The Design of Capacitated Intermodal Hub Networks with Different Vehicle Types

Elif Zeynep Serper[a], Sibel A. Alumur[b]

ezserper@gmail.com, sibel.alumur@uwaterloo.ca

[a] Department of Industrial Engineering, TOBB University of Economics and Technology, Ankara, Turkey
[b] Department of Management Sciences, University of Waterloo, Waterloo, Ontario, Canada

Abstract

In this study, we allow using alternative transportation modes and different types of vehicles in the hub networks to be designed. The aim of the problem is to determine the locations and capacities of hubs, which transportation modes to serve at hubs, allocation of non-hub nodes to hubs, and the number of vehicles of each type to operate on the hub network to route the demand between origin-destination pairs with minimum total cost. Total cost includes fixed costs of establishing hubs with different capacities, purchasing and operational costs of vehicles, transportation costs, and material handling costs. A mixed-integer programming model is developed and a variable neighborhood search algorithm is proposed for the solution of this problem. The heuristic algorithm is tested on instances from the Turkish network and CAB data set. Extensive computational analyses are conducted in order to observe the effects of changes in various problem parameters on the resulting hub networks.

Keywords: Hub location, hub network design, intermodal transportation networks.

1 Introduction

In today's competitive environment, most firms use alternative transportation modes and different types of vehicles within their hub networks. The main reason for using different vehicle types in hub networks is to achieve economies of scale and decrease unit transportation costs by employing vehicles with larger capacities. In this study, we focus on minimizing total cost to efficiently locate hub facilities and design intermodal hub networks.

The motivation of this study arises from the operational characteristics of small parcel delivery networks. Such networks operate on a hub-and-spoke network structure. Hubs are crossdock

1

terminals consolidating flow in which the main operations are unloading, sorting, and loading. Spokes or demand nodes, on the other hand, are typically branch offices. In real-life parcel delivery networks, each branch office is customarily allocated to a single hub mainly due to ease of management. Allocation connections carry low volume traffic and distances are relatively short; therefore, only ground transportation is employed on these connections. Moreover, usually a single vehicle type is used to transfer the flow between branch offices and hubs. Customarily, the flow is collected from hubs in the mornings and delivered to hubs at the end of a day. During the day, the vehicles on these allocation connections are used to deliver and pick-up parcels from customers. Alternative modes of transportation and different types of vehicles with larger capacities are used, on the other hand, while transporting flow between hub facilities. Each firm operates with its own fleet. Additional vehicles can be rented especially to address peak demand periods.

Operations in small parcel delivery are planned daily and the amount of flow to be transported refers to daily demand. The capacities of hubs in parcel delivery networks are usually expressed in terms of the total number of vehicles of each type that can be handled within a day. For example, the number of docks limits the number of trucks that can be loaded and unloaded at a hub during a day. Similarly, the number of airplanes that can be handled within a day is limited.

Various costs are involved in building and operating this hub network including the costs of establishing hubs with different capacities, purchasing, operational and transportation costs of vehicles. In addition, there is material handling costs associated with unloading, sorting, and loading operations at hubs. Material handling cost is dependent on the mode of transportation. For example, the cost of loading and unloading an airplane is different than the cost of loading and unloading a truck. In addition, material handling cost may also depend on the location of a hub as some components of cost, such as minimum wage and electricity rates, can vary from location to location.

In this paper, we model and solve the intermodal hub network design problem with the specifications described above. We develop a mixed-integer programming formulation and a variable neighborhood search heuristic. We solve the model and test the algorithm on instances from the Turkish network and CAB data set (Beasley [6]).

Hub location problems deal with choosing the sites of hub facilities and allocating demand nodes to hubs so as to effectively route the traffic between origin-destination pairs. The reader may refer to Alumur and Kara [1], Campbell and O'Kelly [11], Farahani et al. [15], and Contreras [12] for an overview of hub location studies in the literature.

In this study, in addition to the hub location and allocation decisions, we also consider decisions regarding the design of the inter-hub network; that is, we do not assume a fully interconnected hub network. Examples of hub location studies which incorporate decisions on inter-hub network design include Nickel et al. [24], Campbell et al. [10], Yoon and Current [33], Alumur et al. [2], Contreras et al. [13], Martins de Sa et al. [22], and Alumur et al. [4]. In all of these studies, only one transportation mode is allowed within the hub network to be designed.

The choice for mode of transportation within a hub network is initially discussed by O'Kelly and Lao [28]. Different hub location models are used in the literature to determine locations of hubs in intermodal transportation networks. Examples of such studies are Racunica and Wynter [30] using an uncapacitated multiple allocation hub location model, Limbourg and Jourquin [20] and Ishfaq and Sox [18] both using a p-hub median model. In these studies, inter-hub network is assumed to be complete with a direct link between every hub pair.

The most relevant studies to our problem are the ones jointly considering the location of hubs and the design of intermodal hub networks. Meng and Wang [23] proposed a hybrid genetic algorithm to solve the intermodal hub network design problem for multi-type container transportation with multiple stakeholders including the network planner, carriers, hub operators, and intermodal operators. In our problem, we have a single stakeholder and we solve a single company's hub network design problem.

Alumur et al. [3] developed a mixed-integer programing formulation for multimodal hub location and hub network design problem with different service options. In this study, only one type of service and one transportation mode is allowed between each pair of hubs. In a subsequent study, Alumur et al. [5] modeled and solved a hub location problem designing a hierarchical intermodal network while ensuring time-definite deliveries. In the current study, we do not impose any specific hub network structure or delivery time restrictions.

Most studies on hub location allow using a single transportation mode and a single vehicle type on the hub network to be designed. The studies considering the design of intermodal hub networks, on the other hand, do not model the possibility of employing different types of vehicles with different capacities. To the best of the authors' knowledge, there are not any hub location studies in the literature determining the optimal number of different types of vehicles to operate on an intermodal hub network. Choice for mode of transportation and number of vehicles to be handled at hubs affects both hub location and capacity decisions. Hence, such decisions should be considered simultaneously in designing hub networks.

A related field to our study is service network design. Service network design is concerned with tactical operations and decisions like selection and scheduling of services, specification of

3

terminal operations, and routing of freight using already existing hub facilities. Interested reader may refer to Crainic [14] and Wieberneit [31] for reviews in this area. In this study, we focus on the strategic decisions and do not address any tactical decisions such as the frequency or scheduling of the vehicles.

In most hub location studies, economies of scale is modeled by using a constant discount factor (usually referred as α) to reflect the discounted cost of flow between hubs. However, as pointed out by many studies in the literature (e.g., O'Kelly and Bryan [26], Kimms [19], and Campbell [9]) using flow-independent discounts between hubs results in a misapplication of economies of scale. O'Kelly and Bryan [26] and Bryan [7] suggested using a nonlinear concave cost function to correctly model economies of scale. Camargo et al. [8] proposed a stronger formulation and a Benders decomposition algorithm for the function introduced in O'Kelly and Bryan [26]. Podnar et al. [29] used a threshold-based discounting on links. Campbell et al. [10] introduced hub arc location models where the models locate hub arcs with reduced unit flow costs rather than locating hub facilities. O'Kelly et al. [27] added fixed costs for arcs to better model transportation operations. By the use of fixed costs, the unit cost per flow on an arc decreases as more flow is routed on that arc.

Kimms [19] suggested alternative formulations to model flow-dependent economies of scale in hub networks. Alternatives were introduced using the classical single allocation hub location problem with fixed costs and the effects were illustrated on a seven node numerical example. One approach introduced in Kimms [19] is to determine the optimal number of different types of vehicles to employ and consider transportation cost per vehicle rather than a fixed transportation cost per flow. In this study, we adopt a similar flow-dependent method to correctly model economies of scale. Unlike Kimms [19], we consider hubs with different transportation modes and capacities, and we do not assume a fully interconnected hub network.

The objective function of our model includes purchasing, operational, and transportation costs of different types of vehicles. Purchasing cost is the cost of acquiring a vehicle. Operational cost is incurred for each trip and vehicle, and involves crew cost, maintenance cost, service fee, etc. Purchasing and operational costs constitute a fixed cost in the objective function for each vehicle to be used on the hub network. This is very similar to the idea of using fixed costs for arcs employed, for example, in Alumur et al. [2] and O'Kelly et al. [27]. We additionally consider different vehicle types and transportation modes in our problem. Each vehicle type has a different fuel consumption rate and, hence, a different transportation cost per distance traveled. In order to determine the optimal number of vehicles to employ on the hub network, our model readily considers the trade-off between vehicle costs and capacities; that is, using few

4

large-capacity vehicles versus many small-capacity vehicles.

Determining optimal number of different vehicles to employ is analogous to determining optimal number of different link types to install in network design. Such problems are referred as network loading problems in the literature (Magnanti et al. [21]). The major difference of our problem from network loading problems, apart from hub location and capacity decisions, is that there are no transportation (routing) costs in network loading, there is only fixed costs of installing different links. In our problem, even though flows on the links of the network are known, the vehicle allocation subproblem is not straight forward; it is another combinatorial optimization problem. Whereas in network loading, once the flows on the network are fixed, the decision on the type and number of links to be installed on each connection is trivial. Our problem is more difficult in the sense that there is an additional dimension of transportation costs that needs to be considered for each vehicle on each connection as well as the capacities of hubs.

The outline of this paper is as follows. In the next section, a mathematical formulation for the capacitated intermodal hub network design problem is introduced. In the third section, we present a heuristic algorithm developed for the solution of this problem. The fourth section presents results from the application of the heuristic on the Turkish network and CAB data set. This section also includes a statistical analysis and a method for determining a lower bound. The fifth section contains sensitivity analyses using instances from the Turkish network. The paper ends with some concluding remarks presented in the last section.

2   Mathematical Formulation

In this section, we present a mathematical formulation for the single allocation capacitated intermodal hub network design problem. The aim of the problem is to decide on the locations and capacities of hubs, which transportation modes to serve at these hubs, the allocation of non-hub nodes to hubs, and the number of vehicles of each type to purchase and operate between each pair of hubs. The objective is to minimize total cost which includes the purchasing, operational, and transportation costs of vehicles, fixed costs of establishing hubs with different capacities, and material handling costs.

For the mathematical model, we are given a node set N consisting of n demand nodes and a potential hub set H such that $H \subseteq N$ with h nodes. Hubs may serve different transportation modes from the set M and for each transportation mode they can be established with different capacities from the set Q. Different hub capacities may refer for example to the number of docks to be established for each transportation mode.

5

There are different types of vehicles that can be purchased to serve on the hub network. V_m represents the set of available vehicle types that use transportation mode $m \in M$. The set of all vehicles that can be used is denoted by V such that $V = \bigcup_{m \in M} V_m$. It is assumed that each vehicle operates on a single connection.

Unit transportation costs are dependent on the vehicle type. Different transportation modes are allowed to be used only within the hub network; that is, only for traveling between hubs. For the allocation connections, only one type of vehicle is employed. The rest of the parameters required for the mathematical model are listed below.

w_{ij}	Amount of flow originated at node $i \in N$ destined to node $j \in N$.
c_{ij}	Unit cost of transportation from node $i \in N$ to hub $j \in H$ on allocation connections.
c_{ij}^v	Unit cost of transportation from hub $i \in H$ to hub $j \in H$ using vehicle of type $v \in V$.
fc_j	Fixed cost of establishing a hub at node $j \in H$.
kc_{jq}^m	Cost of installing capacity $q \in Q$ for transportation mode $m \in M$ at a hub established at node $j \in H$.
k_q^v	Maximum number of vehicles of type $v \in V$ that a hub with capacity $q \in Q$ can handle.
pc^v	Purchasing cost of a vehicle of type $v \in V$.
oc^v	Operational cost of a vehicle of type $v \in V$.
u^v	Capacity of a vehicle of type $v \in V$.
mhc_j	Material handling cost of one unit of flow arriving from a demand node to hub $j \in H$.
mhc_j^m	Material handling cost of one unit of flow arriving with transportation mode $m \in M$ to hub $j \in H$ from another hub.
a_i	Number of vehicles required to transport the flow originated at node $i \in N$.
b_i	Number of vehicles required to transport the flow destined to node $i \in N$.

Since each node is allocated to a single hub and only one type of vehicle operates on the allocation connections, a_i and b_i can be calculated as follows.

$$a_i = \left\lceil \frac{O_i}{u} \right\rceil, \qquad b_i = \left\lceil \frac{D_i}{u} \right\rceil$$

where $O_i = \sum_{j \in N} w_{ij}$, $D_i = \sum_{j \in N} w_{ji}$, and u represents the capacity of the vehicle that serves on the allocation links.

It is assumed that all cost values are scaled considering the planning horizon. For example, if the flows reflect the daily amounts, then all the cost values should correspond to daily values.

The decision variables of the mathematical model are

$$
x_{ij} = \begin{cases} 1, & \text{if node } i \in N \text{ is allocated to a hub at node } j \in H, \\ 0, & \text{otherwise.} \end{cases}
$$

($x_{jj} = 1$ indicates that a hub is established at node $j \in H$.)

$$
y_{jq}^m = \begin{cases} 1, & \text{if capacity level } q \in Q \text{ is installed at hub } j \in H \text{ for transportation mode } m \in M, \\ 0, & \text{otherwise.} \end{cases}
$$

$f_{ijk}^v = $ Amount of flow originated at node $k \in N$ and transported from hub $i \in H$ to hub $j \in H$ using vehicle of type $v \in V$.

$z_{ij}^v = $ Number of vehicles of type $v \in V$ used to travel from hub $i \in H$ to hub $j \in H$.

The capacitated intermodal hub network design problem is modeled as

$$
\begin{aligned}
Min \quad & \sum_{i \in N} \sum_{j \in H} c_{ij} a_i x_{ij} + \sum_{i \in N} \sum_{j \in H} c_{ji} b_i x_{ij} + \sum_{i \in H} \sum_{j \in H} \sum_{v \in V} (pc^v + oc^v + c_{ij}^v) z_{ij}^v + \\
& \sum_{j \in H} fc_j x_{jj} + \sum_{j \in H} \sum_{q \in Q} \sum_{m \in M} kc_{jq}^m y_{jq}^m + \sum_{i \in N} \sum_{j \in H} mhc_j O_i x_{ij} + \\
& \sum_{i \in N} \sum_{j \in H} \sum_{k \in N} \sum_{m \in M} \sum_{v \in V_m} mhc_j^m f_{ijk}^v
\end{aligned}
\tag{1}
$$

$$
s.t. \quad \sum_{j \in H} x_{ij} = 1 \qquad \forall i \in N \tag{2}
$$

$$
x_{ij} \le x_{jj} \qquad \forall i \in N, \, j \in H \tag{3}
$$

$$
\sum_{q \in Q} y_{jq}^m \le x_{jj} \qquad \forall j \in H, \, m \in M \tag{4}
$$

$$
\sum_{j \in H, \, j \ne i} z_{ij}^v \le \sum_{q \in Q} k_q^v y_{iq}^m \qquad \forall i \in H, \, m \in M, \, v \in V_m \tag{5}
$$

$$
\sum_{i \in H, \, i \ne j} z_{ij}^v \le \sum_{q \in Q} k_q^v y_{jq}^m \qquad \forall j \in H, \, m \in M, \, v \in V_m \tag{6}
$$

$$
\sum_{j \in H, \, j \ne i} \sum_{v \in V} f_{ijk}^v - \sum_{j \in H, \, j \ne i} \sum_{v \in V} f_{jik}^v = O_k x_{ki} - \sum_{l \in N} w_{kl} x_{li} \qquad \forall i \in H, \, k \in N \tag{7}
$$

$$
\sum_{k \in N} f_{ijk}^v \le u^v z_{ij}^v \qquad \forall i,j \in H, \, i \ne j, \, v \in V \tag{8}
$$

7

$$f_{ijk}^v \geq 0 \qquad \forall\, i,j \in H,\ i \neq j,\ k \in N,\ v \in V \tag{9}$$

$$z_{ij}^v \geq 0 \text{ and integer} \qquad \forall\, i,j \in H,\ i \neq j,\ v \in V \tag{10}$$

$$x_{ij} \in \{0,1\} \qquad \forall\, i \in N,\ j \in H \tag{11}$$

$$y_{jq}^m \in \{0,1\} \qquad \forall\, i \in H,\ q \in Q,\ m \in M \tag{12}$$

In the objective function (1), the first term calculates the transportation cost from demand nodes to hubs and the second term calculates the transportation cost from hubs to demand nodes. The third term accounts for the purchasing, operational, and transportation costs of different types of vehicles that operate between hubs. The fourth term calculates the fixed cost of establishing hubs, the fifth term, on the other hand, sums the cost of establishing different capacity levels at hubs. The last two terms calculate the material handling costs at hubs: firstly for the flow arriving from demand nodes to hubs, secondly for the flow arriving at hubs from other hubs using different transportation modes.

By constraints (2) and (3) every demand node is allocated to a single hub. These are the classical single allocation constraints.

It is possible for a hub to serve different modes of transportation. Constraint (4) ensures that a single capacity level is selected and installed for each mode of transportation to be served at a hub.

Constraints (5) and (6) are the capacity constraints on the number of vehicles. By these constraints, the total number of vehicles that can either arrive at or depart from a hub cannot be greater than the maximum number of vehicles that can be handled at that hub. This constraint can be interpreted as an upper bound on the number of vehicles that can be served due to a limit on the number of available docks at a hub.

Constraint (7) is the flow balance constraint. The amount of flow to be routed on the hub network is determined via this constraint. Constraint (8) calculates the number of vehicles of each type to serve on the hub network from one hub to another.

The rest of the constraints of the model (9)–(12) are the non-negativity constraints and the constraints defining the integer and binary variables.

3 A Variable Neighborhood Search Algorithm

Preliminary analysis with our mathematical formulation showed that it is not possible to solve the model optimally in reasonable times for realistically sized instances using standard commercial solvers. Thus, we developed a heuristic algorithm for our problem. In this section, we present the heuristic algorithm that we developed.

For the heuristic, we decompose the problem into two subproblems: i) hub network design, and ii) vehicle assignment. The first subproblem is comprised of determining the locations of hubs, allocation of non-hub nodes to hubs, and the design of the inter-hub network. The second subproblem, on the other hand, determines the types and number of vehicles to purchase and operate on the hub network, as well as the capacities of the hubs to be established. Both parts of the heuristic consists of a construction and an improvement phase.

The hub network design part of the heuristic is a variable neighborhood search (VNS) algorithm. Interested reader may refer to Hansen et al. [17] for an excellent review on the methods and applications of VNS. According to the classification presented in Hansen et al. [17], the first part of our heuristic falls under the category *Reduced VNS* with some variations. We provide more details of the algorithm in the sequel.

The algorithm starts with a greedy construction phase to generate an initial solution. Then, different neighborhood structures are applied to this initial solution. Six neighborhood structures are considered for the hub network design part of the algorithm. Let $N_i, (i = 1, ..., 6)$ denote the neighborhood structures, and $N_i(n)$ denote the set of solutions in the ith neighborhood of n. For each neighborhood, we generate a point, say n_i, at random (using the uniform random distribution) from the ith neighborhood of n; i.e., $n_i \in N_i(n)$. Vehicle assignment phase of the algorithm is then applied one-by-one to all six neighbors of n: $n_1, ..., n_6$, and the objective function value of each neighbor, $f(n_i)$, is evaluated. A move is made to the best neighbor among the six, even though it does not improve the current objective function value. This avoids getting trapped in a local optima. The algorithm runs until a prescribed number of solutions are generated. The flow chart of the main algorithm (Figure 1) and its pseudocode (Algorithm 1) are presented below. The details of each step of the algorithm is provided in the following sub-sections.

3.1 Initial Solution

The heuristic starts with a greedy construction algorithm to generate an initial solution. There is an additional parameter, p, required for this phase of the algorithm which determines the number of hubs to be established in the initial solution. The number of hubs in the solutions are then changed through the application of different neighborhood structures.

In generating the initial solution, first, $\frac{fc_j}{O_j + D_j}$ value is calculated for each potential hub location; i.e., for all $j \in H$. These values are then sorted in ascending order. The first p locations on this ascending list are selected as hubs.

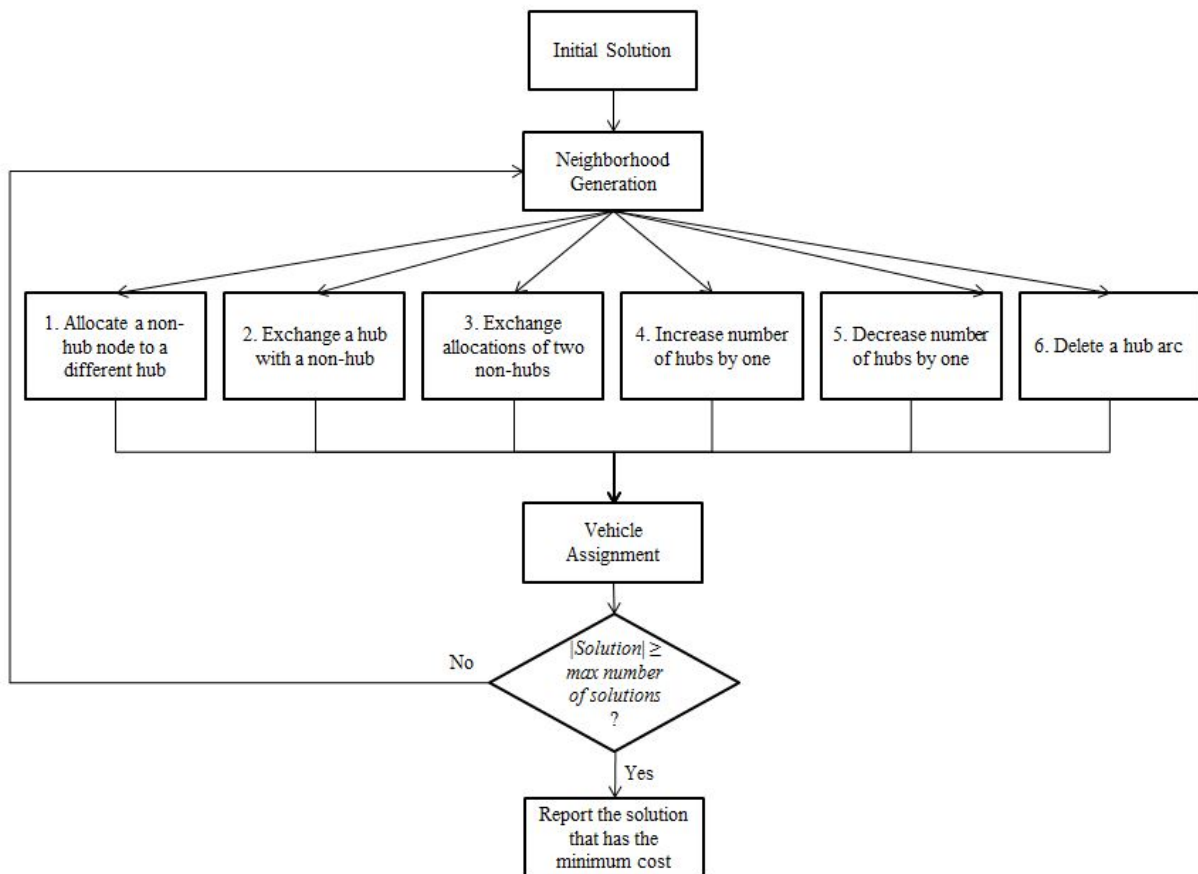Every demand node is allocated to its nearest hub and the hub network is assumed to be

9

Figure 1: Flow chart of the algorithm.

Algorithm 1 Main Algorithm

Input: All parameters required for the model, p, *max number of solutions*

Output: Hub locations and capacities (y_j^{mq}), allocation of non-hub nodes to hubs (x_{ij}), flows on each vehicle type on the hub network (f_{ijk}^v), number of vehicles operating on the hub network (z_{ij}^v)

1: $Solution \leftarrow \emptyset$

2: $n \leftarrow$ Initial Solution

3: **repeat**

4:     $Solution \leftarrow Solution \cup \{n\}$

5:     **for** $i = 1$ to 6 **do**

6:         Generate a point $n_i \in N_i(n)$ at random

7:         Apply Vehicle Assignment to n_i and calculate total cost, $f(n_i)$

8:     **end for**

9:     $i^* \leftarrow \arg\min_{i=1,\ldots,6} f(n_i)$

10:     $n \leftarrow n_{i^*}$

11: **until** $|Solution| \geq$ *max number of solutions*

12: Report the solution that has the minimum total cost in $Solution$

---

complete in the initial solution. Flow between hubs are calculated on this complete hub network structure where each flow originated at a hub is sent on the direct link to the destination hub. Incomplete hub networks are allowed in the succeeding steps of the algorithm. Algorithm 2 provides the pseudocode of the initial solution generation algorithm.

Algorithm 2 Initial Solution

Input: N, H, p, fc_j, O_j, D_j

Output: Hub locations (x_{jj}), allocation of non-hub nodes to hubs (x_{ij}), flows on the hub network (f_{ijk})

1: Calculate $\frac{fc_j}{O_j + D_j}$ for all $j \in H$

2: Sort in ascending order, pick the first p as hubs

3: Assign each non-hub node to its nearest hub

4: Calculate flow from each hub to all other hubs using a complete hub network structure

---

3.2 Neighborhood Generation

We use the mutation operators used in the genetic algorithm proposed by Gomes et al. [16] for generating neighbors. We included an additional procedure for allowing incomplete hub

networks in the solutions. In total, we use six different neighborhood structures. The six different neighborhood structures that are adopted by the algorithm are demonstrated in Figure 2 on an eight-node example.

(a) Initial solution

(b) Allocate a non-hub to a different hub

(c) Exchange a hub with a non-hub

(d) Exchange allocations of two non-hubs

(e) Increase number of hubs by one

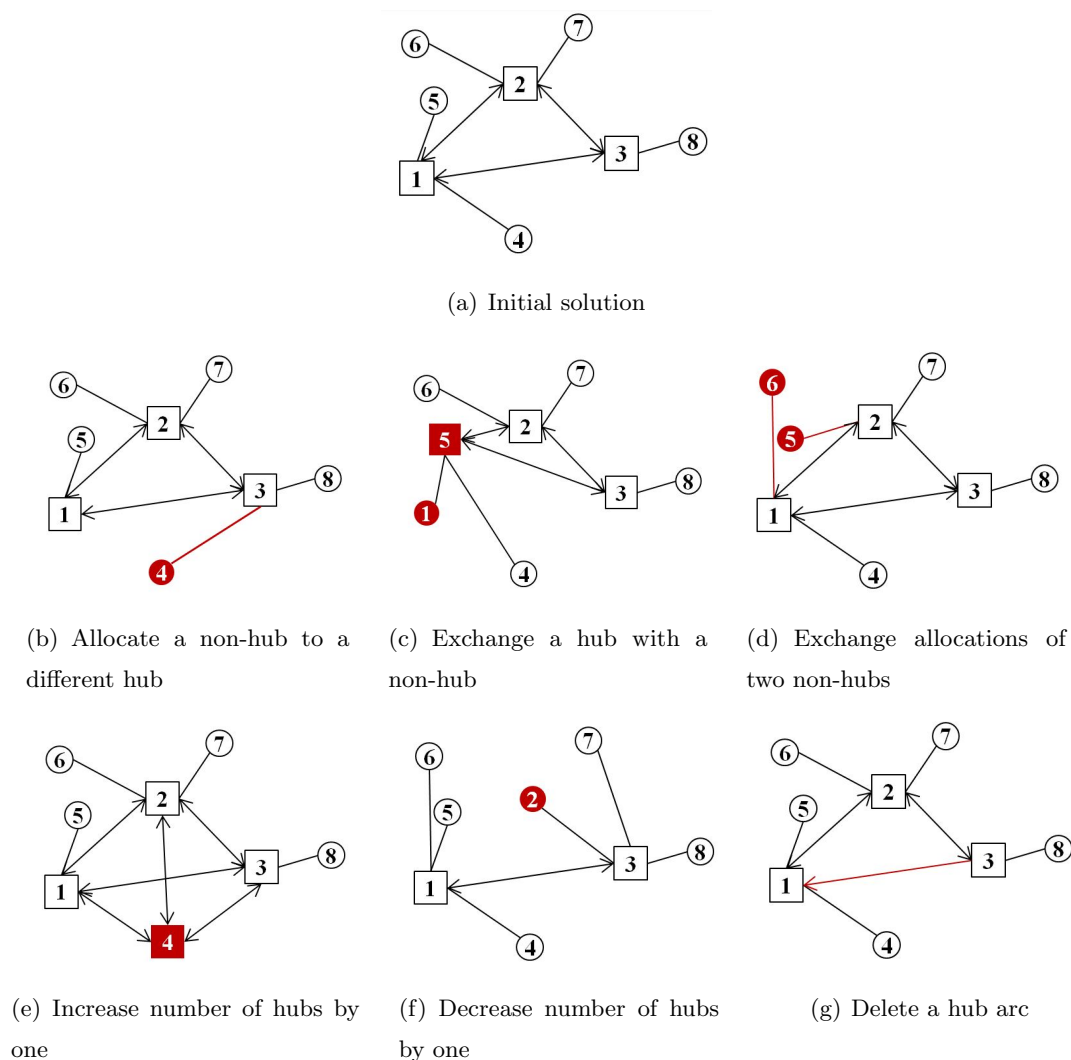(f) Decrease number of hubs by one

(g) Delete a hub arc

Figure 2: Neighborhood structures.

Figure 2(a) represents the initial solution to which different neighborhood structures are applied. In Figure 2(b), non-hub node 4 is allocated to a different hub. In Figure 2(c), hub node 1 and non-hub node 5 are interchanged; i.e., hub node 1 becomes a non-hub and non-hub node 5 becomes a hub. In Figure 2(d), non-hub nodes 5 and 6 exchange the hubs that they are allocated to. Figure 2(e) represents the insertion of a new hub node: non-hub node 4 becomes a hub and the total number of hubs is increased by one. Figure 2(f) shows the removal of a hub, hub node 2 becomes a non-hub and the total number of hubs is decreased by one. Finally,

Figure 2(g) demonstrates the deletion of a hub arc from the hub network. The hub arc directed from hub 1 to hub 3 is deleted.

In order to generate neighbors of a solution, a point is selected at random by using a discrete uniform distribution from each of the neighborhoods. Each neighborhood structure is taken into consideration for every solution. In order to avoid cycling, if an already generated neighbor is found, then another neighbor is searched using the same structure until a given maximum number of iterations is reached. If a new neighbor cannot be found within this limit, then this neighborhood structure is not evaluated for that particular solution.

The hub network should be designed in a way that it is possible to send flow between any pair of demand nodes. So, we need to ensure that the hub network is strongly connected even after the removal of a hub arc. For this purpose, we route all the flow on the network on their shortest paths and initially pick the hub arc that has the least amount of flow routed on it for deletion. When the arc is removed from the hub network, total flow on that arc is routed on the shortest path containing the remaining arcs of the hub network. A hub arc on the path of a re-routed flow is not considered as a candidate for further removal. This way, we ensure the connectivity of the hub network and the feasibility of the solution. This neighborhood structure is applied only if the total number of hubs is greater than or equal to three.

Each neighborhood structure is first applied to the initial solution and six distinct neighbors of the initial solution are generated. Vehicle assignment phase of the algorithm is then applied to each neighbor and objective function value of the six neighbors are calculated. The algorithm moves to the neighbor with the best objective function value among the six. At this step, a move to a worse solution is allowed. Next, six neighbors of the new solution is generated and evaluated, and the algorithm continues in the same manner. Each solution is stored in the *Solution* set. The solution with the best objective function value is reported at the end of the algorithm.

3.3   Vehicle Assignment

Vehicle assignment phase of the algorithm determines the transportation modes to be served at hubs as well as the number of vehicles of each type to serve on the hub network. Since the capacity of a hub is expressed as the maximum number of vehicles of each type to be handled at a hub, the capacity of the hubs for each transportation mode is also determined through the vehicle assignment phase of the heuristic algorithm.

Vehicle assignment is a local search algorithm. It consists of a greedy construction phase followed by an improvement phase.

In the construction phase, the demand of the origin-destination pairs are initially routed on their shortest paths on the available hub network. Then, vehicle assignments are considered for each hub arc starting from the one which has the highest amount of flow. Each arc is evaluated one by one considering the purchasing, operational, and transportation costs of vehicles on that particular connection. Flow on an arc is loaded onto vehicles starting from the least cost vehicle corresponding to that distance and flow. Both the fixed and variable costs are considered at this stage. Flow is allocated to the best vehicle type for a particular connection considering the available capacities. The capacities of the hubs on each end of a hub arc are checked before assigning any vehicles. No new vehicle is assigned to a hub arc for a given transportation mode, if the maximum available capacity of a hub is to be exceeded. If the next vehicle to be assigned to a hub arc exceeds the maximum available capacity of a hub incident to it, then the second best vehicle type is considered for the remaining flow. The flow is allocated to vehicles considering the costs incurred in the objective function, until all the flow on an arc is allocated. If no feasible solution can be found at this stage due to available hub capacities, then this neighbor is discarded from further evaluation.

Once all the flow on the hub network is assigned to vehicles, hub capacities are set to the minimum feasible capacity level for each transportation mode. If hubs were uncapacitated, one could have found the optimal vehicle allocations with a greedy type algorithm for given flows on the hub network. As hubs are capacitated, with a limit on the number of vehicles that can be handled, it is not possible to allocate flow on hub arcs to vehicles optimally by using the greedy approach described above. The order of hub arcs to be considered for vehicle allocation may affect the quality of the resulting solution. In order to obtain better quality solutions, we added an improvement phase to the vehicle assignment part of the algorithm.

The improvement phase of the vehicle assignment algorithm is a random local search with a limit on the number of iterations. A 2-opt type neighborhood structure is considered for this local search. Two distinct hub arcs that are incident to the same hub, which is used at its capacity, are selected randomly. The idea is to interchange the vehicle assignments on these hub arcs. Since the capacity of a hub is defined as the number of vehicles of each type that can be handled at the hub, there is a certain type of vehicle which is used up to its maximum number for a hub used at its capacity. The flow on that specific vehicle type is loaded onto the next best vehicle type for a hub arc, to make room for that type of vehicle on the other hub arc. The interchange is considered one vehicle at a time for the vehicle type limiting the capacity. Note that the removal of a single vehicle may lead to the addition of more than one vehicle of another type as vehicles are of different capacities. Similarly, the addition of one vehicle may result

in the removal of more than one vehicle of another type. For every exchange, the algorithm calculates the necessary number of vehicles to be removed and added. When assigning new vehicles to hub arcs, the capacities of the hubs on each end of the hub arcs are checked once again to ensure the feasibility of the solution. A move is made to a neighbor in this phase only if it improves the current solution. The neighborhood is searched randomly until a prescribed number of maximum iterations.

The algorithm is implemented by using Eclipse Java EE IDE. Extensive computational analysis with the heuristic algorithm on two data sets from the literature are presented in the next section.

4   Application of the Heuristic Algorithm

We tested our heuristic algorithm on the Turkish network and the well-known CAB data set introduced by O'Kelly [25]. The Turkish network contains 81 and CAB data set contains 25 demand nodes. All of the demand nodes were included in the potential hub set ($|H| = |N| = n$). Distance and flow data regarding both of these data sets are readily available in OR Library (Beasley [6]). While comparing our heuristic algorithm with the optimal solutions, we used a smaller version of the Turkish network containing 16 nodes (Yaman et al. [32]).

We conducted interviews with three of the largest cargo companies operating in Turkey in order to obtain data regarding different types of vehicles to operate on the hub network. Only ground and air transportation are employed for small parcel delivery in Turkey. Commonly, four basic types of vehicles are employed. Vans are operated on the allocation connections to carry flow to be transported between demand nodes and hubs. Airplanes, trailers, and trucks on the other hand, are operated on the inter-hub network to carry flow to be transported in-between hubs. The capacities are 3.5, 15, 25, and 200 tons, and the fuel costs are 0.5, 1.0, 1.2, and 6.4 TL/km, for a van, truck, trailer, and airplane, respectively. The total of purchasing and operational costs were taken as 5,000, 364, and 309 TL for airplanes, trailers, and trucks, respectively. As there is no real data for different types of vehicles to be operated on the CAB data set, we used the same set of vehicles with corresponding parameter values.

We tested different values ranging between 50,000 and 150,000 TL for the fixed cost of establishing hubs. We did not differentiate fixed costs for different potential hub locations; i.e., we took $fc_j = fc$ for all $j \in H$. We again used the same values for the CAB data.

For both of the data sets, two different capacities were made available for each transportation mode depending on the maximum number of vehicles that can be served at a hub. These capacities are referred as small and large. For the Turkish network, we additionally generated two

sets of capacities, referred as tight and loose, each set containing again two different capacities for each transportation mode.

Optimal solutions were obtained by solving the mathematical model using CPLEX 12.4. As there is randomness involved in the heuristic algorithm, we took five different runs for each problem instance. For each instance, we report the average results obtained at the end of these five runs. All experiments were run on a HP Z600 workstation with a 2xIntel Xeon 2.40 GHz processor and 48 GB RAM.

We initially tested our heuristic on the 16-node Turkish network. We experimented different parameter settings for the heuristic. For the *max number of solutions*, we tested generating 6,000 and 12,000 solutions. For the number of hubs to be established in the initial solution (p), we tested three different values: $\sqrt{n} - 2$, $\sqrt{n}$, and $\sqrt{n} + 2$, where n is the number of nodes in the network. The results obtained with the loose and tight capacity sets are presented in Table 1 and Table 2, respectively.

The first columns of Tables 1 and 2 list the fixed cost values for establishing hubs. Second and third columns present the results of the optimal solution obtained by using CPLEX. Remaining columns of the tables present results obtained with the heuristic algorithm. For each parameter setting with the heuristic algorithm, we report the solution time in seconds as well as the percent gap from the optimal solution. There are a total of eleven instances in each of the tables. Each instance was run five times. The reported values in Tables 1 and 2 for each instance is the average value obtained at the end of these five runs. The last row of both of the tables report the average solution times and gaps obtained by the heuristic algorithm corresponding to each column.

As expected, percent gaps from the optimal solution decreases when the algorithm searches for more solutions. When the number of solutions generated by the algorithm was doubled from 6,000 to 12,000, the average CPU time requirement of the heuristic went from approximately 7 to 59 seconds with the loose, and from 9 to 39 seconds with the tight capacity set. However, the improvement in the quality of the solutions was not as much. Considering all the instances presented in Tables 1 and 2, the average optimality gaps were 0.75% and 0.63%, for generating 6,000 and 12,000 solutions, respectively.

Observe from both of the tables that the number of hubs in the initial solution affects both the solution time and quality. The optimality gaps and solution times tend to increase when the heuristic starts with an initial solution containing more hubs. The worst performance was obtained when $p = \sqrt{n} + 2$. The performances with $p = \sqrt{n} - 2$ and $p = \sqrt{n}$, on the other hand, were comparable. Especially with some higher fixed cost values, there are instances when

16

Table 1: 16-node Turkish network solutions with the loose hub capacity set

	Optimal solution			Heuristic algorithm													
				Generating 6,000 solutions						Generating 12,000 solutions							
				$p = \sqrt{n} - 2$		$p = \sqrt{n}$		$p = \sqrt{n} + 2$		$p = \sqrt{n} - 2$		$p = \sqrt{n}$		$p = \sqrt{n} + 2$	
Fixed cost	Objective function value	Solution time (sec)		Solution time	Gap (%)	Solution time	Gap (%)	Solution time	Gap (%)	Solution time	Gap (%)	Solution time	Gap (%)	Solution time	Gap (%)
50,000	516,753	669		7	1.08	8	1.03	14	1.03	49	0.95	45	0.99	66	0.99
60,000	557,142	198		7	0.87	8	0.88	14	0.88	48	0.82	50	0.85	59	0.87
70,000	597,142	168		8	0.82	9	0.82	11	0.82	41	0.79	61	0.80	46	0.76
80,000	637,142	117		7	0.77	8	0.76	14	0.85	48	0.65	66	0.73	50	0.77
90,000	677,142	150		7	0.76	8	0.71	11	0.73	42	0.67	60	0.77	50	0.70
100,000	717,142	129		6	0.80	8	0.69	11	0.69	50	0.63	71	0.57	61	0.66
110,000	750,190	15		7	0.15	6	0.49	13	1.26	52	0.09	60	0.48	64	1.04
120,000	770,190	14		7	0.15	7	0.48	12	0.73	58	0.09	53	0.35	68	0.71
130,000	790,190	4		7	0.12	6	0.50	16	1.46	48	0.09	54	0.32	57	0.88
140,000	810,190	4		6	0.14	5	0.54	17	0.65	55	0.06	63	0.14	70	0.60
150,000	830,190	3		6	0.14	7	0.64	14	0.98	57	0.08	51	0.19	62	0.62
Average				7	0.53	7	0.69	13	0.92	50	0.45	58	0.56	59	0.78

Table 2: 16-node Turkish network solutions with the tight hub capacity set

Fixed cost	Optimal solution		Heuristic algorithm											
	Objective function value	Solution time (sec)	Generating 6,000 solutions						Generating 12,000 solutions					
			$p=\sqrt{n}-2$		$p=\sqrt{n}$		$p=\sqrt{n}+2$		$p=\sqrt{n}-2$		$p=\sqrt{n}$		$p=\sqrt{n}+2$	
			Solution time	Gap (%)	Solution time	Gap (%)	Solution time	Gap (%)	Solution time	Gap (%)	Solution time	Gap (%)	Solution time	Gap (%)
50,000	522,237	519	11	0.99	10	1.00	17	0.98	14	0.98	69	0.79	24	0.97
60,000	562,237	169	12	0.92	11	0.90	13	0.97	23	0.92	71	0.74	31	0.91
70,000	602,237	118	11	0.98	8	0.85	14	0.84	18	0.86	62	0.84	54	0.94
80,000	642,237	149	12	0.83	8	0.89	16	0.88	17	0.69	54	0.80	44	0.80
90,000	682,237	109	10	0.84	8	0.80	12	0.84	26	0.82	57	0.61	36	0.75
100,000	722,237	170	9	0.76	7	0.71	13	0.71	25	0.71	58	0.72	22	0.72
110,000	750,790	14	7	0.13	8	0.83	12	0.98	34	0.08	47	0.22	43	1.46
120,000	770,790	9	7	0.15	9	0.52	15	1.49	40	0.10	57	0.36	39	1.09
130,000	790,790	2	7	0.12	7	0.79	14	1.16	28	0.07	60	0.09	45	0.86
140,000	810,790	4	7	0.14	8	0.34	12	1.97	38	0.07	52	0.26	42	1.13
150,000	830,790	3	6	0.12	7	0.35	17	1.37	42	0.06	49	0.11	48	1.07
Average			9	0.54	8	0.73	14	1.11	28	0.49	58	0.50	39	0.97

18

$p = \sqrt{n}$ obtained the lowest gaps. The average optimality gaps of all the instances listed in Tables 1 and 2 are 0.50%, 0.62%, and 0.95%, for $p = \sqrt{n} - 2, \sqrt{n}$, and $\sqrt{n} + 2$, respectively. When the average values are compared, $p = \sqrt{n} - 2$ is evidently superior.

Table 3 presents the average gaps that were obtained as a result of 660 runs with the 16-node Turkish network under different parameter settings. Observe from the table that the optimality gaps were slightly lower at the instances with the loose capacity set. The heuristic algorithm performed a little better when the capacities were loose.

Table 3: Average values obtained with the 16-node Turkish network.

Capacity set	Solutions generated	Average gaps (%)		
		$p = \sqrt{n} - 2$	$p = \sqrt{n}$	$p = \sqrt{n} + 2$
Loose	6,000	0.53	0.69	0.92
	12,000	0.45	0.56	0.78
Tight	6,000	0.54	0.73	1.11
	12,000	0.49	0.50	0.97

When the average gaps were compared under different p values, $\sqrt{n} + 2$ had the worst performance and $\sqrt{n} - 2$ had the best. As expected, optimality gaps decrease when more solutions were generated. However, there was no drastic decrease in the average optimality gaps when the number of generated solutions were doubled from 6,000 to 12,000. In order to observe the effects of parameters on both optimality gaps and solution times, we conducted a statistical analysis with the heuristic algorithm. The results of this statistical analysis are provided in Section 4.1.

For the CAB data set with 25 nodes, we tested only $p = \sqrt{n} - 2$ and $p = \sqrt{n}$, as the performance with $p = \sqrt{n} + 2$ were not satisfactory on the Turkish network. Since the number of nodes were increased from 16 to 25, we increased the number of solutions to be generated by the heuristic algorithm as well. We tested generating 12,000 and 24,000 solutions. Table 4 presents the results that were obtained with the CAB data set.

Observe from the third column of Table 4 that the CPU time requirement by CPLEX increased considerably to solve the model to optimality on the CAB data set. The instance corresponding to the fixed cost value of 50,000 could not be solved optimally in 12 hours. The gap reported by CPLEX at the end of 12 hours was still 8%, thus, we omitted that instance from the computational analysis. This instance clearly justifies the need for a heuristic algorithm. On the average, the heuristic algorithm required 13 seconds for generating 12,000 and around 5 minutes for generating 24,000 solutions.

The last row of Table 4 summarizes the average gaps obtained with the CAB data. Again,

Table 4: CAB data set solutions.

Fixed cost	Optimal solution		Heuristic algorithm							
	Objective function value	Solution time (sec)	Generating 12,000 solutions				Generating 24,000 solutions			
			$p = \sqrt{n} - 2$		$p = \sqrt{n}$		$p = \sqrt{n} - 2$		$p = \sqrt{n}$	
			Solution time	Gap (%)	Solution time	Gap (%)	Solution time	Gap (%)	Solution time	Gap (%)
60,000	859,086	1062	15	1.87	14	2.22	279	1.07	285	1.08
70,000	909,086	1474	13	1.27	13	1.17	298	0.98	301	1.07
80,000	959,086	1124	14	1.03	12	1.10	275	1.08	312	0.99
90,000	1,009,086	889	13	1.02	12	1.33	259	0.63	311	0.77
100,000	1,059,086	1304	12	0.85	13	1.08	285	0.74	320	0.86
110,000	1,109,086	1117	13	0.62	13	0.65	286	0.45	315	0.47
120,000	1,150,660	1012	11	0.72	13	0.67	272	0.37	311	0.40
130,000	1,190,660	1160	11	0.67	15	0.72	274	0.46	301	0.41
140,000	1,230,660	740	13	0.56	16	0.78	272	0.42	296	0.41
150,000	1,270,660	1742	14	0.64	13	0.75	270	0.31	293	0.39
Average			13	0.93	13	1.05	277	0.65	305	0.69

we had lower averages when $p = \sqrt{n} - 2$ compared with $p = \sqrt{n}$. Generating 24,000 solutions naturally resulted in better quality solutions than generating 12,000 solutions. In the best case, our heuristic was able to obtain solutions within 0.65% of optimality on the average. Average gaps with the CAB data turned out to be slightly higher than the average gaps obtained with the Turkish network.

For both of the data sets, observe from Tables 1, 2, and 4 that the quality of the heuristic solutions are dependent on the fixed cost of establishing hubs. The statistical analysis presented in the next section also confirms that the fixed cost has significant effects on both the solution time and the optimality gap. We verified by comparing the optimal hub networks and hub networks of the heuristic solution that the heuristic algorithm is able to find the optimal hub locations and allocations in almost all of the instances. Optimality gaps usually result from vehicle assignments. With lower fixed cost values, more hubs are established and, thus, hub networks contain more connections resulting in higher gaps due to inefficient vehicle assignments. One way to overcome this issue is to consider additional neighborhood structures in the vehicle assignment phase of the algorithm. There is surely room for improvement. Nevertheless, we believe our heuristic algorithm is able to find good quality solutions in reasonable CPU times in its current form.

4.1 Statistical Analyses

We additionally conducted a statistical experiment to observe the effects of parameters on the performance of the heuristic algorithm. We tested the effects of the number of hubs in the initial

solution, number of solutions generated, set of hub capacities, and fixed costs of establishing hubs, on both the solution times and quality. This experiment was conducted using Design-Expert Software Version 9 considering a total of 660 runs taken with the 16-node Turkish network.

The hypotheses for this analysis were as follows:

$H_0 =$ The factor is not significant at the 95% confidence level.

$H_{alternative} =$ The factor is significant at the 95% confidence level.

The critical *p-values*, which represent the significance level of each factor, are shown in Table 5.

Table 5: Statistical analyses with the heuristic.

Factors	*p-values*	
	Solution time	Optimality gap
Number of hubs in the initial solution ($\sqrt{n} - 2, \sqrt{n}, \sqrt{n} + 2$)	< 0.0001	0.0003
Number of solutions generated (6,000, 12,000)	< 0.0001	0.004
Capacity set (loose, tight)	< 0.0001	0.004
Fixed hub establishment cost (50,000–150,000)	< 0.0001	< 0.0001

In Table 5 the first column lists the factors, the second and third columns show the corresponding *p-values* for the effect of each factor on the solution time and optimality gap of the heuristic solution. If the *p-value* of a factor is less than 0.05, the null hypothesis is rejected and the alternative hypothesis is accepted. Otherwise, the null hypothesis cannot be rejected.

According to the results of the statistical analysis, we conclude that each of the factors that were tested has a significant effect on both the solution time and the optimality gap.

4.2 Lower Bound

In order to comment on the solution quality of the heuristic algorithm for larger-sized instances, one needs a good lower bound. We propose solving a relaxed version of the problem in order to obtain lower bounds. To generate a lower bound we relaxed the integer variables in the mathematical model which determine the number of vehicles of each type to use (z_{ij}^v). Only this set of variables were relaxed, the remaining binary variables were not relaxed. We tested the quality of this lower bound on the optimal solutions obtained with the 16-node Turkish network and the CAB data set. Averages of the results are provided in Table 6.

For each data set, Table 6 presents the average CPU time requirements by CPLEX to solve both the original and the relaxed model to optimality under different fixed cost values. The last column of the table lists the percent gap of the lower bound from the optimal value.

Table 6: Performance of the lower bound.

Data set	Optimal solution time (sec)	Relaxed solution time (sec)	Lower bound gap (%)
Turkish network – loose capacity set	134	6	0.30
Turkish network – tight capacity set	115	5	0.38
CAB data	1,162	57	0.47

The average CPU time requirement to obtain the lower bound was a few seconds for the 16-node Turkish network, whereas, it was just less than a minute with the CAB data set containing 25 nodes.

We were able to obtain very good quality lower bounds with the proposed methodology. In the worst case, our lower bounds were within 0.47% of the optimal on the average. On the other hand, for the same instances, the average lower bound obtained by the LP relaxation was 84.75%.

One may utilize this lower bound for developing an exact solution methodology for the problem. However, note that the relaxed problem is still an integer programming formulation due to the binary variables. When the dimension of the problem is large, even the relaxed problem could not be solved exactly. In this case, one needs to develop a solution methodology for the relaxed problem as well, and this is beyond the scope of the current study.

In the next section, we present sensitivity analyses with the 81-node Turkish network. For these instances, unfortunately, we could not solve the relaxed problem optimally within ten hours using CPLEX. We used the best lower bound of the relaxed solution reported by CPLEX at the end of ten hours to comment on the quality of the solutions we obtained with the 81-node Turkish network. The results are presented in the next section.

5   Sensitivity Analyses with the Turkish Network

In this section, we used all the 81 nodes in the Turkish network and analyzed the resulting hub networks under different values of problem parameters. All nodes were included in the potential hub set.

Initially, we did a preliminary analysis with the heuristic using the 81-node Turkish network. We found that using $p = \sqrt{n} - 2 = 7$ hubs in the initial solution results in good quality solutions. In addition, we observed that generating more than 75,000 solutions does not improve the solution quality much. So, we generated a maximum of 75,000 solutions with the heuristic. The heuristic algorithm required 15 minutes on the average to generate 75,000 solutions. We tested

different values for the fixed hub establishment costs and the fixed costs of vehicles (the total of purchasing and operational costs), then we observed the effects of changing these values on the resulting hub networks.

It is computationally very demanding to compute lower bound values for all of the instances that we tested with the 81-node Turkish network. As noted in the previous section, in order to obtain a lower bound, we put a ten hour limit to CPLEX while solving the relaxed problem. Unfortunately, CPLEX could not find the optimal solutions of the relaxed problem within ten hours. We used the best lower bound of the relaxed solution reported by CPLEX at the end of ten hours to comment on the quality of our heuristic solutions. We computed the gaps between this lower bound and the heuristic solution for some selected instances. In these instances, the average gap from the lower bound was 4.5%.

First, we tested different fixed cost values for establishing hubs. Similar to the analysis we did in the previous section, we varied the fixed cost values from 50,000 to 150,000 TL in increments of 10,000. Table 7 shows the results.

Table 7: The effects of changing fixed hub establishment costs.

Fixed cost	Number of hubs	Air hub locations	Ground hub locations	Number of airplanes	Number of trailers	Number of trucks
50,000	9	1, 6, 29, **34**, 35, 63	**1**, **6**, 7, **16**, **29**, **34**, **35**, 42, **63**	9	297	20
60,000	8	6, 25, 27, 34	**6**, 7, **16**, **25**, **27**, **34**, **35**, 42	10	241	30
70,000	7	1, 6, 21, 34, 42	**1**, **6**, 7, **21**, **34**, **35**, **42**	11	311	27
80,000	7	6, 12, 34	**1**, **6**, 7, 11, **12**, **34**, **35**	11	241	43
90,000	6	1, 6, 12, 34	**1**, **6**, **12**, 16, **34**, **35**	10	238	85
100,000	6	1, 6, 21, **34**	**1**, **6**, 21, **34**, **35**, 42	11	230	49
110,000	6	1, 19, 21, 34, 45	**1, 19, 21, 34, 41, 45**	13	223	27
120,000	6	1, 6, 24, 34, 35	**1**, **6**, 16, **24**, **34**, **35**	13	209	58
130,000	6	1, 6, 21, 34, 35	**1**, **6**, **21**, 26, **34**, **35**	14	217	29
140,000	6	6, 27, **34**, 35, 42, 69	**6, 27, 34, 35, 42, 69**	14	215	32
150,000	4	6, 21, 34	**6, 21, 34, 35**	14	145	77

Table 7 analyzes the effects of changing fixed cost values on the resulting hub locations and number of different types of vehicles purchased and operated on the inter-hub network.

Locations of air and ground hubs listed in this table correspond to the node numbers presented in the Turkish network data set (Beasley [6]).

There were two capacity levels available for each transportation mode: small and large. For example, a small air hub can handle 6 airplanes whereas a large one can handle 12. There is cost incurred for installing capacities for each transportation mode. The bold node numbers listed under the columns reporting air and ground hub locations identify large capacity levels. For instance, when the fixed hub cost was 100,000 in Table 7, a small capacity air hub was established at locations 1, 6, 21, and a large capacity air hub at node 34. Similarly, a ground hub with small capacity level was established at locations 21 and 42, whereas a ground hub with large capacity level was established at 1, 6, 34, and 35.

When the fixed costs increase, total number of hubs tends to decrease. For each air hub location in Table 7, a capacity for ground transportation was also installed. Actually, this is a result that we obtained in all of our solutions. That is, air hubs were selected from among the set of ground hub locations. This is a very realistic insight as it should be possible to employ ground transportation between any pair of hubs in real-life.

Hub locations tend to differ with varying fixed costs. Even though fixed cost values were the same for all potential hub locations, some locations were commonly preferred, such as 6, 34, and 35. On the other hand, there were also some nodes which was chosen as a hub only in one instance in Table 7, such as nodes 26, 45, and 63. Commonly preferred nodes generate high amount of flow in the Turkish network. For example, Ankara (6) and Istanbul (34) are the two largest cities in Turkey generating the highest amount of flow and they appeared as a hub in all of our solutions. This result is realistic as every small parcel delivery company operating in Turkey has a hub operating both in Ankara and Istanbul.

When the fixed costs are high, less number of hubs are established and, hence, the amount of flow concentrated at each hub is high. Consequently, hub arcs carry relatively higher amounts of flow. When the amount of flow to be carried on a hub arc is high, using an airplane is usually more advantageous than using trailers or trucks. Observe from Table 7 that when less number of hubs were established, more airplanes were employed on the hub network.

According to our results, the most preferred vehicle for the Turkish network data was trailer. This result coincides with the current networks of the three largest small parcel delivery companies operating in Turkey. The analysis presented in Table 7 may be very valuable for a company in providing managerial insight for determining the number of vehicles of each type to purchase and operate on the hub network.

Next, we varied the total cost of purchasing and operating vehicles to observe the effect of

changing fixed costs of vehicles on the resulting solutions. For this analysis, we took the fixed hub establishment cost to be 100,000. The results are presented in Table 8.

Table 8: The effects of changing fixed costs of vehicles.

Airplane-trailer-truck	Number of hubs	Air hub locations	Ground hub locations	Number of airplanes	Number of trailers	Number of trucks
4000-275-225	6	1, 6, 12, **34**, 35	**1**, **6**, **12**, 16, **34**, **35**	15	197	46
4025-300-250	6	1, 12, 34, 35, 52	**1**, 7, 12, **34**, **35**, **52**	14	213	24
4050-325-275	6	1, 6, 21, **34**, 35	**1**, **6**, **21**, 34, **35**, **42**	13	228	21
4075-350-300	6	6, 25, 27, 34	**6, 25, 26, 27, 34, 35**	12	221	50
5000-364-309	6	1, 6, 21, **34**	**1**, **6**, 21, **34**, **35**, 42	11	230	49
5025-375-315	6	1, 6, 12, 34	**1, 6, 12, 34, 35, 42**	10	215	68
5050-400-325	6	1, 21, 34, 35, 42	**1**, 16, **21**, **34**, **35**, **42**	10	233	34
5075-425-350	6	1, 6, 12, 34	**1, 6, 12, 34, 35, 42**	9	210	93
6000-450-375	6	6, 12, 34	**6**, **9**, **12**, 16, **34**, 80	8	214	92

The fixed costs for each type of vehicle are presented in the first column of Table 8. We ranged the fixed airplane costs from 4,000 to 6,000, trailer costs from 275 to 450, and truck costs from 225 to 375. The combination 5000-364-309 corresponds to the values used in the previous analyses.

In Table 8, for each combination of fixed costs, we present the total number of established hubs, locations of air and ground hubs, and the number of vehicles of each type employed on the hub network. Again, the bold node numbers represent hubs with large capacities.

Similar to our previous analysis, some hub locations were commonly preferred in most of the instances, such as 6, 34, and 35. As can be observed from Table 8, even though the number of hubs to be established is exactly the same, hub locations tend to differ with varying fixed vehicle costs. These results clearly demonstrate the need to jointly consider hub location and vehicle assignment decisions.

As expected, when the fixed cost of an airplane was lower, more airplanes were employed on the hub network. At the instances presented in Table 8, the number of airplanes to be

employed on the hub network ranged between 8–15, the number of trailers between 197–233, and the number of trucks between 21–93. At each instance, the number of vehicles of each type to employ on the hub network was determined considering the capacities and costs of different types of vehicles.

When we compared our findings with the hub networks of the small cargo delivery companies operating in Turkey, we found out that our methodology resulted in establishing less number of hubs compared with their currently operating hub networks. Even though locations of the hubs varied with different values of problem parameters in our results, our solutions usually contained at least one hub from each of the seven geographical districts in Turkey. Figure 3 demonstrates a solution on the map of Turkey.
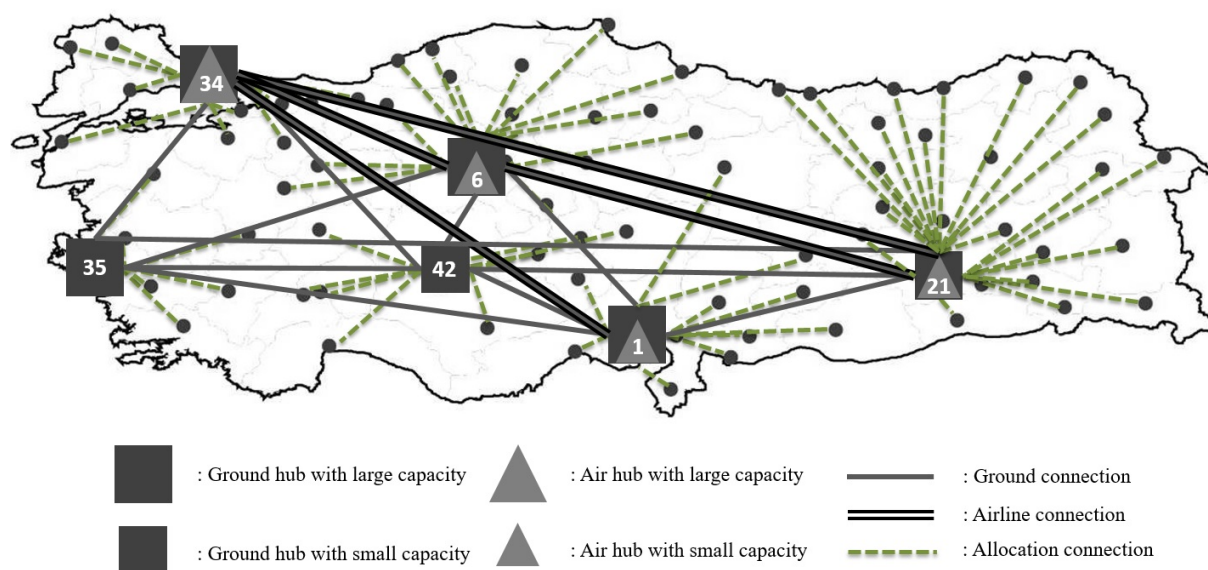

Figure 3: Solution with the fixed cost value of 100,000 on the map of Turkey.

The solution depicted in Figure 3 corresponds to the instance with the 100,000 fixed cost which is presented in the sixth row of Table 7. The figure shows the locations of air and ground hub nodes, airline and ground connections on the hub network as well as the allocation of non-hub nodes to hubs. Four air hubs (1, 6, 21, and 34) and six ground hubs (1, 6, 21, 34, 35, and 42) were established in this solution. Observe from Figure 3 that the locations of the hub nodes are geographically dispersed in the region.

11 airplanes, 230 trailers, and 49 trucks were purchased and operated on the hub network depicted in Figure 3. The number of vehicles of each type operated on each hub arc are presented in Table 9.

Observe from Figure 3 and Table 9 that airplanes were operated only to and from the hub

Table 9: Number of vehicles of each type operated between the hub nodes in the solution depicted in Figure 3.

Hub node	1 A	1 Tra	6 A	6 Tra	6 Tru	21 A	21 Tra	21 Tru	34 A	34 Tra	34 Tru	35 Tra	35 Tru	42 Tra	42 Tru
1	-	-	-	11	-	-	10	-	-	13	-	6	1	-	10
6	-	11	-	-	-	1	8	-	2	2	2	10	-	9	-
21	-	10	1	8	-	-	-	-	2	1	3	10	-	9	-
34	1	6	2	2	6	2	3	3	-	-	-	14	-	14	-
35	-	7	-	10	-	-	10	-	-	13	-	-	-	-	10
42	-	6	-	9	-	-	7	4	-	11	-	-	10	-	-

node 34 (Istanbul). On the other hand, ground connections were present between each pair of hubs. Hence, ground inter-hub network turned out to be complete, whereas, air inter-hub network was incomplete.

Flows on the Turkish network are not symmetrical. Even though they are not symmetrical, the values of w_{ij} and w_{ji} are very close to each other. Because of this, total number of vehicles of each type operated on the hub arcs in each direction were close to each other. For example, in Table 9, 11 trailers were employed between hub nodes 1 and 6 in each direction, similarly, 10 trailers operated between hub nodes 6 and 35 in each direction. This result is very desirable from the cargo companies' perspective as they prefer to transport the flow with the same set of vehicles in the opposite direction the very next day. The analyses and the realistic outcomes presented in this section show that our model and algorithm is a valuable tool for decision making.

6 Conclusion

In this paper, we allowed using alternative transportation modes and different types of vehicles in hub networks to be designed. The contributions of this paper are many fold. We developed a mixed-integer programming formulation of the problem. The model determines the optimal number of vehicles of each type to purchase and operate on the hub network; i.e., the optimal fleet size. We modeled economies of scale correctly to reflect the trade-off between the capacities of different types of vehicles, and the purchasing, operational and transportation costs. The model also considers material handling costs of flow arriving to hubs with different transportation modes. Based on real-life observations, hub capacities were defined, for the first time, as the maximum number of vehicles of each type that can be handled at hubs.

In order to solve realistically sized instances, we developed a variable neighborhood search

algorithm as a solution methodology for this problem. We tested our heuristic algorithm both on the Turkish network and CAB data sets. The heuristic algorithm found good quality solutions in reasonable CPU times. On the average, the heuristic algorithm was able to obtain solutions within 0.49% and 0.65% of optimality with the Turkish network and the CAB data set, respectively.

We conducted extensive sensitivity analyses with the 81 node Turkish network. Our results provided managerial insights for the decisions on the locations and capacities of hubs, transportation modes to serve at these hubs, allocations of non-hub nodes to hubs, and how many vehicles of each type to purchase and operate on the hub network. The developed model and heuristic algorithm were proved to be valuable tools for the decision makers in the design of intermodal hub networks.

For future research, one may consider developing exact solution methodologies for the problem. Incorporating tactical decisions such as the frequency and schedules of the vehicles is another area of research which is definitely worth pursuing. With the integration of tactical operations and consideration of material handling constraints, the interdependency and complex choreography of transportation modes can better be modeled.

Acknowledgement

References

[1] Alumur, S., Kara, B., 2008. Network hub location problems: The state of the art. European Journal of Operational Research 190, 1–21.

[2] Alumur, S., Kara, B., Karasan, O., 2009. The design of single allocation incomplete hub networks. Transportation Research Part B 43, 936–951.

[3] Alumur, S., Kara, B., Karasan, O., 2012. Multimodal hub location and hub network design. Omega 40, 927–939.

[4] Alumur, S., Nickel, S., Saldanha-da-Gama, F., Secerdin, Y., 2015. Multi-period hub network design problems with modular capacities. Annals of Operations Research doi: 10.1007/s10479-015-1805-9.

[5] Alumur, S., Yaman, H., Kara, B., 2012. Hierarchical multimodal hub location problem with time-definite deliveries. Transportation Research Part E 48, 1107–1120.

[6] Beasley, J. E., 1990. OR Library: Hub location. http://people.brunel.ac.uk/˜mastjjb/jeb/orlib/phubinfo.html.

[7] Bryan, D. L., 1998. Extensions to the hub location problem: Formulations and numerical examples. Geographical Analysis 30, 315–330.

[8] Camargo, R. S., G. Miranda, J., Luna, H. P., 2009. Benders decomposition for hub location problems with economies of scale. Transportation Science 43, 86–97.

[9] Campbell, J. F., 2013. Modeling economies of scale in transportation hub networks. Proceedings of the 46th Annual Hawaii International Conference on System Sciences (HICSS), 1154–1163.

[10] Campbell, J. F., Ernst, A. T., Krishnamoorthy, M., 2005. Hub arc location problems: Part I - Introduction and Results. Management Science 51, 1540–1555.

[11] Campbell, J. F., O'Kelly, M. E., 2012. Twenty-five years of hub location research. Transportation Science 46, 153–169.

[12] Contreras, I., 2015. Location Science. Springer, Ch. Hub Location Problems, pp. 279–301.

[13] Contreras, I., Fernández, E., Marín, A., 2010. The tree of hubs location problem. European Journal of Operational Research 202, 390–400.

[14] Crainic, T. G., 2000. Service network design in freight transportation. European Journal of Operational Research 122, 272–288.

[15] Farahani, R. Z., Hekmatfar, M., Arabani, A. B., Nikbakhsh, E., 2013. Hub location problems: A review of models, classification, solution techniques,and applications. Computers & Industrial Engineering 64, 1096–1109.

[16] Gomes, B. N., Marting, A. X., Camargo, R. S., Ramirez, J. A., 2013. An efficient genetic algorithm for the design of hub-and-spoke networks. IEEE Communication Letters 17, 793–796.

[17] Hansen, P., Mladenović, N., Pérez, J. A. M., 2010. Variable neighbourhood search: methods and applications. Annals of Operations Research 175, 367–407.

[18] Ishfaq, R., Sox, C. R., 2011. Hub location-allocation in intermodal logistics networks. European Journal of Operational Research 210, 213–230.

[19] Kimms, A., 2007. Perspectives on Operations Research. Springer, Ch. Economies of scale in hub & spoke network design models: We have it all wrong, pp. 293–317.

[20] Limbourg, S., Jourquin, B., 2009. Optimal rail-road container terminal locations on the European network. Transportation Research Part E 45, 551–563.

[21] Magnanti, T. L., Mirchandani, P., Vachani., R., 1995. Modeling and solving the two facility capacitated network loading problem. Operations Research 43, 142–157.

[22] Martins de Sa, E., Contreras, I., Cordeau, J. F., de Camargo, R., Miranda, G., 2015. The hub line location problem. Transportation Science 49, 500–518.

[23] Meng, Q., Wang, X., 2011. Intermodal hub-and-spoke network design: incorporating multiple stakeholders and multi-type containers. Transportation Research Part B 45, 724–742.

[24] Nickel, S., Schöbel, A., Sonneborn, T., 2001. Mathematics Methods and Optimization in Transportation Systems. Springer, Ch. Hub Location Problems in Urban Traffic Networks, pp. 1–12.

[25] O'Kelly, M., 1987. A quadratic integer problem for the location of interacting hub facilities. European Journal of Operational Research 32, 393–404.

[26] O'Kelly, M., Bryan, D. L., 1998. Hub location with flow economies of scale. Transportation Research Part B 32, 605–616.

[27] O'Kelly, M., Campbell, J. F., de Camargo, R. S., de Miranda, G., 2015. Multiple allocation hub location model with fixed arc costs. Geographical Analysis 47, 73–96.

[28] O'Kelly, M., Lao, Y., 1991. Mode choice in a hub-and-spoke network: A zero-one linear programming approach. Geographical Analysis 23, 283–297.

[29] Podnar, H., Skorin-Kapov, J., Skorin-Kapov, D., 2002. Network cost minimization using threshold-based discounting. European Journal of Operational Research 137, 371–386.

[30] Racunica, I., Wynter, L., 2005. Optimal location of intermodal freight hubs. Transportation Research Part B 39, 453–477.

[31] Wieberneit, N., 2008. Service network design for freight transportation: A review. OR Spectrum 30, 77–112.

[32] Yaman, H., Kara, B. Y., Tansel, B. C., 2007. The latest arrival hub location problem for cargo delivery systems with stopovers. Transportation Research Part B 41, 906–919.

[33] Yoon, M.-G., Current, J., 2008. The hub location and network design problem with fixed and variable costs: formulation and dual-based solution heuristic. Journal of the Operational Research Society 59, 80–89.