NEW OPTIMIZATION MODELS FOR THE HUB COVERING LOCATION PROBLEM
(ANA DAĞITIM ÜSSÜ KAPSAMA PROBLEMİ İÇİN YENİ ENİYİLEME MODELLERİ)

by

NAZMİ ŞENER, M.S.

Thesis

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

in

INDUSTRIAL ENGINEERING

in the

GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

of

GALATASARAY UNIVERSITY

AUGUST 2020

Approval of the thesis:

NEW OPTIMIZATION MODELS FOR THE HUB COVERING LOCATION PROBLEM

submitted by **NAZMİ ŞENER** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Industrial Engineering Department, Galatasaray University** by,

Examining Committee Members:

Prof. Dr. Orhan FEYZİOĞLU (Supervisor)
Department of Industrial Engineering,
Galatasaray University

Assoc. Prof. Dr. Sinan APAK
Department of Industrial Engineering,
Maltepe University

Assist. Prof. Dr. Orhan İlker BAŞARAN
Department of Industrial Engineering,
Galatasaray University

Prof. Dr. Gülçin BÜYÜKÖZKAN FEYZİOĞLU
Department of Industrial Engineering,
Galatasaray University

Assist. Prof. Dr. Fethullah GÖÇER
Department of Industrial Engineering,
Kahramanmaraş Sütçü İmam University

Date:

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude to my thesis advisor and mentor, Prof. Dr. Orhan FEYZİOĞLU, for his invaluable lead, motivation, support, patience, and time. Without his encouragement and supports, I would not have been able to complete this research. Without him, I would not have made it through my dissertation! It is a real privilege to be a student of him. I wish we may continue to strengthen the collaboration for a long time after this study's conclusion.

Special thanks to my ex-committee member, Prof. Dr. Necati ARAS, for his priceless suggestions and encouragement, whose insight and knowledge into the subject matter steered me through this research. During the last years of my study, I am also thankful to Assoc. Prof. Dr. Sinan APAK for taking part as a member of my thesis evaluation committee and his valuable comments and constructive suggestions. I take this opportunity to express my sincere gratitude to my co-workers, Aylin KIZIK, Emine KELEŞ, Nazik TOPRAK, and Ali ATAN. They endure my anxiety and give me moral support during this endeavor. I owe special thanks to Asst. Prof. Dr. Fethullah GÖÇER, my best friend and my Wonderwall during the thesis, supported me and had to put up with my stresses for the past four years of the study. Furthermore, I am also grateful to Mahir ATMIŞ, my friend, and my helper, for coding.

To conclude, my biggest thanks to my family for all the support you have shown me through this research, my father Taner ŞENER, my mother, and my first teacher Ümmühan ŞENER, my sister Nazlı ŞENER and my little brother Enes ŞENER. Without their support, encouragement, and love of my family, this work could not ever be possible.

Nazmi ŞENER

August 2020

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Hub location problems are derived from classical origin-to-destination problems. The difference of hub location problems is the existence of hub nodes that merge/split transportation demands. A hub also acts as a node where demand originates or arrives. The presence of hub nodes simplifies the network structure. Because of the economy of scale, the hub networks' benefit is saving transportation costs between hubs. This kind of problem is commonly encountered in the design of postal distribution, airline passenger transport, and telecommunication networks. In the current literature, hub location problems are classified into hub median, hub center, and hub covering. The least studied of these is the hub covering problems due to the network's lack of full connectivity. For this type of problem, the elements of capacity and transportation costs are generally ignored in studies. For filling this gap in the literature, new deterministic multiple assignment hub covering models have been developed. These proposed models were examined with numerical experiments using benchmark data sets. It was revealed how important it is to consider the elements mentioned above in the hub covering problem. Another issue that has not been studied for hub location problems in general and hub coverage problems, in particular, is the effect of uncertainties that may be encountered. In this thesis, new two-stage stochastic mixed-integer linear optimization models have been developed to examine the effect of uncertainties on transportation costs and transportation demand. The L-shaped algorithm has been adapted to solve these models, and numerical experiments have been performed with the help of benchmark data sets. As a result of these experiments' analysis, it was shown that the uncertainties should be taken into account for the hub covering problem.

Keywords : Hub-and-Spoke Networks, Hub Covering, Stochastic Programming, Uncertainty Modeling, Capacity Constraints

ÖZET

Ana Dağıtım Üssü (ADÜ) yerleşim problemleri klasik kaynaktan hedefe taşıma problemlerinden türetilmiştir. ADÜ yerleşim problemlerinin farklılığı taşıma taleplerini birleştiren/ayrıştıran ADÜ düğümlerinin varlığıdır. ADÜ'ler ayrıca diğerleri gibi talebin oluştuğu veya vardığı düğümlerdir. Birleştirme düğümlerinin varlığı ağ yapısını basitleştirmektedir. ADÜ'ler bulunan ağlarının faydası ölçek ekonomisi sayesinde ADÜ'ler arası taşıma maliyetlerinde elde edilecek tasarruftur. Bu problem türüyle genel olarak posta dağıtım, havayolu yolcu taşımacılığı ve telekomünikasyon ağlarının tasarımında karşılaşılır. Güncel yazında ADÜ yerleşim probleminin üç türü vardır: ADÜ medyan, ADÜ merkez ve ADÜ kapsama. Ağ yapısının tam bağlantılı olmamasından dolayı bunlar arasında en az çalışılmış olan tür ADÜ kapsama problemidir. Bu problem türü için kapasite ile taşıma maliyetleri unsurları çalışmalarda genellikle göz ardı edilmiştir. Yazındaki bu boşluğu doldurmak amacıyla yeni determinist çoklu atamalı ADÜ kapsama modelleri geliştirilmiştir. Önerilen bu modeller, kıyaslama veri kümeleri kullanılarak yapılan sayısal deneylerle incelenmiş, anılan unsurların ADÜ kapsama probleminde dikkate alınmasının ne derece önemli olduğu ortaya konmuştur. Genelde ADÜ problemleri için az, özelde ise ADÜ kapsama problemleri için hiç çalışılmamış bir diğer konu da karşılaşılabilecek belirsizliklerin etkisidir. Bu tez çalışmasında, taşıma maliyetleri ile taşıma talebindeki belirsizliklerin etkisini incelemek amacıyla yeni iki aşamalı rassal karışık tam sayılı doğrusal eniyileme modelleri geliştirilmiştir. Bu modellerin çözümü için L-şekilli algoritma uyarlanmış ve kıyaslama veri kümleri yardımı ile sayısal deneyler yapılmıştır. Bu deneylerin analizi sonucunda ADÜ kapsama problemi için belirsizliklerin göz önüne alınmasının gerekli olduğu gösterilmiştir.

Anahtar Kelimeler : Toplama ve Dağıtım Ağları, Ana Dağıtım Üssü Kapsama, Rassal Eniyileme, Belirsizlik Modelleme, Kapasite Kısıtları

1 INTRODUCTION

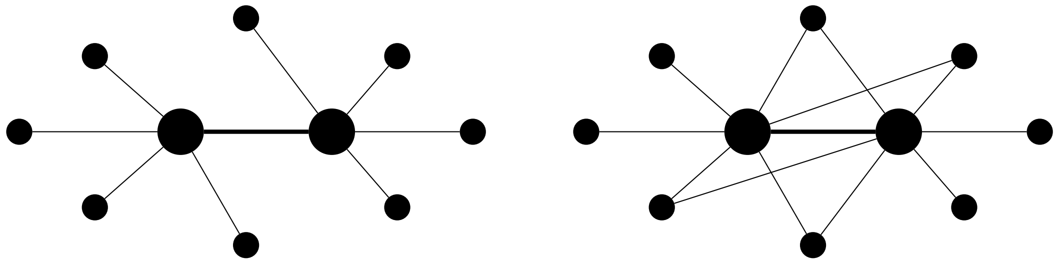
Creating a network structure is a complicated task as determining links, nodes, depots, or hubs always includes several layers. For example, choosing depots or hubs and assigning flows to these points are two dependent layers. Different studies investigate different aspects of this problem type due to complications of this problem type. In this study, the hub location problem is at the focus.

Hub location and vehicle routing problems may seem like the same problem. However, these problems differ from the concept of hubs or depots. Depots are not serving as a demand point. The only duty in the system of it is collecting/distributing goods to the system. Unlike depots, hubs are a combination of a depot and a demand point. This difference causes using these problem types in different application areas.

The hub location problem has several subtypes. According to the size of potential hubs, the problem is divided into two types : (1) a predetermined subset of nodes and (2) all nodes. The first type of problem decreases the size and complexity of the problem. On the other hand, this type ignores other nodes, which can be better hubs than potential subset nodes. Solving the second one may consume more time. However, the problem takes in to account all of the nodes.

According to the number of hubs in which a node can be attached, hub problems can be divided into two categories : *single assignment* and *multiple assignment*. For the single assignment hub location problems, each node must be assigned to only a single hub. On the other hand, a node can be assigned to one or more hubs in multiple assignment problems. A graphical illustration of this case is shown in Figure 1.1. It is obvious to deduct that the hub problem (whether multiple or single allocation) has fewer links comparing other network problems.

The hub location problem has 3 distinct types : hub median, hub center, and hub covering. Hub median problem fundamentally aims to minimize the total flow cost required to deliver service the given n origin/destination nodes, and the certain the



(a) single allocation hub network (b) multiple allocation hub network

Figure 1.1: The difference between single and multiple allocation hub networks

number of hubs (p) to settle. The hub center problem is a minimax version of the hub median problem. In a classical covering problem, the demand of nodes must be covered by a service point. In the hub covering problem, these service points are hubs.

The demand flow has three layers in this problem type. The first layer is delivering demands to hubs (*deliver phase*), the second layer (if necessary) hub to hub transportation (*interhub phase*), the third layer is allocating demand to destination points (*allocation phase*). The discount factors (per unit) of the deliver, interhub and allocation are χ, α and δ respectively. Because of the scale economy, the interhub discount factor must be the least of three factors ($\alpha < \chi$ and $\alpha < \delta$) (O'Kelly and Bryan, 1998). These discount factors are also a benefit of hub problems. In this thesis, χ and δ are assumed to be equal to 1. Because of that, these parameters are omitted from models.

The thesis includes five chapters. In this first chapter, basic notions are provided and hub location problems are briefly defined and classified. A comprehensive literature survey about the hub location problem is summarized and analyzed in detail in the second chapter. A new formulation for the multiple allocation hub covering the flow problem is introduced in Chapter 3. To incorporate some realistic issues, three more models are also developed for investigating link and hub capacity. The validity of the models is shown with computational studies based on two different data sets. Three different two-stage mixed-integer linear optimization model for multiple allocation hub covering flow problem is introduced in Chapter 4. These models are developed to investigate demand and/or transportation cost uncertainties. Different scenarios are created to examine the consequences of uncertainty. It is assumed that the flow demand is Poisson distributed and the transportation cost is normally distributed. An extensive numerical study is given in this chapter. In the last chapter, whole studies in the thesis are summarized and future research possibilities are discussed.

2 LITERATURE REVIEW

Hub location problems, which are crucial for the location science, are investigated over 30 years. Examples of these problems occur in various application areas such as postal networks, freight transportation via land, air or rail, airline passenger transportation, telecommunication, etc.

Campbell and O'Kelly (2012) classify the function of hubs into two groups : (1) a connecting, switching, or sorting function that authorizes flows to be altered at a hub node and (2) a consolidation/breakbulk function that authorizes flows to be amassed and dismissed. The first function is fundamentally focused on arranging flow directions to access the destination point in the shortest possible time. The number of links in the path between origin and destination pairs is decreased remarkably during this occurrence. On the other hand, the second function provides cost reduction for consolidated flows via economy of scale.

Lium et al. (2009) points out that the hub network structure in transportation systems and telecommunication systems are similar from different aspects such as service cost reduction, increasing service density, and appearance of demand uncertainty. Transportation hub problems do not concern setting up new links. Because road infrastructures are public (There is no opportunity to construct new roads or design a new route by air or sea to build new links without public approval and/or government permissions), this issue causes to shift the major focus to the cost of flow via these links. From this aspect, constructing new links is also expensive and difficult to set up in telecommunication hub problems (For example, set up a new fiber optic line can be extremely expensive). Because of this, telecommunication networks are also focused on unit flow costs.

The first articles which address the hub location problem are (O'Kelly, 1986a,b, 1987). However, Hakimi (1964, 1965) are seminal papers for hub location problem without addressing its name. In these papers, origin-destination flows are described, and node optimality for single or p medium location problems, which are the source of hub median location problems' motivation, are proven. A single facility location problem inspires Hakimi (1964) 's study. Because of this, origin-destination flow pairs are transformed to equivalent nodal demands. Similarly, O'Kelly (1986b) transforms the continuous planar single hub location problem to the classical Weber problem. Hakimi (1965) extends a single center to multiple switching centers. In this paper, the network movement is possible between two hubs, the same as in a hub network. This assumption is eliminated via neglected costs between switching centers. After this elimination, Hakimi (1965) focuses on the p median problem.

O'Kelly (1986b) develops a heuristic algorithm that is simply assigning nodes to the closest hub. In this study, two hubs' location in a plane is introduced in the new model with different assumptions about transportation economies pf scale effects on network structure. Also, American Civil Aeronautical Board (CAB) data set is introduced, which is then widely used as a benchmark in the literature. In this data set, the passenger traffic information between 25 US cities is given. Another benchmark data set, Australia Post (AP) data set, is first used by Ernst and Krishnamoorthy (1996). This data set is based on an Australian City's postal delivery service with 200 nodes.

The hub location problem is generally divided into three subcategories according to the objective function type in literature. Consequent sections are created based on this division.

2.1 The Hub Median Problem

Hub location problem is first formulated as an integer quadratic assignment model (O'Kelly, 1987). This formulation is known as the first example of a *single allocation p-hub median problem*. In this model, the objective function minimizes the total flow cost, measured in time, distance, etc., to carry goods or packages from origin to destination. There is n demand creating nodes to connect via p hubs.

Let the variable x_{ik} be 1 if node i is assigned to hub k, and 0 otherwise ; x_{kk} equals 1 when hub k is opened and 0 otherwise. W_{ij} symbolizes the flow amount sourced from node i with a destination node j. C_{ij} represents the unit flow cost between node i and node j. Also, α is an inter-hub cost discount factor, which is due to the economy of scale of inter-hub transfers. Single allocation p hub median problem is given in Eqs.(2.1a)-(2.1e). The objective function in Eq.(2.1a) includes the sum of all flow costs that occurred in the network. Due to the economy of scale, α must be in $[0, 1)$ range. The first constraint in Eq.(2.1b) ensures that if a hub is not opened, the node can be assigned to it. Constraints in Eq.(2.1c) and Eq.(2.1e) ensure single allocation, which means that each node is allowed to be assigned to only one opened hub. Constraint in Eq.(2.1d) enforces to open exactly p hubs.

$$\min \sum_{i,j} W_{ij} \left\{ \sum_k x_{ik} C_{ik} + \sum_m x_{jm} C_{jm} + \alpha \sum_{k,m} x_{ik} x_{jm} C_{km} \right\} \tag{2.1a}$$

$$\text{s.t. } (n+p-1)x_{kk} - \sum_i x_{ik} \geq 0 \qquad \forall k, \tag{2.1b}$$

$$\sum_k x_{ik} = 1 \qquad \forall i, \tag{2.1c}$$

$$\sum_k x_{kk} = p, \tag{2.1d}$$

$$x_{ik} \in \{0, 1\} \qquad \forall i, k. \tag{2.1e}$$

O'Kelly (1987) develops two heuristics to solve this problem with 2,3 and 4 hubs in a subproblem of the CAB data set, including ten nodes. Besides the development of the heuristics, the scale economy is also examined at different levels, and the article points out its effect on the solution procedure.

In subsequent work, O'Kelly (1992) adds the fixed hub opening cost to the hub median problem. Adding this cost eliminates the need for prior setting the number of hubs. As a result, the number of hubs becomes a decision variable. Besides constraint in Eq.(2.1b) is transformed into a simpler form as in Eq.(2.2b) below. Let, f_j be the fixed cost of opening a hub at node j, O_i the total amount of flow originating from node i ($O_i = \sum_j W_{ij}$), D_i the total amount of flow destination node i ($D_i = \sum_i W_{ij}$), and other parameters are same as previous. The fixed cost hub median problem due to O'Kelly (1992) is given below.

$$\min \sum_{i,k} x_{ik} C_{ik} \Big\{ O_i + D_i \Big\} + \sum_{j,m} x_{jm} \sum_{i,k} x_{ik} \Big\{ \alpha W_{ij} C_{km} \Big\} + \sum_{j} f_j x_{jj} \tag{2.2a}$$

$$\text{s.t. } x_{kk} - \sum_{i} x_{ik} \geq 0 \qquad \forall k, \tag{2.2b}$$

$$\text{(2.1c) and (2.1e).} \tag{2.2c}$$

The multiple allocation p-hub median problem is formulated as an integer linear programming initially by Campbell (1992). In that paper, discrete and continuous demands are compared, and their effects on the solution are investigated in detail. As the first integer linear programming model of the hub median problem, the model is given below. Define x_{ijkm} as a percentage of flow sourcing from node i to destination node j via hubs located at node k and m respectively. Let $c_{ijkm} = c_{ik} + c_{mj} + \alpha c_{km}$.

$$\min \sum_{i,j,k,m} w_{ij} c_{ijkm} x_{ijkm} \tag{2.3a}$$

$$\text{s.t. } \sum_{k,m} x_{ijkm} = 1 \qquad \forall i, j, \tag{2.3b}$$

$$x_{ijkm} \leq x_{kk} \qquad \forall i, j, k, m, \tag{2.3c}$$

$$x_{ijkm} \leq x_{mm} \qquad \forall i, j, k, m, \tag{2.3d}$$

$$x_{ijkm} \geq 0 \qquad \forall i, j, k, m, \tag{2.3e}$$

$$\text{(2.1d) and (2.1e).} \tag{2.3f}$$

Klincewicz (1992) develops a heuristic by combining tabu search and a greedy randomized adaptive search procedure to solve the single allocation p-hub median problem. According to the paper, this metaheuristic gives the optimal solution in 90% of the test problems. Skorin-Kapov and Skorin-Kapov (1994) develop an LP-relaxation algorithm combined with tabu search for solving p interacting hub in n nodes. They compared their method with the heuristics developed by O'Kelly (1987) and the metaheuristic due to Klincewicz (1992). The acquired solutions are far better than prior heuristics. On the other hand, CPU times for solving the instances are enormously large due to the allocation phase's burden.

The first integer linear programming model for a single allocation hub median problem is introduced by Campbell (1994). In this formulation, there are $(n^4 + n^2 + n)$ variables including (n^2+n) and (n^4+2n^2+n+1) linear constraints. Campbell (1994) also uses minimum flow amount as flow thresholds for allowing to use a link. If the thresholds are fixed the maximum values, the problem is transformed into a hub median problem with a single allocation setting. Another insight derived from the Campbell (1994) is if the threshold is omitted from links, the decision variables for routing(x_{ijkm}) turn binary because of directing origin-destination pair via least-cost paths. Because of this, there is no restriction needed to turn the decision variables integer.

Klincewicz (1996) uses the dual of the Campbell (1994) formulation. The dual ascending method, which has two parts : attempting to increase summation of dual decision variables constantly, and a dual adjustment procedure for balancing dual decision variables values, is used to solve subproblems in his branch-and-bound algorithm. Comparing the exact solution of branch-and-bound in that paper, solution times are decreased for all instances significantly. Contrary to this decrease, there is a gap (ranging from 0% to 2.2%) between the exact and inexact branch-and-bound solutions.

Ernst and Krishnamoorthy (1996) introduce a new approach for modeling hub median problems with a decreasing number of constraints $(2n^2 + n + 1)$ and variables $(n^3 + n^2$,n^2 of them are integers) which are useful for solving large scale problems. They model inter-hub transportation as a multicommodity flow problem. Let χ be the discount factor for transferring from non-hub to hub nodes, and δ be the discount factor for transferring from hub to non-hub. Define y_{ikl} as the total amount of flow originating from node i routed from hubs k and l, respectively. Using previous decision variables and parameters, they formulate the problem as in Eqs.(2.4a)-(2.4e). The constraint in Eq.(2.4b) ensures that the summation of the input flow amount for a node and flow amount sourced from that node is equal to the summation of the output flow amount that node and the amount of flow destined to that node.

$$\min \sum_{i,k} c_{ik}x_{ik}\Big\{\chi O_i + \delta D_i\Big\} + \sum_{i,k,l} \alpha c_{kl} y_{ikl} \tag{2.4a}$$

$$\text{s.t.} \sum_{l} y_{ikl} - \sum_{l} y_{ilk} = O_i x_{ik} - \sum_{j} w_{ij} x_{jk} \qquad \forall i,k, \tag{2.4b}$$

$$x_{ij} \leq x_{jj} \qquad \forall i,j, \tag{2.4c}$$

$$y_{ikl} \geq 0 \qquad \forall i,k,l, \tag{2.4d}$$

$$\text{(2.1c), (2.1d) and (2.1e).} \tag{2.4e}$$

O'Kelly et al. (1996) primarily focuses on the effects of the inter-hub discount factor on the exact solutions of the hub median problem with both single and multiple allocations. They apply the objective function linearization to solve the problem. They find out that the inter-hub discount factor's different values oblige the decision variables to behave as integer decision variables. Because of that, the hardness of the model increases or decreases related to the inter-hub discount factor.

Skorin-Kapov et al. (1996) find out that values of integer decision variables of the Campbell (1994) formulation's LP relaxation are immensely fractional. Because of this, they reformulate the p-hub median problem with a single allocation setting. This formulation includes $(2n^3+n^2+n+1)$ linear constraints with (n^4+n^2) variables including (n^2) binary variables. Linear relaxation of this formulation is almost close to its integral solutions when the CAB data set is used. They also employed an implicit enumeration search tree to get the optimal solution. From their derived information, the gap between this formulation's relaxation and optimal objective function value is less than 1%. They also solve multiple allocation setting of the hub median problem. This paper is first to attempt to acquire an optimal solution, and they also compare this solution with four different heuristics.

Sohn and Park (1997) introduces the first linear programming model to determine the location of two hubs in the hub median problem with a single allocation setting. They indicate that if the location of hubs is settled, the problem is solved in polynomial times. They transform the problem from the binary quadratic problem, which is introduced by O'Kelly (1987). After transformation, the problem becomes similar to the minimum cut problem. This property enables us to develop efficient solving methods in polynomial times.

Bryan (1998) presents four extensions to the literature : (1) a capacitated model, (2) a minimum threshold model, (3) a model that endogenously determine the number of hubs, and (4) a model that incorporates a flow-dependent cost function for the inter-hub links. The first and second extensions result in disconnected sub-networks, not the desired outcome for hub problems. The third extension transforms the p-hub median problem model like fixed cost addition to the model, which means that number of open hubs is determined to satisfy network demands. The paper's numerical results indicate that an increase in flow costs can cause an increase in total network cost while the inter-hub costs decrease.

Ernst and Krishnamoorthy (1999) introduce two new mathematical models for the capacitated single allocation hub median location problem. Their formulations extend those of Ernst and Krishnamoorthy (1996) and Campbell (1994) in two ways : adding hub opening cost as fixed cost and capacity constraints for each hub. These capacity constraints take only flow into account directly coming from non-hub nodes. They solve these problems with two heuristics. The first one is built on the simulated annealing, and the other heuristic is built on random descent. The optimal solutions are derived via an exact branch-and-bound method with upper bounds obtained by these heuristics. According to the results, the formulation based on Ernst and Krishnamoorthy (1996) 's work is superior to other formulations in both solution times and quality. They also improve their solution times and the gap between optimal and heuristic solutions with their preprocessing approach.

Marianov et al. (1999) propose a mathematical model for the multiple allocation hub median problem in the competitive environment. The model determines the hubs' location to maximize demand flow capture when different service providers are available in the network. The capturing procedure works as if the flow costs of a link of the provider is less than other competitors in the network; the provider captures all flows in that link. On the other hand, if; these are equal, the competitors' share flows equally. They solve the model with a tabu search based heuristic. For measuring the performance of the solution procedure, they use well known AP data set. Comparing the application of solution methodology to the non-competitive version of their model, the competitive model's procedure gives more quality results.

Sasaki et al. (1999) extend the multiple allocation hub median problem with only 1-stop in the route. They solve the problem with the branch-and-bound algorithm and a greedy-type heuristic. They also find out that branch and bound methodology is better when the number of nodes in the system is relatively small. Greedy type heuristic generally approaches the optimal solution when the specified integer parameter necessary for the algorithm is equal to the number of nodes. Also, the heuristic solves the problem quickly when the predetermined parameter equals to 1.

Ebery et al. (2000) extend the capacitated version of the multiple allocation hub median problem, which is introduced by Campbell (1994). To solve this problem, they develop an efficient heuristic based on the shortest path algorithm. They also set up an upper bound derived from a hybrid algorithm that is a linear programming based on a branch-and-bound algorithm. Their new model is solved in less amount of time compared to previous studies. Let Γ_k and f_k capacity and opening cost of hub k, respectively. Also let's define $W_{ij},\ z_{ik},\ y_{ikl},\ x_{ilj}$ as the total flow between node i to j, the flow amount which is transferred from node i to hub k, the flow amount which is transferred from node i through hub k and hub l and the amount of flow originated from node i with destination node j through hub l respectively. Let h_k equals 1 if the hub k is opened and 0 otherwise. Ebery et al. (2000)'s model is given in Eqs.(2.5a)-(2.5f).

$$\min \sum_{i,k,l} \left(\chi c_{ik} + \alpha c_{kl}\right) y_{ikl} + \sum_{i,k,l} \delta c_{lj} x_{ilj} + \sum_{k} f_k h_k \tag{2.5a}$$

$$\text{s.t.} \sum_{l} x_{ilj} = W_{ij} \qquad \forall i, j, \tag{2.5b}$$

$$\sum_{i,l} y_{ikl} \leq \Gamma_k h_k \qquad \forall k, \tag{2.5c}$$

$$\sum_{i,k} y_{ikl} \leq \sum_{i,j} W_{i,j} h_l \qquad \forall l, \tag{2.5d}$$

$$\sum_{k} y_{ikl} = \sum_{j} x_{ilj} \qquad \forall i, l, \tag{2.5e}$$

$$x_{ilj},\ y_{ikl} \geq 0, h_k \in \{0, 1\} \qquad \forall i, k, l, j. \tag{2.5f}$$

Sohn and Park (2000) suggest three formulations for the single allocation hub median problem. They prove that when the number of hubs equals three, the problem becomes an $\mathcal{NP}$-hard problem. They formulate the problem with a mixed-integer programming model. For reaching a feasible solution in considerable time, the polyhedral properties of the model are used. Because of using a fixed hub location, they also treat hub opening costs as a constant. They also find out that this formulation can be adapted to solve the three-terminal cut problem and the three-processor distribution problem.

Ebery (2001) develop an efficient algorithm for solving the large scale single allocation hub median problem with two and three hubs. In this study, the p hub allocation problem is also examined. They propose new mixed-integer programming models with fewer constraints and variables from prior studies. According to the best of authors' knowledge, they solve the largest single allocation problem. The numerical results of the study are superior to the results in the literature. Howbeit, the proposed formulations are effective only with a maximum of three hubs.

Mayer and Wagner (2002) introduce a solution procedure for the uncapacitated multiple allocation hub median problem. This procedure is essentially based on the branch-and-bound algorithm combined with tightened lower bound obtained through an aggregated model formulation and a simple heuristic determined by the upper bounds satisfying the complementary slackness of some constraints. Their procedure solves the problem optimally at most 40 nodes systems.

Boland et al. (2004) propose formulations for the uncapacitated and capacitated multiple allocation hub median problems. For improving the solution quality, preprocessing procedures obtained from computational results, and tightening constraints for prior mixed-integer linear programming formulations are presented. They also use flow covering constraints to improve solution times of the capacitated version of the problem. They prove that their proposed approach decreases computational times significantly to obtain optimal solutions.

Labbé and Yaman (2004) introduce multicommodity flows on the uncapacitated single allocation hub median problem. They add flow variables to the model via two different ways : direct projection and indirect projection. They investigate facets of the model's polyhedron and the linearization of the model for both single commodity and multicommodity cases. They propose reflecting flows to the model by introducing two new facet-defining inequalities. These facets dominate some projection inequalities. This development ensures that the facet inequalities can be detached in polynomial time.

Elhedhli and Hu (2005) combine the traffic congestion concept, which means that the model includes delays and different route preferences due to the enormous amount of traffic with the uncapacitated single allocation hub median problem. This combination adds some nonlinear terms in the objective function. For handling this nonlinearity, a linearization scheme is developed. They use a Lagrangian heuristic to solve the linearized model, which also includes integer decision variables. This heuristic is capable of finding high-quality solutions within acceptable computational times. Their results contribute to the literature's new and realistic insights about hub location problems.

Labbé et al. (2005) use polyhedral properties of the capacitated single allocation hub median problem with fixed hub opening costs, and use a branch and bound technique to solve the problem. In the model, capacity constraints limit the amount of flow transferred via hubs. They investigate the effects of different types of cuts on the problem. Their algorithm gives better results in a high amount of transferred flows and hubs' tight capacities comparing the low flow amounts and loose capacities.

Marín (2005a) deal with hub capacities on the multiple allocation hub median problem. They set up an upper bound to total flow that comes directly from origin. They introduce tight integer programming formulation for solving this problem. They also find out some useful properties of the optimal solution which can be applied to accelerate the computational time. Their experiments show that their model can solve mid-sized problems effectively, and it performs better than the existing literature.

Marín (2005b) investigate the multiple allocation hub median problem based on valid inequalities. They introduce an integer programming model that gives the best quality results compared to the literature available at that time. They use the fact that flow costs between the hubs satisfy triangular inequality. The formulation of the problem can be extended employing powerful valid inequalities acquired through the study of the intersection graph of a combined set packing problem. They use these inequalities in the Lagrangian context to reduce the size of the branching tree. This reduction decreases computational times remarkably.

Topcuoglu et al. (2005) employ a genetic algorithm to solve the uncapacitated single allocation hub median problem. They use both CAB and AP data sets. Their numerical results indicate that the study is significantly superior to prior studies in both solution quality and computational times to solve small, middle, and large-scale problems. Especially for the CAB data set, they found the optimal solution in less running time from existent studies remarkably. In the AP data set, they find eight times better solution than existent solutions.

Cánovas et al. (2007) focus on the uncapacitated hub median problem with multiple allocation. They use dual problem to solve the primal problem. For solving four indexes dual problem, they use a heuristic based on a dual ascending method. This heuristic includes several subroutines and an external linear programming solver embedded into the branch and bound algorithm. The heuristic gives better lower bounds to the branching tree. They use a maximum of 120 nodes of data sets. The numerical experiments indicate that their results are significantly surpassing prior results.

Chen (2007) suggests determining the upper bound for the number of hubs in the single allocation hub median problem with a hybrid heuristic hybridization of tabu search and simulated annealing heuristics. Numerical studies signify that implementing the acquired upper bound for the number of hubs, the proposed heuristic is adequate for acquiring optimal solutions for all small-scaled problems very easily. The results show that the hybrid heuristic is better than a tabu search heuristic and simulated annealing for solving the problem.

Cunha and Silva (2007) develop a hybrid algorithm based on a genetic algorithm for solving the uncapacitated hub median problem with a single allocation for trucking companies that operate less-than-truckload services in Brazil. The model consists of calculating the number of hubs (terminals) and determining the location of terminals. Their model alters from similar models found in the literature in the sense that it authorizes variable scale-reduction factors for the flow costs according to the total amount of flow between terminals. They brace their theoretical study in a real-world application with a known transportation company in Brazil.

da Graça Costa et al. (2008) address the capacitated hub median problem with single allocation. They propose a second objective function to the model, instead of using capacity constraints to restrict the amount of flow that can be received by the hubs. This second objective tries to minimize the time consumed from origins to reach hubs. Two bi-criteria models are developed. In the first model, total time is treated as the second criteria. In the second one, the maximum processing time for the hubs is minimized. These bi-criteria problems give a non-dominated set of solutions. They also compare their findings with classic single objective models.

Rodríguez-Martín and Salazar-González (2008) focus on a problem that includes establishing paths and locating hubs to be used to transport a group of commodities from origins to destinations in a given capacitated hub-spoke network. For examining this network, a mixed-integer programming model is used. The model is solved via a branch-and-cut algorithm combined with two different decomposition techniques. They find out that Double Benders' Decomposition surpasses Benders' Decomposition. However, these approaches are effective on 25 nodes and ten hubs candidates network. Heuristic methods are used for larger problems, and computational experiences indicate that these heuristics derive near-optimal solutions rapidly.

Yaman (2008) pays attention to the hub median problem with single allocation, and this network also has a different structure. Hubs are selected from a given set of nodes, and each hub in the structure is attached to a central hub in this structure. Links are established on the arcs of the arising network to transmit the flow. The study's objective is to determine the hub locations and the links to minimize the link setup cost. For realizing this aim, two different mathematical models and a metaheuristic is developed. The metaheuristic is the combination of Lagrangian Relaxation and local search methods. The solution quality of the metaheuristic is tested, and the formulations are compared in computational experiments.

Silva and Cunha (2009) suggest three different heuristics to solve the uncapacitated hub median problem with single allocation : three different versions of multi-start tabu search heuristic as a part of a two-stage integrated tabu search heuristic. Multi-start heuristics find several different initial solutions, and then these solutions are improved by tabu search. In the two-stage integrated heuristic, tabu search is implemented to enhance both parts (location and allocation). They also use widely known CAB and AP data sets to benchmark their findings. Computational experiments indicate that their methodology is capable of solving optimally 100 nodes networks. The solution times are considerably decreased comparing to existent literature.

Yaman (2009) proposes a special star network design for hub median problems with single allocation, which includes three layers. There is a complete network relating to the so-called central hubs in the first layer of the design. In the second layer, compounds of star networks connecting the remaining hubs to central hubs have occurred. In the third layer, the demand centers to hubs and central hubs are unified with the star networks. In this problem, the main aim is to allocate a predetermined number of hubs and central hubs with minimum flow costs resulting from the network structure. Service quality constraints originating from the delivery times are also added to the model.

Correia et al. (2010) extend the classical capacitated hub median problem by adding the size of hubs to the model. Each candidate hub has a set of capabilities which is selected via available capacity. Different models are suggested for the problem, analyzed in terms of the bound determined by linear programming relaxation. They also search for different inequalities to improve the model. For reducing the size of the problem, various preprocessing tests are introduced.

Ilić et al. (2010) introduce a new general variable neighborhood search heuristic to solve the uncapacitated hub median problem with single allocation. The problem deals with placement hubs to minimize the flow between all origin-destination pairs. Three neighborhoods are used, and data structures for calculating new total flow in the network are updated competently. A new strategy is offered to design a variable neighborhood descent local search in addition to the classical sequential strategy. Their computational experiments indicate that their results give better results than prior studies. This methodology is superior in large examples.

Ishfaq and Sox (2011) has a different perspective on hub median problem with multiple allocation, which includes flow costs covering different transportation modes, modal connectivity costs, and fixed hub cost contains service time requirements. They solve large scale problems, which include 100 nodes with tabu search metaheuristic. The solution acquired from the tabu search is compared with another heuristic, which setup lower bound via Lagrangian Relaxation. In their computational experiments, the intermodal hub problem's performance and the interaction of various factors on the design of the network with service time requirements are measured.

Kratica et al. (2011) deal with modified multicommodity capacitated hub median problem with multiple allocation. This modified model includes fewer variables and constraints than previous studies. Two evolutionary algorithms that adopt binary encoding and standard genetic operators are used to solve this modified problem. The second algorithm also uses a recently conceived heuristic to approximate the objective function value in large scale instances of the problem. The first one is capable of finding optimal solutions in small and middle scale problems. Comparing the existing literature, the second one gives the best results.

Lin et al. (2012) combine a well-known capacitated single allocation p-hub median model with integral constraints on the paths. The allocation of a specific p among a set of candidate hubs is required. The total cost on the capacitated hub-and-spoke network is minimized even though at the same time meeting origin-destination demands and singular path constraints. The problem structure and developed a genetic algorithm using the path for encoding is developed. The algorithm gives 0.1% better lower bounds for the Chinese Cargo data set comparing the lower bounds acquired from Lagrangian Relaxation.

Kratica (2013) introduces a new metaheuristic application to hub median problem with multiple allocation. This metaheuristic is called as electromagnetism like method. They also add a new scaling technique with combining movement based on the attraction-repulsion mechanism, which pushes the heuristic possible better search regions. They use well-known data set to measure the effectiveness of their methodology. According to the computational experiments, their proposed method acquires better results in large scale instances. They also extend their study with a capacitated version of the problem. This methodology reaches the optimal solution which exists in the literature for both uncapacitated and capacitated types of the problem.

Correia et al. (2014) extend the well-known capacitated single allocation hub median problem with adding multiple products framework to the network. The study aims to minimize the summation of setup costs for the hubs, setup costs for each product in each hub, and flow routing costs. They extend existing capacitated single allocation hub models for a single product to multiple products and add various inequalities classes to make the model more powerful via lower bound acquired from Lagrangian relaxation.

He et al. (2015) deal with a special case of uncapacitated hub median problem with multiple allocation. This special case is known as an intermodal hub location problem that includes setting up a set of fully interconnected facilities named intermodal hubs. The hubs serve as a transit and swapping points for circulating flows between predetermined origin and destination nodes. For solving the problem, they develop a combined metaheuristic. This metaheuristic includes a complex combination of branch-and-bound, Lagrangian relaxation, and linear programming relaxation.

Rothenbächer et al. (2016) examine the merged tactical planning of hub locations and the design of a density service network. Various constraints acquired from the real-world (multiple transportation requests, flow time limitations, splitting, and outsourcing opportunities) are taken into account. For conceptualizing, they use a path-based model and solve the model via a branch-and-price-and-cut algorithm. Computational experiments are done in a German rail freight company with real data.

2.2 The Hub Center Problem

The hub center problem is a variant of the classical p-center problem. Likewise, the p-center problem, the hub center problem, is also a min-max type problem. This problem is introduced by Campbell (1994). In the study, the objective of the hub center problem is classified under three types : (1) minimizing the maximum cost for any origin-destination path, (2) minimizing the maximum cost of change in any link, (3) minimizing the maximum cost of change link between a hub and an origin/destination node.

Campbell (1994) claims that the first objective type is occurred due to perishable or time-sensitive goods. For this reason, the cost term of the objective function is related to time. As in the first one, second objective cost terms also include time. Maximum time on each link is important in this type. Items required to preserve and restore processes, which is only able to do in hubs, are examples of goods in the second type. The third type's only difference from the second type is considering different specialties on the hub-to-hub links. Campbell (1994) introduces and examines also both single and multiple allocation settings of the problem. Let p and c_{ijkm} as the number of potential hubs and consumed time amount through the path $(i \rightarrow k \rightarrow m \rightarrow j)$. x_{ijkm} equals one if the flow is assigned to the associated path and zero otherwise. Likewise, y_k equals one if the hub k is opened and zero otherwise. After the definition Campbell (1994) formulates a model for the first type as follows :

$$\min \max \{x_{ijkm} c_{ijkm}\} \tag{2.6a}$$

$$\text{s.t.} \sum_{k} y_k = p, \tag{2.6b}$$

$$\sum_{k,m} x_{ijkm} = 1 \qquad \forall i, j, \tag{2.6c}$$

$$x_{ijkm} \leq y_k \qquad \forall i, j, k, m, \tag{2.6d}$$

$$x_{ijkm} \leq y_m \qquad \forall i, j, k, m, \tag{2.6e}$$

$$x_{ijkm},\ y_k \in \{0, 1\} \qquad \forall i, j, k, m. \tag{2.6f}$$

(Kara and Tansel, 2000) is one of the pioneer studies of p-hub center in the literature. In this study, a new modeling approach is used to solve the p hub center problem with single allocation. This approach is essentially a combinatorial model. They also prove that this model is an $\mathcal{NP}$-complete problem. Comparing their model with Campbell (1994) 's model, they use the CAB data set with 80 combinations. For solving Campbell (1994)' model, three different linearizations are used. They use a linearization technique to solve their model. The computational experiments indicate that their model is significantly better than Campbell (1994) 's.

Pamuk and Sepil (2001) develop a metaheuristic algorithm to solve p center problem with single allocation. The relocation part of the metaheuristic is essentially based on tabu search because tabu search lowers the possibility of trapping to local optima. Two different single allocation schemes are combined for the allocation phase of the algorithm. A greedy search heuristic is attached to the algorithm for improving assignments.

Campbell et al. (2007) focus on the p-hub center allocation part of the p-hub center problem with both single allocation and multiple allocation settings. They also investigate both uncapacitated and capacitated versions of the problems. For analyzing this, they acquire complexity results and formulate with integer programming method. Their major drawing from the computational experiments is when there are less than three hubs in the system, the problem is solved with polynomial-time algorithms. However, larger problems are $\mathcal{NP}$-hard. They use a spanning tree algorithm to find an initial solution.

Ernst et al. (2009) investigate the uncapacitated hub center problem with both single and multiple allocation settings. They also prove that both settings of the problem are $\mathcal{NP}$-hard. They also find out that despite the problem with a given predetermined potential hub set is solved to optimal, the problem is still $\mathcal{NP}$-hard. They introduce a novel integer programming model for both settings of the problem. For solving multiple allocation settings, a branch-and-bound based algorithm is used.

Gavriliouk (2009) introduces aggregation in hub location problems for the uncapacitated p-hub center problem with both single and multiple allocation settings. For solving these problems, they focus on improving error bounds with heuristics. The structure of the error bounds leads to using instructions of aggregation for minimizing these errors.

Meyer et al. (2009) deal with a single allocation p-hub center problem with the third type of objective. For solving the problem, they introduce an exact two-phase algorithm. In the first phase, a set of ideal candidate hub combinations is determined via a branch-and-bound algorithm based on the shortest path. The second part aims to reduce the size of the problem, which gives the optimal solution. Improving the upper bound of the first phase, they develop an ant colony-based heuristic, which is design for the single allocation p-hub center problem. They prove that their solution is better than well-known MIP solvers, such as CPLEX.

Brimberg et al. (2017a) focus on the uncapacitated p-hub center with multiple allocation and the first type of objective. They try to solve two different integer programming models. For solving the problem, they use an essential variable search heuristic. They also introduce new multi-start local search heuristic, and they compare their results with Ernst et al. (2009) 's. According to the computational results, the essential variable search heuristic outperforms the other two methods.

Brimberg et al. (2017b) examine the uncapacitated p-hub center problem with single allocation and the first type of objective. They develop a general variable search heuristic to solve the problem. A nested structure with two local search neighborhoods is embraced for the local search step. They also compare their model with a prior study that is a metaheuristic which is based on ant colony optimization. Computational experiments indicate that their method is superior to the existent method in terms of both solution quality and CPU time.

2.3 The Hub Covering Problem

The hub covering problem is a hub location variant of the well-known set covering problem. In this problem, all demand nodes must be covered by hubs in predefined distances. Likewise, in the hub center problem, Campbell (1994) defines three different covering types : the path (i,j) covered by hubs k and m if (1) the cost of the path does not exceed a predefined value, (2) each link in the path does not exceed a predefined value, (3) distance of the each of hub/non-hub links does not exceed a predefined value.

As in the hub center problem, Campbell (1994) introduces integer programming formulations of the hub covering problems with both single and multiple allocations. Besides these models, the hub maximal covering problem is also proposed. The major objective in the hub covering problem is increasing hub coverage with minimizing hub opening costs. In the hub maximal covering problem, only maximizing the coverage of demands is considered.

Kara and Tansel (2003) introduce a new combinatorial integer programming formulation for the hub covering problems with the single assignment, which are different from existent studies. In the prior studies, there are only nonlinear binary integer formulations. In their study, they linearize these models with three different methods and their model with one method. Computational experiments indicate that the linearized version of their model gives superior results from prior studies.

Define x_{ik} , to be 1 if the node i assigned to hub k, 0 otherwise. If the x_{kk} equals to 1, it means that hub k is opened. $x_{ijkm}(= x_{ik}x_{jm})$ is a binary parameter for route assignment. t_{im} symbolizes the transportation time between node i and node m. α represents the inter hub cost discount factor. The coverage is ensured by a binary parameter, v_{ijkm}, which equals 1 if the transfer time between node i and node j $(t_{ik} + \alpha t_{km} + t_{jm})$ is below a threshold time (β) and 0 otherwise. After defining decision variables and parameters, Kara and Tansel (2003)'s model is given below.

$$\min \sum_{k} x_{kk} \tag{2.7a}$$

$$\text{s.t.} \sum_{k} x_{ik} = 1 \qquad \forall\, i, \tag{2.7b}$$

$$x_{ik} \leq x_{kk} \qquad \forall i, k, \tag{2.7c}$$

$$\sum_{km} v_{ijkm} x_{ijkm} \geq 1 \qquad \forall\, i, j, \tag{2.7d}$$

$$x_{ijkm} \geq x_{ik} + x_{jm} - 1 \qquad \forall\, i, j, k, m \tag{2.7e}$$

$$\sum_{k,m} x_{ijkm} = 1 \qquad \forall i, j, \tag{2.7f}$$

$$x_{ijkm} \geq 0, x_{ik} \in \{0, 1\} \qquad \forall\, i, j, k, m. \tag{2.7g}$$

Tan and Kara (2007) design a hub network for a cargo company based in Turkey. The design includes specific constraints, requirements, and criteria for the hub covering problem. They formulate the problem with integer programming framework. To better satisfy the requirements of the real-world, they propose two different formulations. In the first formulation, they assign weights to hubs. For decreasing all criteria into a single weight, Multiattribute Utility Theory is used. The second formulation adds constraints for an obligatory driving limit for drivers, deadlines for cargo, and different variations of real-world needs. Moreover, they solve the model with large scale instances. Besides these, they introduce a new benchmark data set, which is known as the Turkish Cargo (TR) data set, which includes 81 nodes/cities.

Wagner (2008) proposes an improved formulation to hub covering problems with both single and multiple allocations. The formulations include non-increasing quantitydependent transport time functions for transport links. They solve the formulations via MIP solvers. Computational results provide comparisons with both Campbell (1994)'s and Kara and Tansel (2003)'s formulations.

Alumur and Kara (2009) deal with cargo application of the hub covering problem. They introduce a new model that includes the relaxation of the network's completeness and time bounds on each link in the network. The model set is the single allocation, and the model allows upmost three hubs in a path. Well-known CAB data set and Turkish Cargo data set is used for the computational experiments.

Calık et al. (2009) propose an integer programming formulation to solve the hub covering problem over an incomplete network with a single allocation. The model aims to determine hubs' location, hub connections to be linked between the located hubs, and the assignment of non-hub nodes to hub nodes. For solving this model, the tabu search based metaheuristic is used. CAB and TR data set is employed for the experiments.

Qu and Weng (2009) examine the hub maximal covering problem with multiple allocation. They introduce a new evolutionary algorithm which is called path relinking. They indicate that the algorithm performs well to solve the model. Computational experiments are based on the Chinese aerial freight data set, which includes 82 nodes/cities.

Hwang and Lee (2012) analyze the hub maximal covering problem with single allocation. The study's objective is to find the location of the hubs, assign nodes to determined hubs, and maximize the demand pertained by the time limit for traveling. For formulating the problem, a novel integer programming model is introduced. CAB data set and its subsets are used for computational studies. Two different heuristics, which are distance-based allocation and volume-based allocation heuristics, are employed to solve the problem.

Lowe and Sim (2013) introduce a new formulation to the hub covering problems with single allocation. Prior methodologies do not include flow costs. The major contribution of this paper is adding flow costs to the model. Besides this, they also extend their study by increasing the types of aircraft in the system.

Peker and Kara (2015) introduce a partial coverage concept to the hub maximal covering problem with both single and multiple allocations. In prior studies, nodes are completely covered and not covered by hubs. This study extends prior studies with adding partial coverage of nodes. For formulating the problem, they introduce an original efficient mixed-integer programming model. They use the CAB and TR data set to compare their model's performance. Their model usually outperforms prior models.

2.4 Uncertainty in the Hub Location Problem

Uncertainty is inevitable and cannot be ignored in the real world. For the last decade, recent developments in stochastic programming enabled handling uncertainty. While inferences from small scale problems guide us to manage large scale problems, these developments do not allow us to decrease dramatically computational times to solve large scale problems.

Setting a hub to a specific location is a certain decision, and changing it once settled is very costly. The hub location problems inherently include uncertainties. Uncertainties originate from different sources, such as costs, transportation amounts, demands, etc. The common point of these sources is that their values tend to change over time. In the hub, location problems, flow amounts, and transportation costs are changing over time.

Yang (2009) introduces the first two-stage stochastic programming formulation for the hub location problem. The source of uncertainty is the demand, which changes depending on seasonality. In the first stage, locating hubs is the major objective with different demand scenarios during the planning horizon. In the second stage, determining flow paths and allocation of the nodes is done—different demand levels structure scenarios. Comparing the use of expectation of the expected value problem with stochastic programming, they indicate that stochastic programming methodology is favorable for reflecting demand uncertainty to the hub location problem.

Sim et al. (2009) develop a chance-constrained stochastic programming formulation for the hub center problem with single allocation. Chance constraints are used because of meeting the service level guarantee. Uncertainty in the service level guarantee is occurred due to variability in transportation times. They solve the problem with a metaheuristic, a combination of radial heuristic developed by them and the Teitz-Bart heuristic. In their model, hub opening costs are ignored. Computational study indicates that the Teitz-Bart heuristic random initiation generally gives the worst result than the proposed method.

Contreras et al. (2011) investigate demand and transportation costs uncertainty for the uncapacitated hub median problem. They also replace random variables with their expectations. Demand uncertainty expected value problem is equivalent to stochastic programming formulation. For transportation cost uncertainty, this change is not equivalent to stochastic programming formulation. Because of this, they use Monte Carlo simulation-based algorithm, which combines a sample average approximation method with the Benders decomposition algorithm.

Alumur et al. (2012) focus on different aspects of the uncapacitated hub location problem under uncertainty. The hub opening costs and flow amount transferred between nodes are taken as the source of uncertainty. The sources are analyzed separately and together. They also investigate both single and multiple allocation settings for the problem. They find that the cost uncertainty affects the model higher than demand uncertainty. According to computational experiments, formulations, including uncertainty, is worth investigating.

Hult et al. (2013) develop a stochastic programming framework to the uncapacitated hub center problem with single allocation. Their objective is minimizing the longest path distance in the network. They take travel times as the source of uncertainty because travel time directly affects service quality. They reformulate a stochastic hub center problem efficiently and solve it with an algorithm based on variable reduction and separation algorithm. Computational results show that the model is solved to global optimal when the problem scale is small.

Meraklı and Yaman (2016) analyze the robust uncapacitated hub median problem with multiple allocation under polyhedral demand uncertainty. They reflect uncertainty on the model in two different ways. First of them is the hose model, which accepts only having the information about the upper limit of each node's total flow and the second one is the hybrid model that approves having information about both lower and upper limits of the total flow. They formulate using minimax criteria based linear mixed-integer programming model and solve it via an algorithm that includes two Benders decomposition consecutively.

Qin and Gao (2017) formulate a novel uncapacitated hub median problem with a fixed cost and single allocation with uncertain flow amounts. They develop the model with two different methods according to different decision-makers. The objective of the first method is minimizing the expected cost. The second objective is minimizing α-cost. Because of the computational complexity of the problem, a genetic algorithm is used to solve the model. According to computational study, the genetic algorithm performs efficiently.

Correia et al. (2017) introduce a multi-period stochastic optimization formulation framework to capacitated hub median problem with multiple allocation. The multi-period concept means dividing planning horizons to periods. The source of uncertainty in the model is taken hub capacity amount. This choice aims to investigate the effects of increment and decrement of hub capacity on the model. Due to many constraints and decision variables, it is hard to solve medium or large scale models to optimality with general-purpose solvers. Because of this limitation, they extend the framework and the model.

2.5 Summary Remarks for the Hub Location Problems

In our literature survey, over 70 articles about the hub location problem are examined and classified. Table 2.1 classifies literature with 3 main categories : objective function structure (Median, Center and Covering), allocation type (single and multiple), and parameter type (deterministic and stochastic). It is appropriate to divide the literature between three time periods. Before the year 2000, major contributions in the literature is about the new formulations and several concepts for the hub location problem. One of the breakthrough ideas from this period is adding the interhub discount factor to the model. This idea is suggested by O'Kelly and Bryan (1998). This factor bears from the economy of scale application to the network. As in the economies of scale, there are discounted transportation prices when the transferred amount is increased. As the nature of the hub location problems, hub-to-hub transfer amounts are larger than others. So that, it is generally accepted that multiplying hub-to-hub transfer costs with a factor is reflecting this idea to the model. As this concept, there are different concepts about the problem is observed. Between 2000 and 2009, generally, deterministic models are widely used with all median, center, and covering formulations. Due to computational enhancement, the scale of the problem is increased in these years. However, there is also an increase in problem sizes. Because of this increase, heuristic methodologies also developed. After the year 2009, stochastic formulations are introduced. Comparing the other formulation methods, the method needs comprehensive computation resources. Because of the resource need, the number of these studies is limited as shown in Figure 2.1 and Table 2.1. There are also limited studies for capacitated hub problems which are shown in Figure 2.1.

The number of studies on the hub center and hub covering problems is limited as seen in Figure 2.1 and Table 2.1. There is no stochastic programming formulation of the hub covering problem, including flow costs in the literature, and there is also no multiple allocation studies for this problem. To fulfill these gaps, this thesis presents

— a new mathematical model for the multiple allocation hub covering flow problem (MAHCFP),

— new models for the capacity restricted MAHCFP,

— and a new two-stage stochastic programming model for MAHCFP considering flow demand and flow cost uncertainties.

Table 2.1: Summary of the Literature Survey

Article	Objective			Allocation		Parameter	
	Med	Cen	Cov	Sin	Mul	Det	Sto
Hakimi (1964)	✓	✓		✓		✓	
Hakimi (1965)	✓		✓		✓	✓	
O'Kelly (1986a)	✓			✓		✓	
O'Kelly (1986b)	✓			✓		✓	
O'Kelly (1987)	✓			✓		✓	
O'Kelly (1992)	✓			✓		✓	
Campbell (1992)	✓				✓	✓	
Klincewicz (1992)	✓			✓		✓	
S-Kapov and S-Kapov (1994)	✓			✓		✓	
Campbell (1994)	✓	✓	✓	✓	✓	✓	
Ernst and Krishnamoorthy (1996)	✓			✓		✓	
Klincewicz (1996)	✓				✓	✓	
O'Kelly et al. (1996)	✓			✓	✓	✓	
Skorin-Kapov et al. (1996)	✓			✓	✓	✓	
Sohn and Park (1997)	✓			✓		✓	
Bryan (1998)	✓			✓		✓	
Ernst and Krishnamoorthy (1999)	✓			✓		✓	
Marianov et al. (1999)	✓				✓	✓	
Sasaki et al. (1999)	✓				✓	✓	
Ebery et al. (2000)	✓				✓	✓	
Kara and Tansel (2000)		✓		✓		✓	
Sohn and Park (2000)	✓			✓		✓	
Ebery (2001)	✓			✓		✓	
Pamuk and Sepil (2001)		✓		✓		✓	
Mayer and Wagner (2002)	✓				✓	✓	
Kara and Tansel (2003)			✓	✓		✓	
Boland et al. (2004)	✓				✓	✓	
Labbé and Yaman (2004)	✓			✓		✓	

Continued on next page

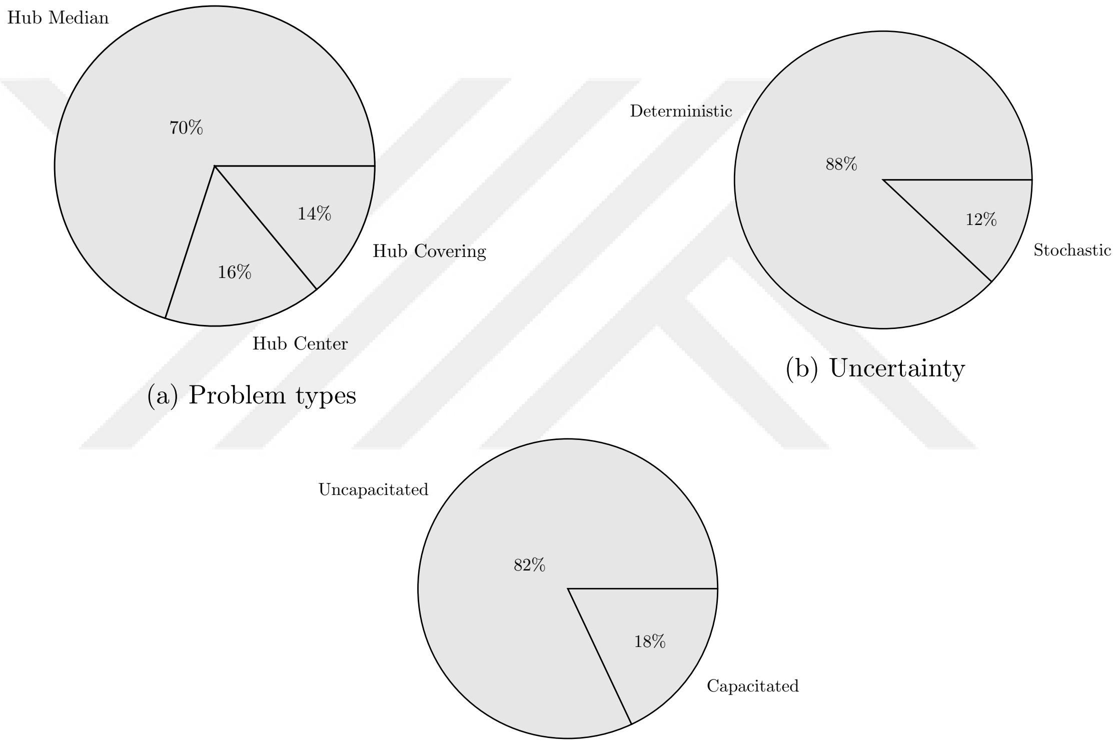
Table 2.1: Summary of the Literature Survey

Article	Objective			Allocation		Parameter	
	Med	Cen	Cov	Sin	Mul	Det	Sto
Elhedhli and Hu (2005)	✓			✓		✓	
Labbé et al. (2005)	✓			✓		✓	
Marín (2005a)	✓				✓	✓	
Marín (2005b)	✓				✓	✓	
Topcuoglu et al. (2005)	✓			✓		✓	
Cánovas et al. (2007)	✓				✓	✓	
Campbell et al. (2007)		✓		✓	✓	✓	
Chen (2007)	✓			✓		✓	
Cunha and Silva (2007)	✓			✓		✓	
Tan and Kara (2007)			✓	✓		✓	
da Graça Costa et al. (2008)	✓			✓		✓	
R-Martín and S-Gonzáles (2008)	✓				✓	✓	
Yaman (2008)	✓			✓		✓	
Wagner (2008)			✓	✓	✓	✓	
Alumur and Kara (2009)			✓	✓		✓	
Calık et al. (2009)			✓	✓		✓	
Ernst et al. (2009)		✓		✓	✓	✓	
Gavrilouk (2009)		✓		✓	✓	✓	
Meyer et al. (2009)		✓		✓		✓	
Qu and Weng (2009)			✓		✓	✓	
Silva and Cunha (2009)	✓			✓		✓	
Sim et al. (2009)		✓		✓			✓
Yaman (2009)	✓			✓		✓	
Yang (2009)	✓			✓			✓
Correia et al. (2010)	✓			✓		✓	
Ilić et al. (2010)	✓			✓		✓	
Contreras et al. (2011)	✓				✓		✓
Ishfaq and Sox (2011)	✓				✓	✓	

Continued on next page

Table 2.1: Summary of the Literature Survey

Article	Objective			Allocation		Parameter	
	Med	Cen	Cov	Sin	Mul	Det	Sto
Kratica et al. (2011)	✓				✓	✓	
Alumur et al. (2012)	✓			✓	✓		✓
Hwang and Lee (2012)			✓	✓		✓	
Lin et al. (2012)	✓			✓		✓	
Hult et al. (2013)		✓		✓			✓
Lowe and Sim (2013)			✓	✓		✓	
Kratica (2013)	✓				✓	✓	
Correia et al. (2014)	✓			✓		✓	
He et al. (2015)	✓				✓	✓	
Peker and Kara (2015)			✓	✓	✓	✓	
Meraklı and Yaman (2016)	✓				✓		✓
Rothenbächer et al. (2016)	✓			✓		✓	
Brimberg et al. (2017a)		✓			✓	✓	
Brimberg et al. (2017b)		✓		✓		✓	
Correia et al. (2017)	✓				✓		✓
Qin and Gao (2017)	✓			✓			✓



(a) Problem types

(b) Uncertainty

(c) Capacity

Figure 2.1: Graphical Representations on the Literature Survey

3 DETERMINISTIC MULTIPLE ALLOCATION HUB COVERING FLOW PROBLEM

Logistics, airline and postal companies are needed to be managed more cost-effective and efficient because of the recent developments in the global economy. Before these days, there were some tackles to cooperation between companies. For example, plane usage was less than today. Technological improvements were not sufficient to develop long-range aircraft with more seats. Support systems (such as luggage delivery system, ticketing, etc.) were not computerized.

On the other hand, the consumption of petroleum was less than today, and petroleum prices were low. Because of these, all airlines tried to fly all points without cooperation with other airlines. According to U.S. Energy Information Administration (2019), known hydrocarbon sources meet the world's demand through 2050. It means that colossal petroleum deficiency will appear shortly. Besides of this inadequacy, there are political instabilities in some of petroleum exporter countries. After political unrest in petroleum exporter countries (such as Iraq, Libya, Syria, and Venezuela), petroleum prices have been boomed. This increment and insufficiency of sources force airline companies to make codesharing agreements to minimize petroleum consumption and operating expenses. Besides petroleum prices, many other factors affect airline companies' profitability and sustainability, such as qualified staff salaries, lounge expenses, etcetera. Besides the airline companies, logistic and postal companies also depend on petroleum prices and other expenses.

For tackling these problems, airlines make codeshare agreements to cooperate and create flight networks more efficient. For example, Turkish Airlines has more than 50 codeshare agreements (Turkish Airlines, 2020). These intercompany flights are connected via significant cities. Inefficient network design causes much more petroleum consumption and inefficient source usage. In the short term, the effect of inefficient network design may not be exact for companies. Many airlines were declared bankruptcy because of bad network design because small cities have not enough demand for creating such complex connections. Because of the hardness of creating such a complex network, hub location problems are vital for the airline industry.

Different hub location problems, which are defined in the prior section, are investigated to create networks. Median and center problems ignore a plane's maximum flight distances. On the other hand, the classical hub covering problem ignores the unit flow costs. Because of this, Lowe and Sim (2013) proposes to combine both advantages of hub covering and median problem in a model. The model is known as the *Hub Covering Flow Problem.* Their model uses single allocation setting. Campbell (1994) defines three types of covering (where i and j represent nodes and k and l represent hubs) :

1. length of a route (from node i to node j through hub k and hub l) does not exceed a value,
2. length of all connections in the network does not exceed a value,
3. length of connections except hub-to-hub connections does not exceed value.

Lowe and Sim (2013) use the third definition of the covering. However, the definition ignores hub-to-hub flight's planes also have limitations.

$$\min \sum_k f_k x_{kk} + \alpha \sum_{i,k,l} c_{kl} y^i_{kl} + \sum_{i,k} c_{ik}(\chi O_i + \delta D_i) x_{ik} \tag{3.1a}$$

$$\text{s.t. } x_{ik} \le A_{ik} x_{kk} \qquad \forall i, k, \tag{3.1b}$$

$$\sum_k x_{ik} = 1 \qquad \forall i, \tag{3.1c}$$

$$O_i Z_{ik} = \sum_j h_{ij} x_{jk} + \sum_l y^i_{kl} - \sum_l y^i_{lk} \qquad \forall i, k, \tag{3.1d}$$

$$\sum_{l, l \ne k} y^i_{kl} \le O_i x_{ik} \qquad \forall i, k, \tag{3.1e}$$

$$y^i_{kl} \ge 0, x_{ik} \in \{0, 1\} \qquad \forall i, k, l. \tag{3.1f}$$

Define x_{ik} , to be 1 if the node i assigned to hub k, 0 otherwise. If the x_{kk} equals to 1, it means that hub k is opened. y^i_{kl} represents the amount of flow originating from node i through hub k and hub l, respectively. Let $h_{ij},\ O_i,\ D_i,\ c_{ij},\ f_k$ demand flow from node i to node j, total demand originating from node i, total demand destined to node i, unit transfer cost from node i to node j, and opening cost for hub k respectively. $\alpha,\ \chi,\ \delta$ represent transfer collection and distribution cost factors, respectively. The coverage is ensured by a binary parameter, A_{ik}, which equals 1 if node i is capable to be served by hub k and 0 otherwise. After defining decision variables and parameters, Lowe and Sim (2013)'s model is given in Eqs.(3.1a)-(3.1f).

In the real world, it is easy to create networks with single allocation. However, most airline company uses multiple allocation. For example, if you want to plan a flight between Istanbul and Auckland via Expedia, there are more than five different hubs to connect two cities with a single stop. Due to this reality, multiple allocation is a more realistic modeling option. Transforming Lowe and Sim (2013)'s model does not reflect multiple allocation setting ultimately. On the other side, their logic behind the covering parameter is instrumental, and it is easy to apply in a mathematical model. After an extensive search for a useful multiple allocation model, a model proposed by Ebery et al. (2000) is more effective than others.

Define x_k as a binary decision variable associated with the location decision of a hub (which means if $x_k = 1$, node k is a hub). $z_{ik},\ q_{ilj}$, and y_{ikl} are decision variables for symbolizing flows. c_{ij} and f_k represent cost parameters for flow and hub opening respectively. $h_{ij},\ O_i$ and D_i illustrate the demand amount from node i to node j, total demand amount which is sent from node i and the total demand amount which is collected from node i respectively. $\alpha,\ \chi$, and δ represent the coefficients per unit of interhub flows, distribution, and collection, respectively. Define Γ_k as the capacity limit for hub k. After definitions, the model of Ebery et al. (2000) is given below.

$$\min \quad \sum_k f_k x_k + \sum_i \left[\alpha \sum_{k,l} c_{kl} y_{ikl} + \chi \sum_k c_{ik} z_{ik} + \delta \sum_{l,j} c_{lj} q_{ilj} \right] \qquad (3.2a)$$

$$\text{s.t.} \quad \sum_k z_{ik} = O_i \qquad \forall i, \qquad (3.2b)$$

$$\sum_l q_{ilj} = h_{ij} \qquad \forall i, j, \qquad (3.2c)$$

$$\sum_i z_{ik} \leq \Gamma_k x_k \qquad \forall k, \qquad (3.2d)$$

$$\sum_l y_{ikl} + \sum_j q_{ikj} - \sum_l y_{ilk} = z_{ik} \qquad \forall i, k, \qquad (3.2e)$$

$$\sum_i q_{ilj} \leq D_j x_l \qquad \forall l, j, \qquad (3.2f)$$

$$z_{ik} \leq O_i x_k \qquad \forall i, k, \qquad (3.2g)$$

$$q_{ilj}, y_{ikl}, z_{ik} \geq 0 \quad x_k \in \{0,1\} \qquad \forall i, k, l, j. \qquad (3.2h)$$

Ebery et al. (2000)'s model is formulated to solve capacitated multiple allocation hub location problems. In this study, uncapacitated formulation and covering concept is also wanted to handle. Because of this, a novel formulation based on this model, and the second definition of the covering is developed in Section 3.1.

3.1 Mathematical Formulation of the Multiple Allocation Hub Covering Flow Problem

In this section, we develop a mathematical model for multiple allocation hub covering flow problems (MHCPF) with the addition of interhub cost discount factor and different capacity constraints. Standard notation for the deterministic problems are given below :

$\mathcal{V}$	set of nodes
h_{ij}	demand flow originating from node $i \in \mathcal{V}$ destined for node $j \in \mathcal{V}$
H	total amount of flow to be sent
O_i	total demand arising from node $i \in \mathcal{V}$
D_j	total demand up to node $j \in \mathcal{V}$
ω_{ik}	flow amount capacity of the links connecting nodes $i, k \in \mathcal{V}$
Γ_k	flow amount capacity of hub $k \in \mathcal{V}$
f_i	cost of establishing a hub at node $i \in \mathcal{V}$
c_{ij}	unit flow cost for the link connecting nodes $i, j \in \mathcal{V}$
d_{ij}	length of the link connecting nodes $i, j \in \mathcal{V}$
A	node coverage matrix (A_{ij} is 1 if node $j \in \mathcal{V}$ can be covered by node $i \in \mathcal{V}$ and 0 otherwise)
B	path coverage matrix ($B_{ikj} = A_{ik}A_{kj}$)
α	interhub flow cost discount factor with $\alpha \in (0, 1)$

It is clear that $O_i = \sum_{j \in \mathcal{V}} h_{ij}$, $D_j = \sum_{i \in \mathcal{V}} h_{ij}$ and $H = \sum_{i,\, j \in \mathcal{V}} h_{ij}$. Hub opening decision is related to the binary decision variable, x_k, which equals 1 if a hub is set to node k, 0 otherwise. Variable z_{ik} denotes the amount of flow sent from node i to hub node k, q_{ilj} the amount of flow sent from node i to node j through hub node l, and y_{ikl} to the amount of flow sent from node i via hub nodes k, and l.

We introduce first the formulation of model **UMHCFP** in Eqs.(3.3a)-(3.3g). The objective in Eq.(3.3a) is to minimize the total cost of establishing hubs and transferring demand through links which are existed in the network by considering the interhub flow cost discount factor. Eq.(3.3b) guarantees that all the demand arisen from node $i \in \mathcal{V}$ is transported through hubs. Eq.(3.3c) ensures that the demand originating from node $i \in \mathcal{V}$ up to node $j \in \mathcal{V}$ is transferred through hubs. Eq.(3.3d) corresponds to the flow preservation constraints at each hub. Eq.(3.3e) and Eq.(3.3f) together guarantees no demand is allowed to transfer directly between non-hub nodes. Finally, Eq.(3.3g) demonstrates the type of decision variables.

$$\min \sum_{k\in\mathcal{V}} f_k x_k + \alpha \sum_{i,k,l\in\mathcal{V}} c_{kl} B_{ikl} y_{ikl}$$

$$+ \sum_{i,k\in\mathcal{V}} c_{ik} A_{ik} z_{ik} + \sum_{i,l,j\in\mathcal{V}} c_{lj} B_{ilj} q_{ilj} \tag{3.3a}$$

$$\text{s.t.} \sum_{k\in\mathcal{V}} A_{ik} z_{ik} = O_i \qquad i \in \mathcal{V}, \tag{3.3b}$$

$$\sum_{l\in\mathcal{V}} B_{ilj} q_{ilj} = h_{ij} \qquad i,j \in \mathcal{V}, \tag{3.3c}$$

$$\sum_{l\in\mathcal{V}} B_{ikl} y_{ikl} + \sum_{j\in\mathcal{V}} B_{ikj} q_{ikj} - \sum_{l\in\mathcal{V}} B_{ilk} y_{ilk} = A_{ik} z_{ik} \qquad i,k \in \mathcal{V}, \tag{3.3d}$$

$$\sum_{i\in\mathcal{V}} B_{ilj} q_{ilj} \leq D_j x_l \qquad l,j \in \mathcal{V}, \tag{3.3e}$$

$$A_{ik} z_{ik} \leq O_i x_k \qquad i,k \in \mathcal{V}, \tag{3.3f}$$

$$q_{ilj}, y_{ikl}, z_{ik} \geq 0 \quad x_k \in \{0,1\} \qquad i,k,l,j \in \mathcal{V}. \tag{3.3g}$$

The following constraints are also introduced in this study :

$$z_{ik} \leq \omega_{ik}(1 - x_i) + H x_i \qquad i,k \in \mathcal{V}, \tag{3.4}$$

$$\sum_{i\in\mathcal{V}} q_{ilj} \leq \omega_{lj}(1 - x_j) + H x_j \qquad l,j \in \mathcal{V} \tag{3.5}$$

$$\sum_{i\in\mathcal{V}} z_{ik} \leq \Gamma_k \qquad k \in \mathcal{V}. \tag{3.6}$$

Constraints in Eq.(3.4) ensure that the amount of flow on the link between node i and hub node k does not surpass the capacity ω_{ik}. In a similar sense, constraints in Eq.(3.5) do not allow an amount of flow to be transported from hub node l to node j outreaching the link capacity ω_{lj}. Constraints in Eq.(3.6) limit the total flow which reaches any hub according to its capacity. When Eqs.(3.4)-(3.6) are added to extend UMHCFP given in Eqs.(3.3a)-(3.3g), the following models can be built :

MHCFP-1 : Eqs.(3.3a)-(3.3g), Eq.(3.4), Eq.(3.5),
MHCFP-2 : Eqs.(3.3a)-(3.3g), Eq.(3.6),
MHCFP-3 : Eqs.(3.3a)-(3.3g), Eq.(3.4), Eq.(3.5), Eq.(3.6).

To compare numerical results, we also associate mathematical models **UMHCP**, **MHCP-1**, **MHCP-2** and **MHCP-3** with UMHCFP, MHCFP-1, MHCFP-2, MHCFP-3 respectively where transportation-related costs in the objective function Eq.(3.3a) are omitted. The constraints of the models are the same.

3.2 Numerical Experiments

For numerical studies, the well-known CAB data set with 25 nodes (Beasley, 1990) and TR data set with 81 nodes (Yetis Kara, 2017) are used. TR data set includes all necessary data for the experiments such as *unit flow costs*, *hub establishing/opening costs*, *intercity links lengths*, and *intercity demand amounts*. Meanwhile, the CAB data set contains only links lengths and demand flows between nodes. It is assumed that unit flow costs are a fraction of the intercity link lengths such that $c_{ij} = d_{ij}/25,000$ for all $i, j \in \mathcal{V}$ and hub establishing costs f_i for all $i \in \mathcal{V}$ are set all equal to $10,000$, $20,000$ or $30,000$ for a given instance. To achieve sensible results in terms of the number of established hubs, the original hub establishing costs of the TR data set are multiplied with 300. CAB data set has a symmetrical structure for all data such that $h_{ij} = h_{ji}$ and $d_{ij} = d_{ji}$ for all $i, j \in \mathcal{V}$. TR data set has a symmetrical structure only for distances.

Interhub flow cost discount factor α is set to 0.2, 0.5 and 0.8 for CAB data set, and 0.4, 0.6 and 0.8 for TR data set. The hub or *node coverage radius* is obtained by multiplying the *coverage ratio* Δ with the longest link's length of the network. Δ should be selected such that the existing network does not contain disconnected sub-networks. Hence, Δ is set to 0.6, 0.7 and 0.8 for CAB data set and 0.55, 0.65 and 0.75 for the TR data set. Then, each element A_{ij} of the *node coverage matrix* is fixed to 1 if the node coverage radius is greater than the length of the link connecting nodes i and j, and 0 otherwise.

Table 3.1: Solution time statistics (in seconds)

Statistics for CAB data set

	Mean Time	Maximum Time	Minimum Time
UMHCFP	31.53	111.33	3.73
MHCFP-1	65.02	196.09	6.69
MHCFP-2	82.08	739.16	4.19
MHCPF-3	107.04	1127.27	6.73

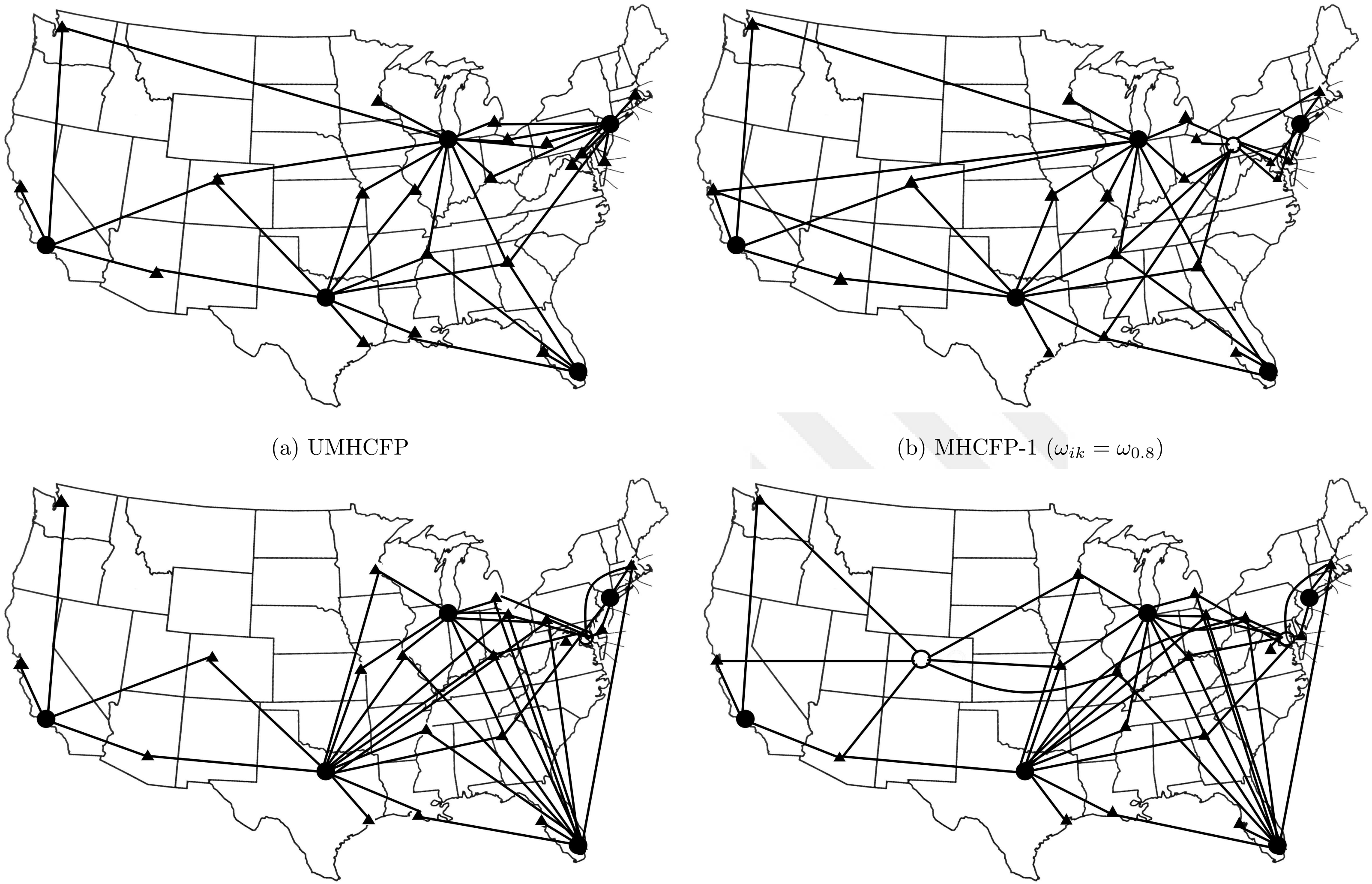
Statistics for TR data set

	Mean Time	Maximum Time	Minimum Time
UMHCFP	26.54	39.27	16.80
MHCFP-1	109.50	332.97	47.03
MHCFP-2	36.45	72.73	18.11
MHCFP-3	68.87	125.19	35.81

Table 3.1 summarizes CPU times of four different types of problems, including different numbers of trials for both CAB and TR data sets. CPU time statistics of UMHCFP, MHCFP-1, MHCFP-2 and MHCFP-3 are acquired from (27, 9), (81,27), (81,27) and (243,81) different settings of parameters for CAB and TR data set respectively. For all problem types, all three statistics of CPU times are better without capacity restrictions. Comparing the limitations on hub and links, the link capacitated problem consumes less time than the hub for minimum CPU time in both data set. The time of MHCFP-2 has better values comparing the MHCFP-1 in the TR data set. The times acquired from the CAB data set show the opposite characteristics of the TR data set times. The solution time statistics of MHCFP-3 are the worst values in the table for the CAB data set. On the contrary, MHCFP-1's statistics are the worst for the TR data set. Each capacity constraint affects the feasible region differently, so CPU time deviates unexpectedly for the distinct constraints.

3.2.1 CAB Data Set Results

Figures 3.1 visualize the network structure of an instance for each proposed model. The figures show that all models follow multiple allocation assignment patterns. 5,6,6 and 7 hubs are opened for UMHCFP, MHCFP-1, MHCFP-2, MHCFP-3 consecutively. It is obvious that adding link capacity causes creating more paths around all



(a) UMHCFP

(b) MHCFP-1 ($\omega_{ik} = \omega_{0.8}$)

(c) MHCFP-2 ($\Gamma_k = \Gamma_{0.2}$)

(d) MHCFP-3 ($\omega_{ik} = w_{0.8}$, $\Gamma_k = \Gamma_{0.2}$)

Figure 3.1: Optimum hub locations for models and CAB data set with $\alpha = 0.50$, $\Delta = 0.70$ and $f_i = 20,000$

USA, and nodes attached more hubs than other models. For MHCFP-2, the east side of the USA has fewer paths comparing other models. Comparing the UMHCFP, there is a more complex network structure in the mid-west of the USA. Constraints of MHCFP-3 supply fewer paths comparing the other capacitated models.

Table 3.2: Summary of UMHCFP Results (CAB data set)

f_i	α	Δ					
		0.80		0.70		0.60	
		NH	IMP	NH	IMP	NH	IMP
10,000	0.80	6	31.67	6	31.43	5	42.62
	0.50	7	42.74	7	42.11	8	50.61
	0.20	9	56.34	9	54.89	10	60.42
20,000	0.80	4	24.79	4	24.38	4	36.28
	0.05	5	32.67	5	32.13	5	41.92
	0.20	5	44.45	5	43.59	5	50.09
30,000	0.08	3	20.08	3	19.89	3	32.17
	0.50	4	25.64	4	24.86	4	35.63
	0.20	5	36.44	5	35.60	5	43.39

Table 3.2 summarizes NH and IMP values of the results of UMHCFP. The minimum value of NH acquired from the results is three, and the maximum value is 10. When the f_i value increases, the NH value decreases. If we decrease α, the NH value rises dramatically. Δ has no meaningful effect on NH values. The best IMP percentage is 60.42, and the worst IMP is 20.08. These values mean that adding flow costs to the covering model lowers total costs for all settings of UMHCFP. The model shows better cost performance when α drops. Fall in Δ levels cause a reverse dome shape effect on IMP, which means that the value decreases first and increases after. An increase in f_i values causes a diminish of IMP values.

Table 3.3 shows the effect of link capacity on the models with NH and IMP values. The NH's maximum value is 12 hubs, and the minimum value of the NH is three hubs. If ω values become more restricted, NH values slightly increase. During the decrease of Δ, NH values increase. The most significant change in NH is observed during the change of α. If the α decreases, NH increases. When f_i climbs up, NH drops. The greatest value of IMP is 70, and the smallest one is 42. Commonly, an increase of f_i causes a fall in IMP values. However, in specific cases, IMP values follow the reverse and normal dome shape pattern. When Δ or ω falls, IMP values create normal or reverse dome shape pattern. The effect of change in α on IMP levels differs according to f_i. For the lower f_i level, a drop of α supplies the increase of IMP. For the middle level, there is no certainty in the patterns. For greater f_i, the fall of α generally causes a drop of IMP. Comparing UMHCFP, NH values are boosted as well as IMP values. Adding flow costs more meaningful for MHCFP-1.

Table 3.4 recaps NH and IMP values for the model with hub capacity. The greatest NH value is 13, and the smallest one is 4. Δ has a slight effect on the model. The other relationships between NH and other parameters are similar to the MHCFP-1. If Γ increases, NH values also increase. Comparing with the prior two models, there are more open hubs. The maximum value of IMP is 70, and the minimum one is 45. When the Δ increases except for the highest capacity level, there is no recognizable pattern for IMP. If Γ decreases, IMP generally falls. If f_i drops, IMP increases. During the drop of α, IMP increases during the lowest level of f_i. For the other levels of f_i, there is no explainable relationship between α and IMP. Comparing the prior models, solving the model with flow costs are more meaningful for MHCFP-2.

Table 3.5 outlines NH and IMP values for MHCFP-3. The greatest NH value is 13, and the smallest one is 4. Hub capacity has a stronger effect on NH from link capacity. For example, if Γ equals to smallest value, ω does not affect NH values. Change of α, Δ, and f_i has a similar effect as prior models. The maximum value of IMP is 74, and the minimum one is 41. When Γ drops, IMP first drops after that increases. This structure is generally observed for ω also. The effects of α and Δ on IMP is not explainable with patterns. Commonly, if f_i increases IMP falls. As it is compared with the prior models, there is no significant change in NH. However, the variation of IMP is much higher than prior variations.

Table 3.3: Summary of MHCFP-1 Results (CAB data set)

ω		$\omega_{0.90}$						$\omega_{0.80}$						$\omega_{0.70}$					
Δ		0.80		0.70		0.60		0.80		0.70		0.60		0.80		0.70		0.60	
f_i	α	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP
10,000	0.80	6	50	7	64	7	59	7	65	7	54	7	59	7	52	7	56	7	57
	0.50	7	54	7	66	8	61	8	65	7	55	8	61	9	56	8	56	8	58
	0.20	10	61	10	70	11	64	10	68	10	59	11	66	11	62	11	59	12	60
20,000	0.80	4	46	4	60	4	55	4	61	4	49	5	56	5	61	5	42	5	57
	0.50	5	46	5	57	5	54	6	59	6	47	6	56	6	60	7	44	6	57
	0.20	6	51	7	57	7	55	7	60	7	49	7	58	7	62	7	50	7	58
30,000	0.80	3	50	3	58	4	54	3	59	3	47	4	57	4	59	4	54	4	58
	0.50	4	48	4	57	4	52	4	55	4	43	4	53	4	57	5	49	5	55
	0.20	5	49	5	58	5	53	5	54	5	42	5	50	6	56	6	46	6	53

Table 3.4: Summary of MHCFP-2 Results (CAB data set)

	Γ	$\Gamma_{0.30}$						$\Gamma_{0.20}$						$\Gamma_{0.10}$					
	Δ	0.80		0.70		0.60		0.80		0.70		0.60		0.80		0.70		0.60	
f_i	α	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP
10,000	0.80	6	61	7	65	7	67	7	64	7	64	7	61	10	61	10	66	10	58
	0.50	7	62	7	67	8	68	8	66	8	66	8	62	12	63	12	66	12	58
	0.20	9	66	9	70	10	70	10	70	10	69	11	65	12	67	12	67	13	60
20,000	0.80	4	57	4	62	4	64	5	57	5	60	6	56	10	51	10	61	10	53
	0.50	6	56	6	61	5	63	6	57	6	60	6	56	10	52	10	59	10	51
	0.20	6	58	6	63	5	63	6	59	6	62	6	58	10	56	10	58	10	50
30,000	0.80	4	55	4	54	4	62	5	59	5	57	5	53	10	50	10	54	10	48
	0.50	4	53	4	51	4	60	5	58	5	55	6	52	10	50	10	52	10	46
	0.20	5	53	5	52	5	59	6	59	6	57	6	53	10	51	10	50	10	45

Table 3.5: Summary of MHCFP-3 Results (CAB data set)

Γ			$\Gamma_{0.30}$						$\Gamma_{0.20}$						$\Gamma_{0.10}$					
Δ			0.80		0.70		0.60		0.80		0.70		0.60		0.80		0.70		0.60	
ω	f_i	α	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP
$\omega_{0.90}$	10,000	0.80	6	71	7	70	7	61	7	56	7	59	7	57	10	68	10	68	10	59
		0.50	7	72	7	72	8	63	8	59	8	61	8	58	12	69	12	68	12	59
		0.20	10	74	10	76	11	66	10	65	10	65	11	62	12	71	12	70	13	61
	20,000	0.80	4	58	4	54	4	58	5	52	6	55	6	56	10	63	10	63	10	53
		0.50	6	58	6	53	6	57	6	53	6	55	6	55	10	62	10	62	10	52
		0.20	6	60	7	54	7	58	6	56	7	58	7	55	10	62	10	62	10	51
	30,000	0.80	4	66	4	51	4	61	5	49	5	52	6	51	10	60	10	55	10	50
		0.50	4	64	4	49	4	58	6	48	6	51	6	51	10	58	10	54	10	48
		0.20	5	63	5	50	5	57	6	51	6	53	6	51	10	57	10	53	10	47
$\omega_{0.80}$	10,000	0.80	7	64	7	65	7	67	7	67	7	60	7	63	10	67	10	65	10	56
		0.50	8	66	7	67	8	67	9	68	8	61	8	63	12	68	12	64	12	57
		0.20	10	70	10	71	11	68	10	71	10	64	11	66	12	70	12	66	13	60
	20,000	0.80	4	57	5	56	5	62	6	64	6	57	6	49	10	62	10	59	10	51
		0.50	6	56	6	56	6	61	6	63	7	56	6	47	10	61	10	58	10	50
		0.20	7	59	7	59	7	60	7	64	7	57	7	45	10	61	10	57	10	50
	30,000	0.80	4	52	4	63	4	60	5	51	5	48	6	53	10	55	10	54	10	50
		0.50	4	50	5	61	5	57	6	48	6	45	6	52	10	55	10	52	10	48
		0.20	6	51	5	62	5	55	6	46	6	45	6	52	10	56	10	51	10	47
$\omega_{0.70}$	10,000	0.80	7	56	7	58	7	64	7	68	7	54	7	62	10	67	10	64	10	67
		0.50	9	58	8	60	8	65	9	69	9	55	8	63	12	68	12	64	12	65
		0.20	11	64	11	64	12	68	11	71	11	59	12	65	12	71	12	67	13	64
	20,000	0.80	5	52	5	54	5	63	6	62	6	62	6	57	10	61	10	59	10	62
		0.50	6	51	7	53	6	61	7	61	7	61	6	56	10	61	10	59	10	58
		0.20	7	54	7	55	7	62	7	61	7	61	7	57	10	62	10	59	10	55
	30,000	0.80	4	56	4	58	4	59	5	52	5	47	6	46	10	57	10	49	10	53
		0.50	5	54	5	56	5	56	6	51	6	43	6	44	10	56	10	47	10	49
		0.02	6	55	6	57	6	55	6	52	6	41	6	42	10	57	10	47	10	47

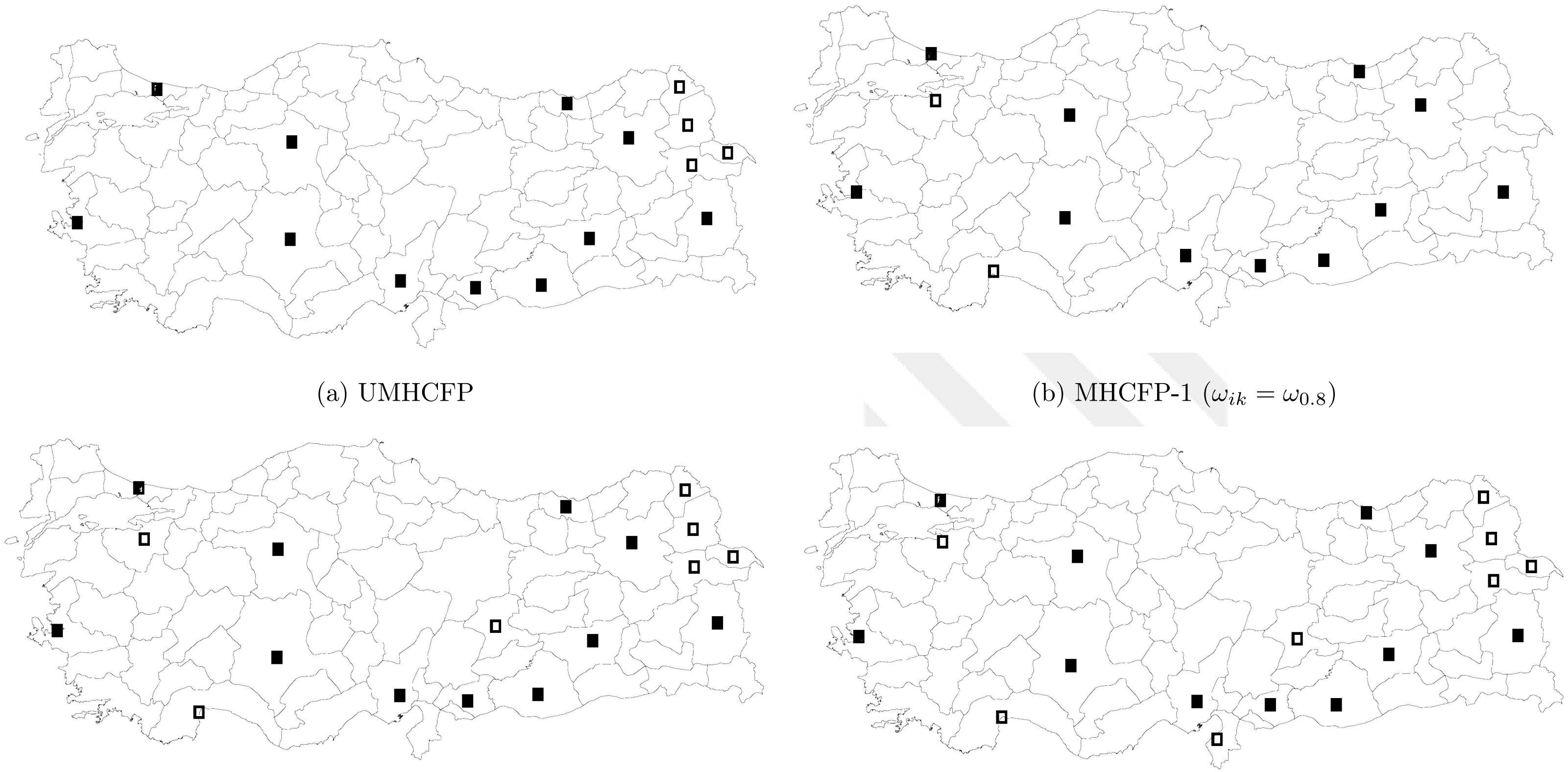
3.2.2 TR Data Set Results

Figures 3.2 visualize the position of opened hubs for an instance for each proposed model. 15, 13, 18 and 19 hubs are opened for UMHCFP, MHCFP-1, MHCFP-2, MHCFP-3 consecutively. The same opened hubs for all models are shown with filled rectangles. Empty rectangles do not occur in all solutions. The lowest number of opened hubs have occurred in MHCFP-1. The common pattern for UMHCFP, MHCFP-2, and MHCFP-3 is more hubs in the northeastern of Turkey. Another drawing from the figures is that approximately no hubs are opened in northern Turkey.

Table 3.6: Summary of UMHCFP Results (TR data set)

α	Δ					
	0.75		0.65		0.55	
	NH	IMP	NH	IMP	NH	IMP
0.80	7	42.97	7	76.85	11	79.33
0.60	7	44.61	15	77.77	15	80.35
0.40	21	47.96	24	79.68	25	82.01

Table 3.6 summarizes NH and IMP values of the results of UMHCFP. The minimum value of NH acquired from the results is seven, and the maximum value is 25. When α decreases, NH increases. Also, this pattern has occurred for Δ. The best IMP percentage is 82.01, and the worst IMP is 42.97. These values mean that adding flow costs to the covering model lowers total costs for all settings of UMHCFP, and the model performs better with TR data set comparing CAB. IMP shows similar reactions with NH to the change in parameters.



(a) UMHCFP

(b) MHCFP-1 ($\omega_{ik} = \omega_{0.8}$)

(c) MHCFP-2 with ($\Gamma_k = \Gamma_{0.2}$)

MHCFP-3 with ($\omega_{ik} = \omega_{0.8},\ \Gamma_k = \Gamma_{0.2}$)

Figure 3.2: Optimum hub locations for models and TR data set with $\alpha = 0.60$ and $\Delta = 0.65$

Table 3.7 shows the effect of link capacity on the models with NH and IMP values. The NH's maximum value is 28 hubs, and the minimum value of the NH is eight hubs. When ω decreases, NH values increase. The effect of ω surpasses the effect of Δ. Because of this, Δ does not affect on the NH.

On the other hand, NH shows the same response to a change in α with UMHCFP. The greatest value of IMP is 87.90, and the smallest one is 70.22. Adding flow cost to MHCFP-1 is more meaningful than UMHCFP. If α decreases, IMP values drop as well. Except for the strictest capacity level, when Δ increases, IMP first decreases after increases.

Table 3.8 recaps NH and IMP values for the model with hub capacity. The greatest NH value is 34, and the smallest one is 9. Comparing the MHCFP-1 Δ has a significant effect on NH for MHCFP-2. When Δ, α or Γ drops, NH increases. The greatest value of IMP is 97.21, and the smallest one is 88.61. It is obvious that MHCFP-2 outperforms both prior models with cost performance. When Γ or α increases, IMP falls. For the greatest Γ value, when Δ decreases, IMP shows dome-shaped behavior. For the other values, Δ and IMP move in the same direction.

Table 3.9 outlines NH and IMP values for MHCFP-3. The greatest NH value is 35, and the smallest one is 10. Γ has a more significant effect on NH than ω. If the Γ increases, NH decreases. Also, when ω decreases, NH increases slightly. Except for the smallest value of Γ, if Δ decreases, NH increases. For the smallest value of Γ, there is a negative relationship between Δ and NH. When α decreases, NH increases. It is obvious that the average NH for all instances reaches the greatest value at MHCFP-3. The maximum value of IMP is 97, and the minimum one is 85. Commonly, when Δ or α falls, IMP also drops. Change of ω level has nearly no effect on IMP. On the contrary, if Γ decreases, IMP increases.

Table 3.7: Summary of MHCFP-1 Results (TR data set)

ω	α	Δ					
		0.75		0.65		0.55	
		NH	IMP	NH	IMP	NH	IMP
$\omega_{0.90}$	0.80	8	79.13	8	78.10	8	87.90
	0.60	10	76.18	10	75.15	10	85.44
	0.40	25	73.00	25	72.08	25	82.12
$\omega_{0.80}$	0.80	9	77.62	9	75.83	9	87.28
	0.60	13	74.74	13	73.00	13	84.80
	0.40	26	71.75	26	70.22	26	81.48
$\omega_{0.70}$	0.80	9	85.46	9	85.43	9	82.40
	0.60	13	83.25	13	83.22	13	80.16
	0.40	28	80.53	28	80.51	28	77.65

Table 3.8: Summary of MHCFP-2 Results (TR data set)

Γ	α	Δ					
		0.75		0.65		0.55	
		NH	IMP	NH	IMP	NH	IMP
$\Gamma_{0.30}$	0.80	9	92.11	12	93.00	11	90.95
	0.60	13	90.63	16	91.82	17	89.83
	0.40	24	88.61	26	90.40	28	88.62
$\Gamma_{0.20}$	0.80	12	95.74	13	93.98	14	93.44
	0.60	14	94.82	18	92.78	18	92.20
	0.40	26	93.41	28	91.13	29	90.49
$\Gamma_{0.10}$	0.80	18	97.21	17	96.81	17	94.53
	0.60	23	96.67	25	96.21	24	93.59
	0.40	33	95.87	33	95.37	34	92.31

Table 3.9: Summary of MHCFP-3 Results(TR data set)

Γ		$\Gamma_{0.30}$						$\Gamma_{0.20}$						$\Gamma_{0.10}$					
Δ		0.75		0.65		0.55		0.75		0.65		0.55		0.75		0.65		0.55	
ω	α	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP
$\omega_{0.90}$	0.80	10	93	12	92	14	92	13	96	13	93	14	92	18	97	17	97	17	93
	0.60	13	91	17	91	18	90	15	95	18	92	18	90	25	96	28	97	25	92
	0.40	26	89	28	90	30	89	27	94	29	91	30	88	34	95	34	96	34	91
$\omega_{0.80}$	0.80	11	95	12	95	14	89	13	96	13	94	14	90	19	96	18	96	18	94
	0.60	14	94	17	94	18	88	16	95	19	93	18	88	25	96	28	96	24	92
	0.40	26	92	28	93	30	87	27	94	29	91	30	87	34	95	34	95	34	91
$\omega_{0.70}$	0.80	13	95	13	94	14	89	14	96	14	94	14	89	20	96	19	97	19	93
	0.60	16	94	19	93	18	88	17	95	20	93	19	87	25	96	28	96	25	92
	0.40	27	92	28	91	30	87	27	93	30	91	30	85	35	95	34	95	34	91

4 STOCHASTIC MULTIPLE ALLOCATION HUB COVERING FLOW PROBLEM

E-commerce has been rapidly gaining share from the global economy since the last two decades. The share of online shopping in total retail sales is increased from 4.2 % to 8.5 % between 2014 and 2018 (Ekonomist Online, 2018). In 2020, it is expected that the share reaches 13%. The volume of parcel delivery has been expanded accordingly. In Turkey, all major parcel delivery corporations increase their revenues and profits. Corporations do not want to lose their customer. Because of that, they invest in creating a network with minimum cost and maximum productivity.

Especially, the volume of parcels increases unexpectedly during the major discount events (i.e., New Year, Valentine's Day) or a health crisis (i.e., COVID-19). Unexpected demand surge also complicates storage activities. These variations highly disturb parcel distributors' daily routines. Delivery times double or sometimes triple of regular times. As a result, the number of unsatisfied customers and churn rate increases.

Besides the fluctuation in demand, there are a lot of uncertainty sources for transportation costs. For example, labor costs may increase when there is a lack of labor. During transportation, transport vehicles (plane, ship, lorry, etc.) generally use petroleum. Petroleum prices change daily. Changes in demand also affect transportation costs.

Uncertainties are unavoidable in real life. Handling uncertainty is a very challenging task considering mathematical optimization models. There are many ways to reflect uncertainty in the models. In this chapter, we formulate the multiple allocation hub covering flow problem as a two-stage stochastic mixed-integer linear optimization model. For showing the appropriation of the model, we also solved models with the expectation of expected value methodology.

4.1 Mathematical Formulation

There are limited stochastic programming studies on hub location and none on hub covering with flow costs. In this section, we develop a mathematical model for stochastic multiple allocation hub covering flow problem (SMHCFP) with three different variations for flow cost uncertainty, demand uncertainty, and demand and cost uncertainty (SMHCFP-C, SMHCFP-D, SMHCFP-DC). We want to note that the associated models are developed based on our deterministic model given in the previous chapter.

Decision variables for all variants are the same. Hub opening decision is related to the binary decision variable x_k, which equals 1 if a hub is set to node k and 0 otherwise. Variable $z_{ik}(s)$ denotes the number of goods sent from node i to hub node k for scenario s, $q_{ilj}(s)$ the number of goods sent from node i to node j through hub node l for scenario s, and $y_{ikl}(s)$ the number of goods sent from node i via hub nodes k and l for scenario s. Common parameters of the 3 variants of SMHCFP are given below.

$h_{ij}(s)$	amount of goods to be sent from node i to node j for scenario s
$O_i(s)$	total amount of goods to be sent from node i $(= \sum_{j \in V} h_{ij})$ for scenario s
$D_j(s)$	total amount of goods to be sent to node j $(= \sum_{i \in V} h_{ij})$ for scenario s
$c_{ij}(s)$	unit transportation cost from node i to node j scenario s

Without determining the location of hubs, we cannot form paths and flow costs. Because of this, we set x as the first-stage decision variable. Eq.(4.1a) is the objective of the first stage problem, which is the summation of expectation of second-stage problem concerning scenarios and hub opening costs. Other constraints are equivalent to the deterministic model. For reflecting cost and demand uncertainty, related parameters become second stage parameters. The realizations of the scenarios are in the second stage. Using the location-allocation problem analogy, the first stage decision must be a location decision. So, determining the hub location is the first stage decision. After location, we decide to demand allocations. So, demand related decision variables must be a second stage decision.

The mathematical model of **SMHCFP-DC** is given as :

$$\min \sum_{k} f_k x_k + E_\xi[Q(x, \xi(s))] \tag{4.1a}$$

$$\text{s.t. } x_k \in \{0, 1\} \qquad \forall k. \tag{4.1b}$$

where

$$Q(x, s) = \min \sum_{i} \left[\sum_{k} c_{ik}(s) z_{ik}(s) + \alpha \sum_{k,l} c_{kl}(s) y_{ikl}(s) + \sum_{l,j} c_{lj}(s) q_{ilj}(s) \right] \tag{4.2a}$$

$$\text{s.t. } \sum_{k} A_{ik} z_{ik}(s) = O_i(s) \qquad \forall i, \tag{4.2b}$$

$$\sum_{l} B_{ilj} q_{ilj}(s) = h_{ij}(s) \qquad \forall i, j, \tag{4.2c}$$

$$\sum_{l} B_{ikl} y_{ikl}(s) + \sum_{j} q_{ikj}(s) - \sum_{l} B_{ilk} y_{ilk}(s) - z_{ik}(s) = 0 \qquad \forall i, k, \tag{4.2d}$$

$$\sum_{i} B_{ilj} q_{ilj}(s) \le D_j(s) x_l \qquad \forall l, j, \tag{4.2e}$$

$$A_{ik} z_{ik}(s) \le O_i(s) x_k \qquad \forall i, k, \tag{4.2f}$$

$$q_{ilj}(s), y_{ikl}(s), z_{ik}(s) \ge 0 \qquad \forall i, k, l, j \tag{4.2g}$$

4.2 Solution Algorithm

Two-stage stochastic linear optimization models with fixed recourse are widely used for modeling uncertainty in an optimization problem. This type of model is introduced by Beale (1955) and Dantzig (1955) and can be represented in Eqs.(4.3a)-(4.3d).

A distinction is made between the first stage and second stage. The first stage decisions are represented by the $(n_1 \times 1)$ vector x. Corresponding to x are the first stage vectors and matrices c, b, and A, of sizes $(n_1 \times 1)$, $(m_1 \times 1)$, and $(m_1 \times n_1)$ respectively. In the second stage, several random events $s \in S$ may realize. For the given realization s, the second stage problem data $q(s)$, $h(s)$, and $T(s)$ become known, where $q(s)$ is $(n_2 \times 1)$, $h(s)$ is $(m_2 \times 1)$ and $T(s)$ is $(m_2 \times n_1)$.

$$\min z = c^T x + E_\xi[\min q(s)^T y(s)] \tag{4.3a}$$

$$\text{s.t. } Ax = b \tag{4.3b}$$

$$T(s)x + Wy(s) = h(s) \tag{4.3c}$$

$$x, y(s) \in \mathcal{R} \tag{4.3d}$$

Each component of q, T, and h is thus a random variable. Let $T_i(s)$ be i th row of $T(s)$. Piecing together the stochastic components of the second stage data, we obtain a vector $\xi^T(s) = (q(s)^T,\ h(s)^T,\ T_1(s),\ ...., T_{m_2}(s))$, with potentially up to $N = n_2 + m_2 + (m_2 \times n_1)$ components. A single random event s influences several random variables, here, all components of ξ.

Let also $\Xi \subset \Re^N$ be the support of ξ, that is, the smallest closed subset in $\Re^N$ such that P(Ξ)=1. As just said, when the random event s is realized, the second stage problem data, q, h, and T become known. Then, the second stage decision $y(s)$ or $y(s, x)$ must be taken. The dependence of y on s is completely different from the dependence of q or other parameters on s. It is not functional but simply indicates that the decisions that the decision y are typically not the same under different realizations of s. They are chosen so that constraints Eq.(4.3c) and Eq.(4.3d) hold almost surely, i.e., for all $s \in S$ except perhaps for sets with zero probability.

The objective function Eq.(4.3a) contains a deterministic term $c^T x$ and the expectation of the second stage objective $q(s)^T y(s)$ taken over all realizations of the random event s. The second stage term is the more difficult one because, for each s, the value of $y(s)$ is the linear program's solution. To stress this fact, one sometimes uses the notion of a deterministic equivalent program. For the given realization s, let

$$Q(x, \xi(s)) = \min_{y} \left\{ \; q(s)^T y \mid Wy = h(s) - T(s)x, \; y \geq 0 \; \right\} \tag{4.4}$$

be the second stage value function. Then define the expected second stage value function

$$\mathscr{Q}(x) = E_{\xi} Q(x, \xi(s)) \tag{4.5}$$

and the deterministic equivalent program

$$\min z = c^T x + \mathscr{Q}(x) \tag{4.6a}$$

$$\text{s.t. } Ax = b \tag{4.6b}$$

$$x \in \mathcal{R}. \tag{4.6c}$$

This representation of a stochastic program clearly illustrates that a deterministic formulation's significant difference is in the second stage value function. If that function is given, a stochastic program is just an ordinary nonlinear program.

Formulation Eqs.(4.3a)-(4.3d) is the simplest form of a stochastic two-stage program. Extensions are easily modeled. For example, if the first stage or second stage decisions are to be an integer, constraint Eq.(4.3d) can be replaced by a more general form :

$$x \in X, \; y \in Y, \tag{4.7}$$

where $X = Z_{+}^{n_1}$ and $Y = Z_{+}^{n_2}$. Similarly, nonlinear first stage and second stage objectives or constraints can easily be incorporated.

There are three subtypes of two-stage stochastic problems : complete recourse, relatively complete recourse, and incomplete recourse problem. If any feasible value of the first stage decision variable causes the second stage's infeasibility, the model is classified as an incomplete recourse problem. Suppose all first-stage decision variables of our model equal zero, the second stage of our problem becomes infeasible. Because of this, our problem is an incomplete recourse stochastic problem.

For solving this type of problem, we use an L-Shaped algorithm. The basic idea of the L-Shaped algorithm is to approximate the nonlinear term in the objective of these problems. A general principle behind this approach is that the recourse function involves a solution of all second stage recourse linear programs because of the nonlinear objective term. We want to avoid numerous function evaluations. Therefore, we use that term to build a master problem in x, but we only evaluate the recourse function precisely as a subproblem.

In the L-Shaped algorithm, there are two types of cuts : feasibility and optimality. Optimality cuts must be added in the whole L-shaped algorithm used solution methodologies. However, feasibility cuts are only used if the problem is an incomplete recourse problem. Because our problem is an incomplete recourse problem, we must add feasibility cuts to the solution algorithm. The L-shaped algorithm is given as Algorithm 1 below.

Algorithm 1 L-Shaped Algorithm (Birge and Louveaux, 2011)

Input $r = t = \nu = 0$

1: Set $\nu = \nu + 1$. Solve the first stage linear optimization model Eqs.(4.8a)-(4.8e) :

$$\min z = c^T x + \theta \tag{4.8a}$$

$$\text{s.t. } Ax = b \tag{4.8b}$$

$$D_l x \geq d_l \qquad l = 1, ..., r, \tag{4.8c}$$

$$E_l x + \theta \geq e_l \qquad l = 1, ..., t, \tag{4.8d}$$

$$x \in \mathcal{Z},\ \theta \in \mathcal{R} \tag{4.8e}$$

Assume $(x^\nu,\ \theta^\nu)$ is an optimal solution for the model. If there is not Eq.(4.8d) in the model, set θ^ν to $-\infty$ and ignore θ^ν in the computation of x^ν.

2: Control $x \in K_2$. If not, solve the linear for $k = 1, ..., K$

$$\min w' = e^T \nu^+ + e^T \nu^- \tag{4.9a}$$

$$\text{s.t. } Wy + I\nu^+ - I\nu^- = h_k - T_k \tag{4.9b}$$

$$y, \nu^+,\, \nu^- \tag{4.9c}$$

where $e^T = (1, ..., 1),\ h_k,\ T_k$ are stochastic parameters in the second stage, and y is a second stage decision variable, until for any k, the optimal value $\omega' > 0$. In this situation, let σ^ν be the associated dual variables and define $D_{r+1} = (\sigma^\nu)^T T_k$, and $d_{r+1} = (\sigma^\nu)^T h_k$ to create a feasibility cut (which represented in Eq.(4.8c). Set $r = r + 1$, add to the first stage problem and return to Step 1. If for all k, $\omega' = 0$, go to Step 3.

3: For all possible scenarios $(k = 1, ..., K)$ solve the second stage linear optimization model,

$$\min w = q_k^T y \tag{4.10a}$$

$$\text{s.t. } Wy = h_k - T_k X^\nu \tag{4.10b}$$

$$y \geq 0. \tag{4.10c}$$

Let the π_k^ν be the dual variables derived from the optimal solution of an instance k of the model shown in Eq.(4.10b). Define $E_{s+1} = \sum_{k=1}^K p_k (\pi_k^\nu)^T T_k$ and $e_{s+1} = \sum_{k=1}^K p_k (\pi_k^\nu)^T h_k$. Let $\omega^\nu = e_{s+1} - E_{s+1}$. If $\theta^\nu - w^\nu \geq TOL$, stop; the optimal solution is found. Otherwise set $s = s + 1$, add optimality cut which is described in Eq.(4.8d) and go back to Step 1.

4.3 Numerical Studies

After this introductory information, the results of the prior demand uncertainty and cost uncertainty with different distributions, and both uncertainties are given below. We analyze 50, 25, and 10 randomly generated equally possible scenarios for each instance. To eliminate bias and analyze the model's behavior, more stable ten instances are generated for each setting. When the MP=% VSS/RP value is higher, the RP problem is more useful, and the validity of the stochastic programming model is more valid. Another performance measure is the mean of the percentage of the hub costs over the total cost, which is symbolized as HCR. The hub opening cost ($f_i = f$) for all potential hubs is selected 30,000 and 50,000. Two levels of interhub discount factor are analyzed : 0.8, 0.4. In some sense of the model's deterministic version, the covering radius is chosen as 0.8, 0.7, and 0.6. We also represent MNH, which stands for the mean number of hubs over instances as a performance measure. To examine the effect of demand uncertainty on the VSS, a random demand data generation approach is implemented for all models and instances.

All instances of the problem are formulated using Python 2.7 and are solved with Gurobi 8.0.1 on a workstation with dual Intel Xeon E5-2670 (2.6 GHz each) processors and 256 GB of RAM. All run-time statistics are given in real-time.

4.3.1 Results of SMHCFP-C

It is assumed that flow costs are distributed normally. The mean of the distribution (μ) is assumed to be equal to deterministic flow costs. The standard deviation (σ) is obtained by multiplying the mean with a scalar. The scalar is selected to 0.1 and 0.2. The results of the computational studies are summarized in Tables 4.1 and 4.2.

Using scenario-based stochastic programming is meant for different levels of parameters. Besides, the mean of MP is more significant than zero ; choosing EEV methodology is preferable for some instances. The mean of MP values changes between 0.33 and 2.33. Lowering f values with an increasing number of S generally increases MP values. More significant values of α cause small MP levels for nearly all instances. Commonly, when σ increases, MP increases too. The effect of Δ MP changes frequently depends on other parameters.

HCR values deviate between 12.80 and 31.76. These ratios are moderate levels for a network problem. There is a significant gap between HCR values for hub opening cost value changes. For small instances, a higher f value provides higher HCR values. If the instances become more extensive, the gap is decreasing significantly. When the α value decreases, the HCR value increases. Δ value causes different behaviors on HCR values under various parameters.

It is expected and observed that an increase in scenario number causes a boost in run times. The mean run time differs between 247 and 3735 seconds. If α increases, run time increases enormously. The relationship between Δ and run times do not follow a classical pattern. When σ becomes smaller, solving the model consumes less time. The relation between f and run times follows the same pattern too.

MNH values differ between 2.00 and 4.50. The most significant change in MNH value is observed during the change of f. Smaller values of f cause higher MNH values. When Δ covers smaller areas, MNH usually increases as expected. The same pattern is observed between α and MNH. Generally, dropping σ decreases MHN values. In most cases, increasing S causes higher MNH values.

4.3.2 Results of SMHCFP-D

Demands are generated from Poisson distribution. When generating random demands, each time the mean of the distribution is selected, such that it is equal to the deterministic CAB data set demand. The summary of the results is given in Table 4.3.

According to the minimum values of MP, using scenario-based stochastic optimization formulation for handling demand, uncertainty is meaningful for all parameters comparing the EEV approach. MP values change between 0.57 and 5.71. When a

Table 4.1: Results of SMHCFP-C ($\sigma = 0.1\mu$)

			MP						HCR		Run Times						MNH	
f			30,000			50,000			30,000	50,000	30,000			50,000			30,000	50,000
S	α	Δ	Mean	Min	Max	Mean	Min	Max			Mean	Min	Max	Mean	Min	Max		
10	0.8	0.8	0.66	0.00	1.20	0.59	0.00	1.16	15.91	23.97	414.85	372.21	484.13	247.22	206.16	281.67	2.00	2.00
	0.8	0.7	0.53	0.09	1.15	0.34	0.05	0.69	17.31	22.76	408.97	296.61	485.67	228.69	184.17	277.34	2.30	2.00
	0.8	0.6	1.46	0.45	2.18	0.54	0.00	1.77	18.78	21.03	449.94	388.26	509.82	267.72	213.89	304.10	2.70	2.00
	0.4	0.8	1.57	0.54	3.12	0.66	0.07	0.94	29.23	24.05	485.69	355.42	582.19	347.58	302.73	394.80	3.50	2.00
	0.4	0.7	0.93	0.39	1.65	0.53	0.18	1.06	32.32	24.34	381.11	309.20	443.05	299.88	242.63	355.59	4.00	2.10
	0.4	0.6	1.32	0.00	2.86	1.46	0.05	1.91	26.43	27.57	525.85	348.47	677.92	332.34	270.33	390.97	3.60	2.60
25	0.8	0.8	0.83	0.16	2.24	0.54	0.12	1.57	15.94	23.97	1143.97	811.78	1323.54	673.00	583.00	812.24	2.00	2.00
	0.8	0.7	0.33	0.00	0.80	0.30	0.00	0.73	15.03	22.76	1196.02	1124.21	1331.40	614.10	472.61	759.61	2.00	2.00
	0.8	0.6	2.04	0.87	3.07	0.64	0.00	2.95	19.58	21.05	1138.33	936.81	1508.87	764.52	581.12	845.63	2.80	2.00
	0.4	0.8	1.60	0.76	2.28	1.22	0.24	2.34	31.76	24.22	1298.62	1102.91	1463.03	804.00	696.45	972.77	3.80	2.00
	0.4	0.7	1.27	0.25	2.83	0.43	0.00	0.83	29.21	23.19	855.71	648.00	1087.41	679.53	599.93	778.55	3.60	2.00
	0.4	0.6	1.77	0.77	3.34	0.96	0.17	1.95	26.56	30.79	1136.10	911.40	1307.26	715.76	636.50	865.76	3.60	2.90
50	0.8	0.8	1.15	0.00	2.71	1.41	0.15	2.98	14.25	20.82	2509.35	1742.38	3607.88	1187.39	960.88	1369.89	2.10	2.00
	0.8	0.7	1.11	0.45	2.18	1.35	0.00	3.39	16.38	19.44	3581.87	2840.89	4661.12	1578.54	1007.10	2144.08	2.60	2.00
	0.8	0.6	1.64	0.15	3.14	0.37	0.00	1.25	17.42	20.98	2465.66	1645.76	3215.09	1273.61	1135.06	1363.76	2.50	2.00
	0.4	0.8	1.69	0.03	2.47	1.79	1.11	1.99	31.51	21.02	4256.35	3556.14	5851.14	1921.77	1758.59	2111.94	4.40	2.00
	0.4	0.7	2.22	0.78	3.07	1.33	0.35	2.85	28.97	22.46	4285.53	3372.27	5571.72	2507.37	1820.68	2985.08	4.30	2.30
	0.4	0.6	1.47	0.18	2.67	0.74	0.11	2.07	28.67	27.54	2476.07	2200.14	2908.89	1655.40	1427.00	1880.71	3.90	2.60

Table 4.2: Results of SMHCFP-C ($\sigma = 0.2\mu$)

			MP						HCR		Run Times						MNH	
f			30,000			50,000			30,000	50,000	30,000			50,000			30,000	50,000
S	α	Δ	Mean	Min	Max	Mean	Min	Max			Mean	Min	Max	Mean	Min	Max		
10	0.8	0.8	0.70	0.00	1.78	0.79	0.14	1.59	15.11	22.91	608.35	421.82	784.45	292.90	246.33	348.96	2.00	2.00
	0.8	0.7	0.85	0.00	2.10	0.76	0.29	1.38	15.71	21.71	697.34	517.07	847.67	291.35	227.66	411.43	2.20	2.00
	0.8	0.6	1.02	0.00	2.38	1.18	0.00	3.41	15.75	20.13	793.85	487.65	1034.25	316.76	270.29	389.53	2.40	2.00
	0.4	0.8	1.66	0.60	2.47	0.61	0.00	1.49	29.76	29.95	628.81	553.36	713.61	450.68	398.00	550.51	3.70	2.60
	0.4	0.7	1.23	0.00	2.75	1.47	0.39	3.19	28.72	31.24	528.78	386.91	655.48	314.68	272.25	371.65	3.70	2.80
	0.4	0.6	1.55	0.51	2.73	1.61	0.08	2.98	26.68	25.47	752.74	585.00	915.20	391.34	325.30	468.89	3.80	2.50
25	0.8	0.8	0.54	0.00	1.58	0.84	0.34	1.43	17.36	22.94	1837.98	1536.25	2194.85	806.09	659.90	1001.90	2.30	2.00
	0.8	0.7	1.05	0.00	2.35	0.87	0.00	1.88	17.92	21.77	1692.76	1407.90	1989.16	754.94	609.89	871.69	2.50	2.00
	0.8	0.6	1.10	0.38	1.94	0.87	0.04	1.90	17.05	20.06	2048.75	1334.00	2394.62	843.54	613.66	1013.26	2.60	2.00
	0.4	0.8	1.92	0.92	2.85	0.43	0.13	0.78	29.86	27.64	1477.09	1129.54	1839.49	1059.02	863.26	1241.61	3.70	2.40
	0.4	0.7	1.52	0.00	2.72	0.72	0.00	1.32	29.58	25.49	1301.69	1089.82	1780.77	962.47	812.64	1106.01	3.80	2.30
	0.4	0.6	2.33	1.75	2.70	1.65	0.00	3.10	28.31	27.54	1557.87	1324.35	1757.81	924.50	675.63	1142.04	4.00	2.70
50	0.8	0.8	1.04	0.37	1.87	0.74	0.35	1.04	12.80	19.61	4231.95	3677.72	5061.33	1742.58	1553.88	1901.13	2.00	2.00
	0.8	0.7	1.36	0.40	2.52	1.46	0.00	3.00	14.33	18.46	5472.38	4479.58	6124.73	2000.23	1713.08	2480.60	2.40	2.00
	0.8	0.6	1.61	0.00	2.50	0.73	0.00	1.69	18.47	20.01	3632.79	2979.76	5163.67	1876.36	1556.03	2387.47	2.80	2.00
	0.4	0.8	1.64	0.69	2.64	1.23	0.62	2.80	30.96	19.94	5640.12	4315.56	7123.57	2534.82	2004.28	2876.81	4.50	2.00
	0.4	0.7	1.92	0.93	3.56	0.85	0.00	2.33	28.95	22.21	6093.03	4590.63	7537.86	3563.88	2679.32	4443.54	4.50	2.40
	0.4	0.6	1.45	0.00	2.35	1.33	0.00	2.70	26.64	26.40	3735.28	3104.06	5241.19	2034.58	1669.77	2276.59	3.80	2.60

hub covers larger areas, the scenario-based approach becomes less meaningful. If α decreases, MP values drop slightly. Increasing f value generally causes falling of MP values. When the S value rises, the MP value drops commonly.

Demand uncertainty provides more stable HCR values comparing to the results of the cost uncertainty model. HCR values differ from 20.04 and 33.82. It means that the total hub establishing cost does not exceed one of the third of the total cost. Increasing f causes greater HCR values. There is not a standardized relationship between Δ and HCR. This situation was also observed between S and HCR. If α rises, HCR also enlarges too.

This model type consumes much more run times comparing to the results of the cost uncertainty model. The run times changes between 235.87 and 41938.82 seconds. Changing f has the strongest effect on run times. Decreasing f may cause increasing run times to increase to five times. If the covered area becomes smaller, run times become shorter. When the α decreases, on the contrary, run times increase. It is expected to the rise of S causes to raise in the run times.

Uncertain demands cause to open more hubs than uncertainty in costs. Lowering f causes to open more hubs. Changing Δ has different effects on MNH according to f values. For smaller f values, increasing Δ causes opening more hubs. On the contrary, greater f values with smaller Δ causes to open more hubs. There is not an explainable relationship between S and MNH values. Greater α values provide to open fewer hubs.

4.3.3 Results of SMHCFP-DC

For observing both uncertainties in the same model, two scenario creation methodology combined. This process causes different scenario structures in some S levels. For example, transportation costs ($\mathbf{c}^1$, $\mathbf{c}^2$, $\mathbf{c}^3$, $\mathbf{c}^4$, $\mathbf{c}^5$) and demands($\mathbf{h}^1$, $\mathbf{h}^2$) are combined to create 10 scenarios as an example to the methodology (This type has "a" extension). Another methodology example is the combination of transportation costs ($\mathbf{c}^1$, $\mathbf{c}^2$) with demands($\mathbf{h}^1$, $\mathbf{h}^2$, $\mathbf{h}^3$, $\mathbf{h}^4$, $\mathbf{h}^5$) (This type has "b" extension). Because of using two different methodologies, there are a demand surpassed method

Table 4.3: Results of SMHCFP-D

			MP						HCR		Run Times						MNH	
f			30,000			50,000			30,000	50,000	30,000			50,000			30,000	50,000
S	α	Δ	Mean	Min	Max	Mean	Min	Max			Mean	Min	Max	Mean	Min	Max		
10	0.8	0.8	2.94	1.76	3.71	2.32	1.27	3.11	20.91	30.50	6558.28	5146.32	9761.82	1415.58	1112.07	1767.81	3.00	3.00
	0.8	0.7	3.00	2.32	3.64	3.04	2.80	3.31	20.86	29.63	4658.60	3624.91	5952.06	928.55	842.78	1049.56	3.00	2.90
	0.8	0.6	4.32	3.41	4.88	5.20	3.82	5.65	20.38	20.04	2644.17	2194.63	3986.46	432.16	376.26	567.66	3.00	2.00
	0.4	0.8	2.68	2.30	3.25	1.85	0.89	2.58	31.85	33.12	6407.85	5019.89	7114.86	939.74	769.66	1185.28	4.10	3.00
	0.4	0.7	2.62	2.12	2.88	1.85	0.55	2.48	31.54	32.82	4561.38	3711.09	5535.26	677.78	544.58	910.46	4.10	3.00
	0.4	0.6	2.22	1.82	2.63	4.40	3.68	5.30	33.77	31.66	3977.61	3302.76	4633.88	335.47	235.87	408.86	4.60	3.00
25	0.8	0.8	3.08	2.18	3.76	2.51	1.80	3.10	20.93	29.53	9497.72	7262.51	12534.15	2556.93	2007.93	3111.25	3.00	2.90
	0.8	0.7	3.18	2.34	3.72	2.80	2.15	3.28	20.90	29.55	6901.10	5340.39	8685.97	1831.86	1637.48	2042.46	3.00	2.90
	0.8	0.6	3.90	2.85	4.91	5.46	5.16	5.71	20.27	21.06	5582.30	3580.70	8524.25	748.97	699.02	780.52	3.00	2.10
	0.4	0.8	2.78	2.41	3.44	2.04	0.82	2.58	31.10	33.18	9565.07	7646.37	11130.21	1612.77	1427.45	2166.49	4.00	3.00
	0.4	0.7	2.30	2.11	2.77	2.22	0.57	2.63	32.97	32.95	7913.21	6967.67	8681.55	1182.38	1041.84	1662.34	4.30	3.00
	0.4	0.6	2.29	1.94	2.66	4.11	3.70	5.45	33.78	31.52	6256.40	5385.56	7111.51	759.26	525.98	858.36	4.60	3.00
50	0.8	0.8	2.96	2.09	3.56	2.49	1.97	3.11	20.91	30.55	33944.69	27540.10	41938.82	14360.21	12293.98	15932.45	3.00	3.00
	0.8	0.7	2.92	2.27	3.61	2.77	2.10	3.16	20.85	28.55	27190.09	22312.02	33719.07	10813.05	9582.47	12507.10	3.00	2.80
	0.8	0.6	4.14	3.66	4.71	5.47	5.14	5.68	20.31	21.06	7787.37	6147.84	9933.37	1494.71	1272.17	1813.29	3.00	2.10
	0.4	0.8	2.58	2.19	3.30	1.64	0.89	2.53	31.86	33.05	34918.20	29107.92	39497.54	11202.87	9000.48	13624.98	4.10	3.00
	0.4	0.7	2.33	1.87	2.86	2.36	1.77	2.54	31.45	32.99	29493.94	26329.37	33623.32	7214.64	6826.40	8593.48	4.10	3.00
	0.4	0.6	2.37	1.87	2.93	3.96	3.77	4.87	30.88	31.45	10512.30	8980.53	12579.64	1447.03	1205.08	1565.80	4.20	3.00

and a cost surpassed method for 10 and 50 scenarios. For 25 scenarios, there are not any surpassed sides.

Comparing the single source of uncertainty models, the results of the model indicate that using scenario-based stochastic optimization methods are more meaningful. The MP values deviate between 1.94 and 13.05. If σ increases, MP decreases. When α falls, the MP value also falls. If Δ drops, the MP value first drops, after increases generally. Increasing f provides an increment in MP. The demand surpassed method gives better results than the cost surpassed method. Increasing scenario numbers makes the MP better.

HCR values acquired from the results of the model are located between HCR of SMHCFP-D and SMHCFP-C. HCR differs from 15.14 and 32.50. If σ becomes smaller, the deviation of HCR occurs in a smaller range. Increasing α causes worsened HCR levels in common. When f is increasing, HCR is generally increasing. There is not any explainable relationship between HCR and Δ. If Increasing or decreasing S has approximately no effect on HCR values. There is also approximately no difference in HCR values for two scenario creation methods.

The shortest run times are observed in SMHCFP-DC comparing with the prior results. Run times differ from 113 to 2420. Increasing σ causes an increase in run times. If S increases, it is expected and observed that the run times are also increased. It is generally observed that if the model covers less area, the run times become smaller. When f drops, run times climbs up. α has no explainable relation with run times. There is an ignorable difference between the methods for run times.

MHN deviates between 2.00 and 4.00. When σ increases, MNH increases commonly. The significant difference in MHH is observed during the change of α. Decreasing α causes a dramatic increase in MNH. When f rises, MNH drops. Changing S or the method has a relatively very small effect on MNH. Generally, covering more area provides opening fewer hubs.

Table 4.4: Results of SMHCFP-DC ($\sigma = 0.1\mu$)

			MP						HCR		Run Times						MNH	
f			30,000			50,000			30,000	50,000	30,000			50,000			30,000	50,000
S	α	Δ	Mean	Min	Max	Mean	Min	Max			Mean	Min	Max	Mean	Min	Max		
10a	0.8	0.8	5.54	4.86	7.30	5.73	5.28	6.31	15.95	23.96	239.39	159.38	293.69	149.26	137.61	167.17	2.00	2.00
	0.8	0.7	4.77	2.94	7.23	4.51	3.13	6.01	15.22	23.02	201.59	113.22	291.89	122.03	66.13	177.80	2.00	2.00
	0.8	0.6	5.61	4.57	6.55	7.22	6.11	7.96	20.21	21.15	236.70	196.11	297.55	112.95	95.89	139.03	2.90	2.00
	0.4	0.8	3.95	3.32	4.63	5.21	4.66	5.68	30.01	24.09	271.10	223.75	331.20	161.57	135.12	181.50	3.60	2.00
	0.4	0.7	4.11	3.08	5.47	4.30	3.20	5.63	29.26	24.49	179.95	130.35	229.85	139.03	109.61	204.25	3.60	2.10
	0.4	0.6	3.92	2.73	4.82	4.81	4.32	6.14	25.79	25.45	216.32	192.50	233.72	136.53	116.68	151.39	3.50	2.40
10b	0.8	0.8	8.28	7.28	10.20	8.58	7.53	9.23	16.75	23.94	219.46	155.40	288.16	119.38	89.16	151.47	2.10	2.00
	0.8	0.7	6.27	4.68	7.56	5.91	4.60	7.49	16.72	22.98	203.58	128.44	303.60	102.94	79.78	124.64	2.20	2.00
	0.8	0.6	7.82	6.52	8.39	10.81	9.50	13.05	17.41	21.22	218.42	174.06	266.83	105.25	90.68	142.01	2.50	2.00
	0.4	0.8	5.84	4.93	7.34	7.93	7.23	8.48	30.09	24.12	232.45	157.92	335.49	146.27	126.39	166.96	3.60	2.00
	0.4	0.7	5.73	4.16	6.74	6.12	4.95	8.19	28.47	29.20	184.75	126.24	251.69	135.22	90.92	166.22	3.50	2.50
	0.4	0.6	5.26	3.93	6.24	7.60	6.44	9.06	23.59	26.67	191.13	140.13	244.29	121.56	99.24	153.80	3.20	2.50
25	0.8	0.8	8.42	7.68	10.07	8.77	8.17	10.08	15.99	23.98	522.19	403.87	573.53	316.35	285.32	354.33	2.00	2.00
	0.8	0.7	5.86	4.55	7.63	5.77	4.58	7.67	16.62	22.91	534.25	419.33	660.04	296.59	184.02	393.25	2.20	2.00
	0.8	0.6	7.75	6.80	9.31	10.56	9.50	11.73	16.70	21.14	546.12	366.56	629.28	301.10	256.63	354.21	2.40	2.00
	0.4	0.8	5.62	5.07	6.74	7.88	7.19	8.78	28.38	24.11	629.01	497.19	761.35	418.27	338.88	465.45	3.40	2.00
	0.4	0.7	5.62	4.41	6.81	5.79	5.16	7.01	29.22	29.07	448.48	380.67	532.13	346.04	270.55	413.59	3.60	2.50
	0.4	0.6	5.12	3.90	6.37	7.24	6.43	8.40	24.30	23.36	544.90	440.98	731.97	348.98	290.82	396.71	3.30	2.20
50a	0.8	0.8	7.94	7.56	9.03	8.69	8.06	9.55	15.91	23.95	1155.00	907.42	1322.51	675.41	607.23	760.88	2.00	2.00
	0.8	0.7	6.20	4.93	7.91	5.51	4.52	8.00	15.15	22.83	1090.42	823.49	1524.99	696.89	467.65	837.62	2.00	2.00
	0.8	0.6	8.23	6.28	9.43	10.78	9.97	11.54	17.45	21.15	1111.20	860.11	1527.50	605.45	440.18	709.80	2.50	2.00
	0.4	0.8	5.47	4.98	6.23	7.71	7.31	8.16	32.50	24.06	1394.73	1158.59	1657.65	912.96	785.66	1095.56	3.90	2.00
	0.4	0.7	5.61	4.71	6.80	6.01	5.14	8.25	30.84	25.64	943.08	747.03	1082.52	697.21	525.48	781.73	3.80	2.20
	0.4	0.6	5.45	4.45	6.49	7.30	6.61	8.59	28.05	24.38	1107.31	877.32	1524.50	700.22	599.31	836.70	3.80	2.30
50b	0.8	0.8	9.49	8.52	10.67	9.89	9.08	11.10	16.82	24.03	1079.66	894.13	1301.74	652.69	543.22	839.49	2.10	2.00
	0.8	0.7	6.82	5.63	7.85	6.05	5.01	6.99	15.16	22.84	998.68	705.84	1369.39	612.81	464.88	765.32	2.00	2.00
	0.8	0.6	8.78	7.25	10.21	11.44	10.58	12.23	17.45	21.09	1189.01	810.71	1636.98	589.35	509.26	670.41	2.50	2.00
	0.4	0.8	6.21	5.64	6.85	8.59	7.99	9.28	29.24	24.08	1336.53	1162.64	1507.02	878.47	726.56	1049.85	3.50	2.00
	0.4	0.7	6.00	5.01	7.00	6.28	4.93	7.44	30.00	25.57	966.13	788.28	1258.57	824.13	736.84	931.50	3.70	2.20
	0.4	0.6	5.35	4.31	6.26	8.29	7.56	8.98	27.90	28.37	1080.94	969.87	1260.78	664.29	538.55	752.39	3.80	2.67

Table 4.5: Results of SMHCFP-DC ($\sigma = 0.2\mu$)

			MP						HCR		Run Times						MNH	
f			30,000			50,000			30,000	50,000	30,000			50,000			30,000	50,000
S	α	Δ	Mean	Min	Max	Mean	Min	Max			Mean	Min	Max	Mean	Min	Max		
10a	0.8	0.8	5.14	4.51	5.98	6.27	4.93	7.72	15.14	22.92	316.39	224.47	380.17	166.67	128.54	198.12	2.00	2.00
	0.8	0.7	4.31	2.64	5.51	4.59	3.15	6.65	16.51	21.85	320.22	221.65	418.19	140.39	107.17	169.97	2.30	2.00
	0.8	0.6	5.38	3.84	7.15	6.84	6.03	7.80	17.21	20.07	294.81	242.62	349.06	146.36	130.75	174.92	2.60	2.00
	0.4	0.8	3.70	2.25	4.79	5.09	4.48	5.92	28.91	25.40	306.90	227.10	426.18	218.34	177.03	263.16	3.60	2.20
	0.4	0.7	3.67	1.94	5.25	4.39	3.17	5.72	28.01	25.55	233.24	161.26	299.17	190.91	163.47	238.81	3.60	2.30
	0.4	0.6	3.81	2.44	4.84	5.66	4.57	6.37	28.19	28.60	280.40	209.87	376.81	177.20	153.67	206.35	4.00	2.80
10b	0.8	0.8	7.68	5.30	9.70	9.57	8.18	11.39	15.93	23.00	315.11	189.31	467.79	151.10	100.15	189.09	2.10	2.00
	0.8	0.7	5.74	4.56	6.99	6.88	5.28	9.02	16.57	22.02	302.15	217.19	351.95	123.56	94.22	155.63	2.30	2.00
	0.8	0.6	7.45	6.03	8.60	10.76	9.55	12.96	16.58	20.24	300.16	222.68	378.41	129.35	114.56	150.11	2.50	2.00
	0.4	0.8	5.45	4.06	6.41	7.78	6.84	8.78	29.08	25.48	292.73	202.97	385.70	187.58	114.35	249.48	3.60	2.20
	0.4	0.7	5.41	3.47	6.56	6.09	4.48	8.24	29.73	26.75	209.49	175.13	297.91	167.07	118.74	213.42	3.80	2.40
	0.4	0.6	5.09	3.41	6.22	7.83	6.36	9.56	28.26	26.58	275.35	219.01	370.24	161.93	128.67	194.63	4.00	2.60
25	0.8	0.8	7.74	6.87	9.24	9.33	7.74	10.73	16.67	22.93	833.63	657.88	1003.75	426.61	306.88	578.11	2.20	2.00
	0.8	0.7	5.85	4.67	6.66	5.89	4.15	7.43	16.50	21.73	904.33	656.74	1129.90	400.54	293.19	503.02	2.30	2.00
	0.8	0.6	7.53	5.86	9.14	10.51	9.27	11.64	17.88	20.14	767.97	594.87	942.55	443.04	319.25	523.73	2.70	2.00
	0.4	0.8	5.18	3.63	6.34	7.47	6.82	8.36	30.57	25.39	790.06	676.15	991.28	505.14	371.18	597.86	3.80	2.20
	0.4	0.7	4.81	3.45	5.33	6.28	4.71	8.32	27.94	23.36	651.60	497.31	765.89	440.56	336.93	503.38	3.60	2.10
	0.4	0.6	4.86	3.94	5.41	7.57	6.82	8.35	25.36	28.49	756.04	608.77	923.82	527.06	441.97	646.62	3.60	2.80
50a	0.8	0.8	7.96	6.89	10.19	9.80	8.27	11.30	15.22	23.05	1489.42	893.30	1985.65	840.78	657.35	1029.95	2.00	2.00
	0.8	0.7	5.77	4.78	7.10	6.61	4.62	7.80	15.77	21.88	1773.01	1138.58	2280.23	702.98	497.79	811.47	2.20	2.00
	0.8	0.6	6.86	5.69	8.30	10.65	9.10	11.71	16.39	20.11	1849.37	1559.08	2176.24	853.72	626.57	1144.41	2.50	2.00
	0.4	0.8	5.64	5.05	6.41	7.49	6.65	7.97	30.71	25.38	1415.77	1049.88	1679.27	1093.68	962.72	1359.50	3.80	2.20
	0.4	0.7	4.95	3.85	5.82	6.23	4.93	7.19	27.20	30.03	1204.24	962.10	1448.98	914.84	680.69	1073.48	3.50	2.70
	0.4	0.6	5.28	3.74	6.31	7.82	6.32	8.86	28.26	27.48	1429.13	1203.11	2039.05	899.06	687.77	1061.58	4.00	2.70
50b	0.8	0.8	8.48	7.38	9.69	10.57	8.58	11.80	15.16	23.00	1588.74	1222.48	1872.80	774.19	653.62	976.51	2.00	2.00
	0.8	0.7	6.66	5.32	8.53	6.72	5.13	9.31	16.56	21.79	1540.53	1239.94	1832.45	762.92	529.42	977.52	2.30	2.00
	0.8	0.6	8.00	6.89	9.16	11.52	10.32	13.26	16.51	20.13	1643.69	1218.17	1989.52	745.54	542.12	880.16	2.50	2.00
	0.4	0.8	5.72	4.55	6.95	8.59	7.79	9.82	29.01	26.63	1593.83	1324.29	2157.07	1009.39	833.97	1195.89	3.60	2.30
	0.4	0.7	6.14	4.07	7.05	7.22	5.90	8.74	31.34	24.56	1132.98	839.50	1555.83	823.43	616.80	1112.88	4.00	2.20
	0.4	0.6	5.12	3.05	7.40	8.35	7.09	10.24	26.73	25.46	1613.63	1164.39	2420.28	912.34	675.80	1257.45	3.80	2.50

5 CONCLUSIONS AND PERSPECTIVES

Every vehicle used for freight or passenger transportation has a distance range due to energy consumption. Even if it is possible to replenish consumed energy to extend the range, this may be economically undesirable. Especially when a vehicle carries less than truckload or is empty in its return trip, significant economic losses occur. Another fact is that it is much more beneficial to use low volume vehicles to collect/distribute small amounts of goods in the short-range, and high volume vehicles to mass transport these goods in the long-range. These types of economic incentives justify the design of hub networks. A different hub covering location problems are the subject of this thesis.

For reducing the number of constraints, mathematical models with three-indices variables are developed. Unlike most of the studies on hub covering in the literature, transportation and hub opening costs are included in the models' objectives. Moreover, even all vehicles and hubs have limited capacities in real life, and this case is also a less considered issue in hub covering problems. Three capacitated deterministic integer linear models are introduced to fill this gap in the literature.

Well-known and widely used TR and CAB data sets are chosen to validate deterministic models. Different coverage radii and interhub discount factors are used for numerical experiments. As hub setting costs are not available, different cost sets are generated for the CAB data set. The results show that it is crucial to include flow costs while designing hub networks. Another drawing from the results is capacity, whether link or hub significantly affects the model.

In real life, everything may become uncertain quickly. Handling uncertainty is a complicated problem for mathematical modeling because of computational limitations. In this thesis, three models are developed for handling uncertainty for MHCFP. The first and second models aim to handle transportation cost and flow demand uncertainties, respectively. The third model includes two uncertainties together. Two-stage stochastic mixed-integer linear optimization modeling formulation is used. The L-

Shaped algorithm is used for solving two-stage stochastic linear models. For showing the efficiency of the methodology, the models are also solved with EEV methodology.

CAB data set is used for numerical studies of stochastic optimization models. 10, 25, and 50 independent scenarios are created for examining models. For decreasing bias of created scenarios, ten independent data sets are used. Like in deterministic models, different parameters are used for numerical studies. Normal and Poisson distributions are used for the cost and demand uncertainties, respectively. Comparing the deterministic models, solving stochastic models consumes more memory and takes more time. Solving the models with an L-Shaped algorithm reflects the stochasticity of the models better than EEV's in a significant amount of scenarios.

Many future research perspectives can be conceived after this study. As it is computationally very time consuming to solve large instances of the hub covering flow problem to optimality, a natural extension is to develop efficient heuristic and/or meta-heuristic solution approaches. As the underlying network for hub covering problems is not complete ; in other words, each network node is not directly connected to all others, every selected hub-nodes set does not produce a feasible solution. This is a challenging issue to overcome in devising a solution procedure.

The addition of capacity-related constraints to a mathematical model may render it infeasible in general. As the hub covering flow problem is not feasible for all possible hub-node sets, capacity consideration can further complicate the design of problem-specific or the adaptation of generic solution approaches. Meanwhile, not only the capacities of hubs or links but also their effective usage can be a concern. This situation can be managed by altering the capacity constraints proposed in this study, or by maximizing capacity usage as a secondary objective in a multi-objective setting. Similarly, linear or nonlinear objectives involving various measures such as congestion at hubs or carbon emissions of vehicles can be formulated.

Instead of being the model's parameters, hub and link capacities can also be decision variables. Variables associated with the capacities of hubs can be continuous and bounded, binary to select among a set of capacity levels, or discrete to identify the right amount of modular capacities. Meanwhile, continuous link capacities can be increased at a cost. Another interesting line of research can be to relate hub covering radius with the hub capacity, i.e., a hub with a larger capacity can cover more nodes. It is not difficult to foresee that additional variables will negatively affect the efforts to find the optimum solution, while total cost will be significantly reduced.

The proposed stochastic programming models in this thesis have the objective to minimize the total expected network design and operating costs under demand and cost uncertainties. While this risk-neutral approach is appropriate most of the time, it can not avoid rare but extreme situations. For a real network designed as the solution of associated hub covering flow model, these extreme situations can sometimes be disastrous or even may lead to life loss. Preventing those cases requires a risk-averse measure such as conditional value-at-risk in the model's objective instead of expectation. Finally, not only uncertainties on demand and costs but also on the network structure can be taken into account. For example, some nodes can be prohibited from being a hub, or some links can be blocked in some scenarios. Minimizing the adverse effects of these uncertainties to obtain a most robust network structure can be a worthwhile research direction.

REFERENCES

Alumur, S. A., Nickel, S. and da Gama, F. S. (2012). Hub location under uncertainty, *Transportation Research Part B : Methodological* **46**(4) : 529 – 543.

Alumur, S. and Kara, B. Y. (2009). A hub covering network design problem for cargo applications in turkey, *The Journal of the Operational Research Society* **60**(10) : 1349–1359.

Beale, E. M. L. (1955). On minimizing a convex function subject to linear inequalities, *Journal of the Royal Statistical Society : Series B (Methodological)* **17**(2) : 173–184.

Beasley, J. E. (1990). Or-library.
URL: *http ://people.brunel.ac.uk/ mastjjb/jeb/orlib/files/phub4.txt*

Birge, J. and Louveaux, F. (2011). *Introduction to Stochastic Programming*, second edn, Springer New York.

Boland, N., Krishnamoorthy, M., Ernst, A. T. and Ebery, J. (2004). Preprocessing and cutting for multiple allocation hub location problems, *European Journal of Operational Research* **155**(3) : 638 – 653.

Brimberg, J., Mladenovic, N., Todosijevic, R. and Urosevic, D. (2017a). A basic variable neighborhood search heuristic for the uncapacitated multiple allocation p-hub center problem, *Optimization Letters* **11**(2) : 313–327.

Brimberg, J., Mladenovic, N., Todosijevic, R. and Urosevic, D. (2017b). General variable neighborhood search for the uncapacitated single allocation p-hub center problem, *Optimization Letters* **11**(2) : 377–388.

Bryan, D. (1998). Extensions to the hub location problem : Formulations and numerical examples, *Geographical Analysis* **30**(4) : 315–330.

Calık, H., Alumur, S. A., Kara, B. Y. and Karasan, O. E. (2009). A tabu-search based heuristic for the hub covering problem over incomplete hub networks, *Computers & Operations Research* **36**(12) : 3088 – 3096.

Campbell, A. M., Lowe, T. J. and Zhang, L. (2007). The p-hub center allocation problem, *European Journal of Operational Research* **176**(2) : 819 – 835.

Campbell, J. F. (1992). Location and allocation for distribution systems with transshipments and transportion economies of scale, *Annals of Operations Research* **40**(1) : 77–99.

Campbell, J. F. (1994). Integer programming formulations of discrete hub location problems, *European Journal of Operational Research* **72**(2) : 387 – 405.

Campbell, J. F. and O'Kelly, M. E. (2012). Twenty-five years of hub location research, *Transportation Science* **46**(2) : 153–169.

Chen, J.-F. (2007). A hybrid heuristic for the uncapacitated single allocation hub location problem, *Omega* **35**(2) : 211 – 220.

Cánovas, L., García, S. and Marín, A. (2007). Solving the uncapacitated multiple allocation hub location problem by means of a dual-ascent technique, *European Journal of Operational Research* **179**(3) : 990 – 1007.

Contreras, I., Cordeau, J.-F. and Laporte, G. (2011). Stochastic uncapacitated hub location, *European Journal of Operational Research* **212**(3) : 518 – 528.

Correia, I., Nickel, S. and da Gama, F. S. (2010). The capacitated single-allocation hub location problem revisited : A note on a classical formulation, *European Journal of Operational Research* **207**(1) : 92 – 96.

Correia, I., Nickel, S. and da Gama, F. S. (2014). Multi-product Capacitated Single-Allocation Hub Location Problems : Formulations and Inequalities, *Networks and Spatial Economics* **14**(1) : 1–25.

Correia, I., Nickel, S. and da Gama, F. S. (2017). A stochastic multi-period capacitated multiple allocation hub location problem : Formulation and inequalities, *Omega* pp. –.

Cunha, C. B. and 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.

da Graça Costa, M., Captivo, M. E. and Clímaco, J. (2008). Capacitated single allocation hub location problem—a bi-criteria approach, *Computers & Operations Research* **35**(11) : 3671 – 3695.

Dantzig, G. B. (1955). Linear programming under uncertainty, *Management Science* **1**(3-4) : 197–206.

Ebery, J. (2001). Solving large single allocation p-hub problems with two or three hubs, *European Journal of Operational Research* **128**(2) : 447 – 458.

Ebery, J., Krishnamoorthy, M., Ernst, A. and Boland, N. (2000). The capacitated multiple allocation hub location problem : Formulations and algorithms, *European Journal of Operational Research* **120**(3) : 614 – 631.

Ekonomist Online (2018). Lojistik sektörüne e-ticaret dopingi.
URL: *https ://www.ekonomist.com.tr/haberler/lojistik-sektorune-e-ticaret-dopingi.html*

Elhedhli, S. and Hu, F. X. (2005). Hub-and-spoke network design with congestion, *Computers & Operations Research* **32**(6) : 1615 – 1632.

Ernst, A. T., Hamacher, H., Jiang, H., Krishnamoorthy, M. and Woeginger, G. (2009). Uncapacitated single and multiple allocation p-hub center problems, *Computers & Operations Research* **36**(7) : 2230 – 2241.

Ernst, A. T. and Krishnamoorthy, M. (1996). Efficient algorithms for the uncapacitated single allocation p-hub median problem, *Location Science* **4**(3) : 139 – 154.

Ernst, A. T. and Krishnamoorthy, M. (1999). Solution algorithms for the capacitated single allocation hub location problem, *Annals of Operations Research* **86** : 141–159.

Gavriliouk, E. O. (2009). Aggregation in hub location problems, *Computers & Operations Research* **36**(12) : 3136 – 3142.

Hakimi, S. L. (1964). Optimum locations of switching centers and the absolute centers and medians of a graph, *Operations Research* **12**(3) : 450–459.

Hakimi, S. L. (1965). Optimum distribution of switching centers in a communication network and some related graph theoretic problems, *Operations Research* **13**(3) : 462–475.

He, Y., Wu, T., Zhang, C. and Liang, Z. (2015). An improved {MIP} heuristic for the intermodal hub location problem, *Omega* **57, Part B** : 203 – 211.

Hult, E., Jiang, H. and Ralph, D. (2013). Exact computational approaches to a stochastic uncapacitated single allocation p-hub center problem, *Computational Optimization and Applications* **59**(1) : 185–200.

Hwang, Y. H. and Lee, Y. H. (2012). Uncapacitated single allocation p-hub maximal covering problem, *Computers & Industrial Engineering* **63**(2) : 382 – 389.

Ilić, A., Urošević, D., Brimberg, J. and Mladenović, N. (2010). A general variable neighborhood search for solving the uncapacitated single allocation p-hub median problem, *European Journal of Operational Research* **206**(2) : 289 – 300.

Ishfaq, R. and Sox, C. R. (2011). Hub location–allocation in intermodal logistic networks, *European Journal of Operational Research* **210**(2) : 213 – 230.

Kara, B. Y. and Tansel, B. (2000). On the single-assignment p-hub center problem, *European Journal of Operational Research* **125**(3) : 648 – 655.

Kara, B. Y. and Tansel, B. C. (2003). The single-assignment hub covering problem : Models and linearizations, *The Journal of the Operational Research Society* **54**(1) : 59–64.

Klincewicz, J. G. (1992). Avoiding local optima in thep-hub location problem using tabu search and grasp, *Annals of Operations Research* **40**(1) : 283–302.

Klincewicz, J. G. (1996). A dual algorithm for the uncapacitated hub location problem, *Location Science* **4**(3) : 173 – 184.

Kratica, J. (2013). An electromagnetism-like metaheuristic for the uncapacitated multiple allocation p-hub median problem, *Computers & Industrial Engineering* **66**(4) : 1015 – 1024.

Kratica, J., Milanović, M., Stanimirović, Z. and Tošić, D. (2011). An evolutionary-based approach for solving a capacitated hub location problem, *Applied Soft Computing* **11**(2) : 1858 – 1866.

Labbé, M. and Yaman, H. (2004). Projecting the flow variables for hub location problems, *Networks* **44**(2) : 84–93.

Labbé, M., Yaman, H. and Gourdin, E. (2005). A branch and cut algorithm for hub location problems with single assignment, *Mathematical Programmming* **102**(2) : 371–405.

Lin, C.-C., Lin, J.-Y. and Chen, Y.-C. (2012). The capacitated p-hub median problem with integral constraints : An application to a chinese air cargo network, *Applied Mathematical Modelling* **36**(6) : 2777 – 2787.

Lium, A.-G., Crainic, T. G. and Wallace, S. W. (2009). A study of demand stochasticity in service network design, *Transportation Science* **43**(2) : 144–157.

Lowe, T. J. and Sim, T. (2013). The hub covering flow problem, *The Journal of the Operational Research Society* **64**(7) : 973–981.

Marianov, V., Serra, D. and ReVelle, C. (1999). Location of hubs in a competitive environment, *European Journal of Operational Research* **114**(2) : 363 – 371.

Marín, A. (2005a). Formulating and solving splittable capacitated multiple allocation hub location problems, *Computers & Operations Research* **32**(12) : 3093 – 3109.

Marín, A. (2005b). Uncapacitated euclidean hub location : Strengthened formulation, new facets and a relax-and-cut algorithm, *Journal of Global Optimization* **33**(3) : 393–422.

Mayer, G. and Wagner, B. (2002). Hublocator : an exact solution method for the multiple allocation hub location problem, *Computers & Operations Research* **29**(6) : 715 – 739.

Meraklı, M. and Yaman, H. (2016). Robust intermodal hub location under polyhedral demand uncertainty, *Transportation Research Part B : Methodological* **86** : 66 – 85.

Meyer, T., Ernst, A. and Krishnamoorthy, M. (2009). A 2-phase algorithm for solving the single allocation p-hub center problem, *Computers & Operations Research* **36**(12) : 3143 – 3151. New developments on hub location.

O'Kelly, M. and Bryan, D. (1998). Hub location with flow economies of scale, *Transportation Research Part B : Methodological* **32**(8) : 605 – 616.

O'Kelly, M., Bryan, D., Skorin-Kapov, D. and Skorin-Kapov, J. (1996). Hub network design with single and multiple allocation : A computational study, *Location Science* **4**(3) : 125 – 138.

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.

Pamuk, F. S. and Sepil, C. (2001). A solution to the hub center problem via a single-relocation algorithm with tabu search, *IIE Transactions* **33**(5) : 399–411.

Peker, M. and Kara, B. Y. (2015). The p-hub maximal covering problem and extensions for gradual decay functions, *Omega* **54** : 158 – 172.

Qin, Z. and Gao, Y. (2017). Uncapacitated p-hub location problem with fixed costs and uncertain flows, *Journal of Intelligent Manufacturing* **28**(3) : 705–716.

Qu, B. and Weng, K. (2009). Path relinking approach for multiple allocation hub maximal covering problem, *Computers & Mathematics with Applications* **57**(11–12) : 1890 – 1894.

Rodríguez-Martín, I. and Salazar-González, J. J. (2008). Solving a capacitated hub location problem, *European Journal of Operational Research* **184**(2) : 468 – 479.

Rothenbächer, A.-K., Drexl, M. and Irnich, S. (2016). Branch-and-price-and-cut for a service network design and hub location problem, *European Journal of Operational Research* **255**(3) : 935 – 947.

Sasaki, M., Suzuki, A. and Drezner, Z. (1999). On the selection of hub airports for an airline hub-and-spoke system, *Computers & Operations Research* **26**(14) : 1411 – 1422.

Silva, M. R. and Cunha, C. B. (2009). New simple and efficient heuristics for the uncapacitated single allocation hub location problem, *Computers & Operations Research* **36**(12) : 3152 – 3165.

Sim, T., Lowe, T. J. and Thomas, B. W. (2009). The stochastic -hub center problem with service-level constraints, *Computers & Operations Research* **36**(12) : 3166 – 3177.

Skorin-Kapov, D. and Skorin-Kapov, J. (1994). On tabu search for the location of interacting hub facilities, *European Journal of Operational Research* **73**(3) : 502 – 509.

Skorin-Kapov, D., Skorin-Kapov, J. and O'Kelly, M. (1996). Tight linear programming relaxations of uncapacitated p-hub median problems, *European Journal of Operational Research* **94**(3) : 582 – 593.

Sohn, J. and Park, S. (1997). A linear program for the two-hub location problem, *European Journal of Operational Research* **100**(3) : 617 – 622.

Sohn, J. and Park, S. (2000). The single allocation problem in the interacting three-hub network, *Networks* **35**(1) : 17–25.

Tan, P. Z. and Kara, B. Y. (2007). A hub covering model for cargo delivery systems, *Networks* **49**(1) : 28–39.

Topcuoglu, H., Corut, F., Ermis, M. and Yilmaz, G. (2005). Solving the uncapacitated hub location problem using genetic algorithms, *Computers & Operations Research* **32**(4) : 967 – 984.

Turkish Airlines (2020). Code share flights.
URL: *https ://www.turkishairlines.com/en-tr/any-questions/codeshare-flights/index.html*

U.S. Energy Information Administration (2019). Frequently asked questions.
URL: *https ://www.eia.gov/tools/faqs/faq.php*

Wagner, B. (2008). Model formulations for hub covering problems, *The Journal of the Operational Research Society* **59**(7) : 932–938.

Yaman, H. (2008). Star p-hub median problem with modular arc capacities, *Computers & Operations Research* **35**(9) : 3009 – 3019.

Yaman, H. (2009). The hierarchical hub median problem with single assignment, *Transportation Research Part B : Methodological* **43**(6) : 643 – 658.

Yang, T.-H. (2009). Stochastic air freight hub location and flight routes planning, *Applied Mathematical Modelling* **33**(12) : 4424 – 4430.

Yetis Kara, B. (2017). Turkish hub data set.
URL: *http://www.bilkent.edu.tr/{~}bkara/dataset.php*

BIOGRAPHICAL SKETCH

Nazmi ŞENER was born in Van, Turkey in 1989. He graduated from Bursa Tofaş Fen Lisesi formerly known as Bursa Ali Osman Sönmez Fen Lisesi in 2005. He completed his undergraduate studies in Industrial Engineering from Dokuz Eylül University in 2010. He attended academician training program of Higher Education Council of Turkey in same year. He completed extensive English learning program at METU in 2011. He gained MSc degree in Industrial Engineering at Galatasaray University in 2014. He is currently a research assistant in Galatasaray University and pursues PhD degree.

PUBLICATIONS

— **Sener, N**, Feyzioglu, O. ; Two Stage Stochastic Optimization Model for Multiple Allocation Hub Covering Problem, 48th International Conference on Computers Industrial Engineering (CIE48) , December 02-05, 2018, Auckland, New Zealand.

— **Sener, N**, Feyzioglu, O. ; Multiple Allocation Stochastic Hub Covering Problem with Flow Costs Uncertainty, 11th International Conference on Industrial Engineering and Industrial Management (CIO'17), July, 05-06, 2017, Valencia, Spain.

— **Sener, N**, Feyzioglu, O. ; A Risk Averse Stochastic Programming Approach for the Appointment Scheduling Problem, Proceedings of GOR/DASIG 2013, September 23-24, 2013, Hamburg, Germany.

— Turan, H. H., **Sener, N.** ; An Analysis of Profit Maximization Problem of Telecommunications Intermediaries under Uncertainty, The International IIE Conference, June 26-28, 2013, Istanbul, Turkey.

— Turan, H. H., **Sener, N.**; A Possibilistic Mathematical Programming Approach for Backbone Selection and Demand Allocation Problem in Telecom Networks, Proceedings of the 10th International FLINS Conference on Uncertainty Modelling in Knowledge Engineering and Decision Making (FLINS 2012), August 26-29, 2012, Istanbul, Turkey.

— **Sener, N.**, Turan, H. H.; An Integration of TOPSIS Method and Interval Type II Fuzzy Sets to Evaluate Municipality Investment Alternatives : The Case of Istanbul Municipality, The 25th European Conference on Operational Research (EURO XXV), July 8-11, 2012, Vilnius, Lithuania.