016. Robust models for transportation service network design. *Transportation Research Part B: Methodological* 94, 378–386.
- O’Kelly, M. E., 1986a. Activity levels at hub facilities in interacting networks. *Geographical Analysis* 18 (4), 343–356.
- O’Kelly, M. E., 1986b. The location of interacting hub facilities. *Transportation Science* 20 (2), 92–106.
- O’Kelly, M. E., 1987. A quadratic integer program for the location of interacting hub facilities. *European Journal of Operational Research* 32 (3), 393–404.
- O’Kelly, M. E., 1992. Hub facility location with fixed costs. *Papers in Regional Science* 71 (3), 293–306.
- O’Kelly, M. E., Bryan, D., 1998. Hub location with flow economies of scale. *Transportation Research Part B: Methodological* 32 (8), 605–616.
- O’Kelly, M. E., Miller, H. J., 1994. The hub network design problem: a review and synthesis. *Journal of Transport Geography* 2 (1), 31–40.
- Pedersen, M. B., Crainic, T. G., Madsen, O. B., 2009. Models and tabu search metaheuristics for service network design with asset-balance requirements. *Transportation Science* 43 (2), 158–177.
- Peker, M., Kara, B. Y., Campbell, J. F., Alumur, S. A., 2016. Spatial analysis of single allocation hub location problems. *Networks and Spatial Economics* 16 (4), 1075–1101.

- Podnar, H., Skorin-Kapov, J., Skorin-Kapov, D., 2002. Network cost minimization using threshold-based discounting. *European Journal of Operational Research* 137 (2), 371–386.
- Pooley, J., Stenger, A. J., 1992. Modeling and evaluating shipment consolidation in a logistics system. *Journal of Business Logistics* 13 (2), 153.
- Skorin-Kapov, D., Skorin-Kapov, J., O’Kelly, M., 1996. Tight linear programming relaxations of uncapacitated p-hub median problems. *European Journal of Operational Research* 94 (3), 582–593.
- Sung, C. S., Song, S. H., 2003. Integrated service network design for a cross-docking supply chain network. *Journal of the Operational Research Society* 54 (12), 1283–1295.
- Tan, P. Z., Kara, B. Y., 2007. A hub covering model for cargo delivery systems. *Networks* 49 (1), 28–39.
- Teypez, N., Schrenk, S., Cung, V.-D., 2010. A decomposition scheme for large-scale service network design with asset management. *Transportation Research Part E: Logistics and Transportation Review* 46 (1), 156–170.
- Wang, D. Z., Lo, H. K., 2008. Multi-fleet ferry service network design with passenger preferences for differential services. *Transportation Research Part B: Methodological* 42 (9), 798–822.
- Wieberneit, N., 2008. Service network design for freight transportation: a review. *OR spectrum* 30 (1), 77–112.
- Yaman, H., Kara, B. Y., Tansel, B. Ç., 2007. The latest arrival hub location problem for cargo delivery systems with stopovers. *Transportation Research Part B: Methodological* 41 (8), 906–919.

Yaman, H., Karasan, O. E., Kara, B. Y., 2012. Release time scheduling and hub location for next-day delivery. *Operations research* 60 (4), 906–917.

APPENDIX

List of 20 Cities in the CAB Dataset

Table 1: 20 cities with corresponding node number

No.	City	No.	City
1	Atlanta	11	Kansas City
2	Baltimore	12	Los Angeles
3	Boston	13	Memphis
4	Chicago	14	Miami
5	Cincinnati	15	Minneapolis
6	Cleveland	16	New Orleans
7	Dallas- Fort Worth	17	New York
8	Denver	18	Philadelphia
9	Detroit	19	Phoenix
10	Houston	20	Pittsburgh