Hub Location and Network Design with Fixed and Variable Costs

James F. Campbell  
University of Missouri – St. Louis  
campbell@umsl.edu

Gilberto de Miranda Jr., Ricardo S. de Camargo  
Universidade Federal de Minas Gerais  
miranda@dep.ufmg.br, rcamargo@dep.ufmg.br

Morton E. O’Kelly  
Ohio State University  
okelly.1@osu.edu

Abstract

Transportation hub networks are important for many large-scale logistics and supply chain systems. Basic hub location models rely on some key assumptions about the transportation cost that limits their practicality. We present a new model for hub location and network design that uses fixed and variable transportation costs on all arcs, fixed costs for hubs, and also allows direct arcs. This approach allows the costs for the network components to drive the design of the network. We provide the basic model formulation and some illustrative results to show the types of optimal networks that are generated. The results document the wide range of network features that can be generated from the general cost structure in the model.

1. Introduction

Hub location problems have become a popular research topic for design of large scale transportation networks in logistics and supply chain systems. Hub networks route some (or all) of the origin-destination (OD) flows via hub nodes that provide a switching, sorting or connecting function, and possibly a consolidation and breakbulk (or deconsolidation) function. A typical OD path in a hub network involves travel in a small vehicle (e.g., a small or medium truck) on a collection arc from the origin to a hub, travel between hubs on a transfer arc in a larger vehicle (e.g., large trucks or trains), and finally travel again in a small vehicle on a distribution arc from the last hub to the destination. This path can be described as origin-hub-hub-...-hub-destination. Such an OD path could visit many hubs, though the cost for multiple transfers at hubs tends to limit the number of visits to hubs in practice. Note that some OD paths may not include the transfer arc component (origin-hub-destination), and direct paths (origin-destination) are possible as well, when warranted by sufficient demand or service needs.

The advantages of hub networks stem from the possible lower costs and higher service levels. Lower costs can result from consolidating flows of freight, parcels or passengers on larger vehicles to exploit the strong economies of scale in transportation. Higher levels of service can result from more frequent OD connections than in a network with only direct OD trips, where the demand for particular OD pairs may be insufficient to provide reasonable or competitive service (e.g., at least daily). Routing via the hub nodes allows flows for multiple OD pairs to be consolidated on the same vehicle, thereby better utilizing the available vehicle (weight and volume) capacity.

As discussed in [5,6], basic hub location models since their introduction have used a simple model for the transportation cost. Travel on collection and distribution arcs is charged a standard constant transportation rate per unit distance per unit flow (e.g., $ per pound per mile). The cost rate for travel on transfer arcs between hubs is discounted relative to a standard transportation cost by a constant factor $g<1$. Note that the discount for transfer arc travel is independent of the actual flows through the network. For direct arcs, if they exist, there is also a constant rate per unit distance per unit flow, though this rate may be different than the standard rate for collection and distribution arcs to reflect the different nature of direct OD transport.

This basic form of hub location models presumes that the optimal flows in the network will be concentrated on the transfer arcs (due to their lower cost rate) and therefore warrant a cost discount from the economies of scale. However, it has long been known [3,7,17,18,20,22] that this basic assumption in hub location models is often violated, in that some of the discounted transfer arcs carry smaller flows than other non-discounted arcs. This mismatch between the idealized model of transportation cost and the actual flows is rather common [5], and it can lead to unreasonable and unrealistic network designs. Several remedies have been presented to better model the transportation costs, including flow dependent cost...
discounts [3,4,11,22], incomplete hub networks [2,7,8,9], alternate allocation schemes [24], vehicle based models [16] and models with fixed and variable transportation cost components [17,25]. However, the majority of hub location research, and its many variants, has continued to use a simple model of an assumed flow-independent discount on all transfer arcs. The simple transportation cost model described above reduces the network design component of the problem, as locating the hubs defines the complete backbone network that connects the hubs.

Two key features observed in real-world hub networks are the incomplete nature of the backbone network, and the use of direct OD arcs where appropriate. In this paper we describe a new model for hub location that uses fixed and variable transportation costs on all arcs, fixed costs for hubs, and also allows direct arcs. This allows the costs for establishing and using the network components to drive the design of the network, as occurs in practice, without a need for simplifying assumptions of particular topologies or path length restrictions.

The following section discusses modeling of costs. Section 3 presents the mathematical programming formulation of our model and our solution approach. Section 4 provides computational results to illustrate the behavior of the model, and Section 5 is the conclusion.

2. Modeling costs

The motivating idea behind our new models with a generalized cost structure is to allow the costs to determine the structure of the optimal network, rather than using simplifying assumptions (e.g., a complete backbone is an unintended consequence of the assumed fixed independent cost structure) or imposing artificial restrictions on the topology (e.g., limiting the number of transfer arcs in the backbone or forcing single or multiple allocation). This would seem to better reflect the real-world concerns where there is considerable flexibility in designing hub networks, but each component has costs associated with its establishment and operation.

An alternative approach to modeling transportation in these types of networks is at the vehicle level, by allowing several types of vehicles, each with its own cost structure (fixed and variable cost) and capacity, and then assigning enough vehicles of the various types to arcs to satisfy demand at lowest cost. That approach is used in [16] where there are not specified hub nodes, but rather a concentration of flows through particular nodes in the solution, which suggests “hub-like nodes”. This approach raises the important issue of “What is a hub?”. The common approach used in hub location models defines a hub as the endpoint of a transfer arc. As indicated earlier we can view hubs as performing sorting and/or consolidation/breakbulk functions. Note that sorting of various types may occur at several locations in the network. For example, offering direct services may require sorting at the non-hub origin to separate the direct flows from those going via one or more hubs. If a carrier makes the strategic decision to offer direct services, then they are incurring the costs for this sorting at the origin. Alternatively, a decision can be made to just send everything without any sorting on a collection arc to a hub for sorting. That is a viable approach, especially in freight transport and postal operations.

In our models hubs are particular nodes that perform different functions than non-hubs in allowing switching, sorting, connecting, and/or consolidation and breakbulk activities. This implies there is different infrastructure at hubs compared to non-hubs, though the exact type and magnitude depends on the application (e.g., airline passenger terminals vs. highly automated sorting centers for parcel shipments). Our models are strategic location models that include fixed costs for hubs and seek to locate hub nodes. Thus, in our model fixed costs are associated with the hubs, and the variable costs for operations at the hubs are captured in the arc costs. See [15] for a different approach with more detailed modeling of hub operations as a multistage queueing process.

The simple transportation cost model used in the vast majority of hub location models is based on the OD paths, where the cost to transport demand $W_{ij}$ (e.g., weight of freight or number of passengers) from origin $i$ to destination $j$ along the path from origin $i$ to hub $k$ to hub $m$ to destination $j$ is modeled as

$$W_{ij} (d_{ik} + ad_{km} + d_{mj}),$$

where $d_{ij}$ is the distance from $i$ to $j$, the collection and distribution arc cost rates are 1, and the transfer arc cost rate is $a$. With this approach, the total cost is given by summing the cost for all OD pairs.

O’Kelly at al. [23] introduced a hub location model with generalized costs that distinguishes four types of arcs: direct OD arcs, denoted type 0, provide a direct trip between an origin and destination; collection arcs, denoted type 1, are from an origin to a hub; transfer arcs, denoted type 2, are from one hub to another hub; and distribution arcs, denoted type 3, are from a hub to a destination. Associated with travel on each of these arc types is a fixed cost per unit distance $f^0, f^1, f^2$, and $f^3$, and a variable cost per unit distance per unit flow $b^0, b^1, b^2$, and $b^3$. In this model, the cost incurred on an arc depends on the total flow on the arc, so the cost on the arc cannot simply be decomposed for each
OD pair. Thus, the cost for the path described above \((i-k-m-j)\) must consider the total flow on each of the arcs, denoted \(\text{flow}_{ij}\):

\[
\left( f^1 + b^1 \times \text{flow}_{ik} \right) d_{ik} + \left( f^2 + b^2 \times \text{flow}_{km} \right) d_{km} \\
+ \left( f^3 + b^3 \times \text{flow}_{mj} \right) d_{mj}
\]

The total flow on each arc includes the flow for all the OD pairs that use that arc. Costs for direct arcs are more straightforward, as the only flow on a direct arc is for the associated OD pair. The use of fixed and variable costs for travel on the arcs can be mapped to the use of different types of vehicles for different arc types (e.g., using larger trucks on transfer arcs). Transportation economies of scale are created in this model from spreading the fixed costs over the flow on the arc, so that the cost per unit flow decreases as the flow increases. (O’Kelly [21] uses similar approach to model aircraft fuel burn.) This cost model is similar to [17], but different than models that use non-linear (or piecewise linear) costs to model economies of scale (e.g., [3,4,22]).

As in many hub location models, we associate hub fixed costs with establishing hubs. These costs are incurred to provide the switching, sorting, connecting, consolidation, and breakbulk activities. Note that activities and costs at the hubs may depend on whether a hub is the first one encountered, or a subsequent one. For example, consider a freight shipment. At the first hub there may be the following actions: unload the collection arc vehicle, identify the destination, print a label, affix the label, route to the departure arc (transfer or distribution arc), load the vehicle for departure, etc. At the second hub visited in a path there may be the following actions: unload transfer arc vehicle, read label, route to departure arc (transfer or distribution), load vehicle for departure, etc. Thus, the collection arcs may well have a different cost structure than transfer or distribution arcs to reflect this differential activity at the associated hubs. In the context of airline passengers, the general functions of a hub are the same, but the actual operations are different. For example, the check-in process, including baggage handling, seems different for passengers than for freight, though abstractly the operations are the same. Also, and importantly, passengers can sort themselves by finding the appropriate departure gates; so passenger sorting is more or less the same at each hub (it does depend on the size of the hub). However, the loading process for the passengers may be different at a hub than at a non-hub due to different aircraft types and gate infrastructure.

In our model, every OD flow travels on a path that either begins with a collection arc and ends with a distribution arc, or is a direct arc. However, the model does allow degenerate collection and distribution arcs from a node to itself. Thus every shipment (or passenger) enters the system at some node which is the starting point of either a direct arc or a collection arc. Several actions occur at this node before traversing the arc, which can be grouped as “Checking in the shipment or passenger”. This requires some equipment and staff with associated fixed and variable costs. This cost is incurred by all shipments/passengers and as long as it does not depend on the arc type, then we can remove it from the model as it does not influence the network design (assuming all flow is carried via the network). Following this step, further actions and costs may be associated with particular arcs as a shipment/pasenger flows through the network.

For a direct arc, consider the following six actions: (1) sorting the shipments/passengers and sending them to the departure door, (2) loading the shipments/passengers on the vehicle, (3) servicing the vehicle, including provision of the vehicle, driver/pilot and fuel, (4) providing staff and services for the shipments/passengers during movement on the arc, (5) moving the vehicle across the arc, and (6) unloading the shipments/passengers from the vehicle. The sorting step is independent (in general) of the vehicle on the arc and is needed only if there is more than one vehicle departing the origin at a particular time. If however, there is more than one vehicle departing the origin in a particular time frame, then there will be initial sorting at the origin node to match the passenger (shipment) with the proper vehicle. The other steps for a direct arc described above may depend on the particular vehicle used, so the associated costs may depend on that vehicle. At the end of these steps the passenger or shipment is at the destination node. From there, the passengers will transport themselves outside the system at presumably no cost to the carrier, while shipments may incur a cost to be moved and made available for pickup by the consignee (alternately, these costs could be included in the last step).

Collection arcs have a set of associated actions similar to those on a direct arc. Sorting at the beginning of a collection arc would not be needed in a single allocation hub network, as all flows from an origin are sent to the assigned hub for sorting. At the end of these steps for a collection arc, the passenger/shipment is at its first hub. A degenerate collection arc (from a node to itself) is a special case and none of steps (1)-(6) are needed, although the initial check-in at the origin still occurs.

From the hub at the end of a collection arc, the passenger/shipment will travel on either a transfer arc or a distribution arc, which may be degenerate. To accomplish this there is first a sorting and possibly a consolidation activity that connects the collection arc
with the subsequent transfer or distribution arc. Depending on the particular operations, these activities could occur in an aggregated fashion on all passengers or shipments that have arrived up to a certain time point – or it could occur more continuously as passengers/shipments arrive at the hub. The first sort for a shipment (passenger) may be different than subsequent sorts, so we associate the additional costs of this first sort (over the costs for subsequent sorts) with the collection arc on which a shipment/passenger arrived at the hub.

For transfer arcs, the set of actions required is similar to the six items described for direct arcs and collection arcs. However, the first action for a transfer arc may include sorting and consolidation. Further, the sorting here may be different (e.g., more automated) than the possible sort at the origin, as it will be the second sort for the passenger/shipment, if it was sorted at the origin. The sort at the hub may also involve a much larger volume than an origin sort due to consolidation, so the fixed and variable costs for sorting at the hub may be different than at the origin. Further, there may be consolidation activities to load the incoming flows at the hub on the outbound vehicle on the transfer arc. These loading activities and costs on a transfer arc may differ from those on other arcs due to the different vehicles involved. At the end of the transfer arc, the passenger/shipment is at a hub again (its second or later hub), and from there the passenger or shipment will travel on either another transfer arc or a distribution arc.

Finally, the actions and costs for a distribution arc are handled similarly to the descriptions above. Passengers/shipments to be sent on a distribution arc may require a breakbulk (deconsolidation) action (e.g., unloading containers) if they arrived on a transfer arc. At the end of the distribution arc, the passenger or shipment is at its destination node, and from there the passengers will transport themselves outside the system at presumably no cost to the carrier, while shipments may incur a cost to be moved and made available to the consignee. A degenerate distribution arc is a special case where the breakbulk activity may be needed, but the other actions above are not needed.

The above discussion highlights some of the complexities associated with properly accounting for all relevant actions and costs, and assigning them to the elements of the network. While a finer level of detail in the modeling might include more complex models for various activities depending on the types of arcs being connected at the hub, that is left for future research.

This model described above is illustrated on a small network shown in Figure 1, and exploded into three layers in Figure 2. Suppose there are 8 nodes (denoted 1-8) that are all origins and, except for node 5, are destinations as well. There are 4 hubs located at nodes 1, 3, 6 and 7 which are connected with four transfer arcs shows as double lines. Node 5 as an origin is allocated to multiple hubs (at nodes 1 and 6) and there is a direct arc (dashed line) from node 8 to node 4.

In Figure 2, arcs from the bottom level to the middle level are collection arcs, and arcs from the middle level to the top level are distribution arcs. Vertical arcs in and out of the hubs are degenerate collection and distribution arcs. The dashed arc from 8 to 4 is a direct arc. The vertical (degenerate) arcs incur no costs, but the collection and distribution arcs between layers, and the horizontal transfer arcs in the middle layer all incur costs as described earlier, which are encapsulated in the fixed and variable components.
3. Model formulation

The hub location model with these fixed and variable arc costs was initially formulated in [23] using flow variables that track the flow from each origin (introduced in [13]) and four sets of binary arc variables to handle the fixed costs for each arc (node pair) $i,j$. The binary variables were $y_{ij}^0$, $y_{ij}^1$, $y_{ij}^2$, and $y_{ij}^3$ to indicate the installation of a direct arc, collection arc, transfer arc, and distribution arc, respectively. The non-negative flow variables used were: $Z_{ik}$ to track the flow on the collection arc from origin $i$ to hub $k$, $Y_{ikm}$ to track the flow from origin $i$ to the transfer arc from hub $k$ to hub $m$, and $X_{ijk}$ to track the flow from origin $i$ on the transfer arc from hub $k$ to destination $j$. In addition, the binary variable $z_k$ was used to indicate whether or not node $k$ was a hub. This formulation approach was used in [23] to solve problems with up to 25 nodes (using CPLEX), though solution times for larger problems were excessive.

Because of the difficulty in solving larger problems, we developed an improved formulation that has nice properties from a computational standpoint and allows solution of considerably larger and more challenging instances. This new formulation is based on a different approach and uses a notation similar to Gelerah and Nickel [14] that views flows on arcs in terms of the OD pair. More comprehensive details on this formulation are provided in [19], along with an extension to consider hop constrained models. Similar to above, every OD path begins with a collection (or head) arc and ends with a distribution (or tail) arc. The collection and distribution arcs may be connected directly (i-k-j) or via one or more transfer arcs (e.g., i-kj-kj-kj). Our formulation uses the following non-negative flow variables: $h_{ij}$ to track the flow for OD pair $i,j$ on the collection arc into hub $k$, $x_{ikm}$ to track the flow for OD pair $i,j$ on the transfer arc from hub $k$ to hub $m$, and $t_{ijm}$ to track the flow for OD pair $i,j$ on the distribution arc from hub $k$ to destination $j$. In addition, the binary variables $z_k$ for hub location and $y_{ij}^0$, $y_{ij}^1$, $y_{ij}^2$, $y_{ij}^3$ for arc installation are used. To represent the fixed costs for hub $k$, we use $F_k$. Our formulation of the hub location model with fixed and variable arc costs, and fixed costs for hubs is as follows:

\[
\text{Min } \sum_k F_k z_k + \sum_{i,j} d_{ij}\left(\sum_{m,k} f_0 + b_{ij} W_{ij}\right) y_{ij}^0 +
\]

\[
\sum_{i,j} d_{ij} f_1 y_{ij}^1 + \sum_{m,k} d_{km} f_2 y_{km}^2 + \sum_j d_{mj} f_3 y_{mj}^3
\]

\[
+ \sum_{i,j} W_{ij} d_{ij}^2 h_{ij} + \sum_{i,j} W_{ij} d_{ij}^2 x_{ikm} +
\]

\[
+ \sum_{i,j} W_{ij} d_{ij}^2 t_{ijm}
\]

subject to

\[
\sum_{m \neq j} t_{ijm} + \sum_{k \neq j} x_{ijh} + h_{ij} + y_{ij}^0 = 1 \quad \forall i,j, i \neq j
\]

\[
h_{ijm} + \sum_{k \neq j} y_{ijm} = x_{ikm} + t_{ijm} \quad \forall i,j,m, i \neq j
\]

\[
t_{ij} + \sum_{m \neq j} x_{ijm} = z_i \quad \forall i, j, i \neq j
\]

\[
h_{ijk} + \sum_{m \neq k} y_{ijk} = z_k \quad \forall i,j,k, i \neq j, k \neq i, k \neq j
\]

\[
h_{ijk} + \sum_{k \neq j} y_{ijk} = z_j \quad \forall i,j,k, i \neq j
\]

\[
x_{ikm} \leq y_{ik}^2 \quad \forall i,j,k,m i \neq j, k \neq i, m \neq i, k \neq m
\]

\[
t_{ij} \leq y_{ij}^3 \quad \forall i,j,m, i \neq j, m \neq j
\]

Line 1 of the objective function includes the fixed costs for hubs and the direct arc fixed and variable costs. Line 2 of the objective function is the fixed arc costs for collection, transfer and distribution arcs. Line 3 of the objective function is the variable arc costs for collection and transfer arcs, and line 4 of the objective function is the variable arc costs for distribution arcs. Constraint (1) indicates that the flow for OD pair $i,j$ either travels on a direct arc, a collection arc, or a distribution or transfer arc from some other node $m$. Constraint (2) is the divergence equation for node $i$. Constraint (3) establishes a hub at node $i$ for the appropriate transfer and distribution arc. This ensures that if the origin (node $i$) is a hub, then the flow for OD $i,j$ travels to the destination node $j$ either directly via a distribution arc, or via some intermediate hub $m$. Constraint (4) establishes a hub at node $k$ for the appropriate collection and transfer arc. Constraint (5) establishes a hub at node $j$ for the appropriate collection and transfer arc. This ensures that if the flow for OD $i,j$ is routed via a node $k$ on a collection arc or a transfer arc, then $k$ must be a hub. Constraints (6-8) link the flow variables to the binary arc variables. This formulation is similar to that in [13], but smaller and stronger.

1063
4. Computational results

The formulation presented in the preceding section allows the particular topology of the hub network to be determined by the relative values of the fixed and variable costs in the network. This flexibility creates a very challenging problem to solve to optimality. Our approach uses a Benders Decomposition algorithm described in detail in [19]. Benders Decomposition has been successfully used in solving large and difficult hub location problems (e.g., [4,10,12]). The Benders master problem addresses the network design by minimizing fixed costs for hubs and arcs, and provides a lower bound. The Benders subproblem addresses flows in the hub network with fixed values for the design variables. It provides cuts to add iteratively to the master problem, and upper bounds. Our algorithm also includes some special procedures to select strong Benders cuts, similar to [12] to help solve large-scale instances. Computational work is performed using CPLEX 12.5 to solve the Benders master program and the associated subproblems on a Dell PowerEdge T620 workstation, equipped with two Intel Xeon E5-2600v2 processors and 96 Gbytes of RAM.

Table 1 provides optimal results for selected problems using the CAB25 data set of flows between 25 major US cities. The first two columns provide the fixed and variable arc costs, respectively (where the collection and distribution arc costs are the same, $f^i = f^d$ and $b^i = b^d$). For all problems the fixed hub costs were set to 10,000,000. The first number in the third column of Table 1 indicates the number of hubs located and whether this was found optimally (indicated by a “*”) or was fixed as an input. The second number in the third column of Table 1 indicates the number of transfer arcs connecting the hubs. The two numbers in the fourth column are the number of direct arcs and the number of collection and distribution arcs, respectively. Columns 5, 6 and 7 provide the hub locations, the cpu seconds, and the total cost, respectively.

The first five rows of results in Table 1 are instances where the fixed arc costs provide an incentive for direct arcs and a disincentive for transfer arcs, but the opposite is true for the variable arc costs. The second five rows of results use the same variable arc costs, but equal fixed arc costs. The last five rows of results use equal fixed arc costs and equal variable arc costs. The cpu times vary from about 350 seconds to almost 40 minutes for one instance. In general, the instances with a specified number of hubs were easier than those where the optimal number of hubs was to be determined (e.g., compare rows 1 and 5 of Table 1).

The results show the expected behavior where more arcs are established in the network when the fixed and variable arc costs are reduced (compare the first five rows with the remaining ten rows). The results also document the increasing use of direct arcs as their variable costs are reduced (last five rows vs. middle five rows). The optimal hubs used in the first ten rows are somewhat similar, and are similar to those found for other hub location problems with this data set (e.g., the heavy use of Atlanta (1), Los Angeles (12), New York (17)). However, the optimal hubs in the last five rows are rather different, as here the variable costs do not provide a cost incentive for consolidating flows. There is a greater centralization of hubs in the last five rows, with Denver being the western-most hub and east-coast hubs rarely used. In contrast, in the first 10 rows of Table 1, Los Angeles (12) is always a hub and there is almost always an east-coast hub (usually New York (17)). The incentive for using hubs in the last five rows of Table 1 stems from the ability to connect all nodes at lower fixed cost, not from economies of scale modeled with a variable cost discount on transfer arcs.

Figure 3 provides an illustration of the results for row 1 of Table 1. In this network, the heavy black links are transfer arcs, the dotted blue links are collection and distribution arcs, and the thin red links are direct arcs. This figure shows the six hubs connected by six transfer arcs in a rather sparse incomplete backbone network due to the high costs for the transfer arcs. Note the use of direct arcs in the lower right to connect nearby Tampa (24) and Miami (14), and nearby Houston (10) and Dallas (7), without a need for a lengthy trip via a hub. However, the high variable costs for direct arcs, along with moderate direct arc fixed costs, prevents the widespread use of direct arcs in the network. Some paths have a large amount of circuitry, such as in the upper left to connect Seattle (23) and San Francisco (22) via the hubs at Denver (8) and Los Angeles (12). The longest paths through this network include six arcs (e.g., from the eastern-most city Boston (3) to the western-most city San Francisco (22)). Note that long paths are undesirable for passenger transportation, but may be reasonable for freight shipments. To address path length issues, we have developed a more complex version of this model that includes hop constraints (see [19]) to control the level of service. Not surprisingly, this model is even more challenging to solve.

Figure 4 illustrates the optimal network for the last row Table 1. The impact of the reduced fixed arc costs (1500 for all arcs in Figure 4 vs. 2500-3500 for the arcs in Figure 3) results in a more fully meshed network with eight transfer arcs connecting the six hubs and ten nodes being allocated to two hubs. The lower fixed cost for direct arcs (1500 in Figure 4 vs. 2500 in Figure
3), along with the lack of incentive for consolidation from equal variable arc costs, results in the addition of direct arcs for west-coast OD pairs, and a corresponding shift in the location of western hub from Los Angeles (12) in Figure 3 to Denver (8) in Figure 4. Although there are no hop constraints in Figure 4, it does provide paths that are at most five arcs long (compared to six arc paths in Figure 3). Figure 4 is one illustration of some of the complexities of the network that can result from the added flexibility in our model.

These figures are just two examples of optimal hub networks from our model. The use of fixed and variable costs for all arcs help capture some of the intricate tradeoffs in network design and allows the generation of networks that may better reflect real-world settings. We have successfully solved problems to optimality with up to 80 nodes from the CAB100 data set [19]. These can be very challenging problems, depending on the relative magnitudes of the various cost components, with several hundred branch and bound nodes and solutions times up to 96 hours (see [19] for details).

5. Conclusion

We have developed a new model for hub location that employs fixed and variable costs for all arcs, along with fixed costs for hubs. The model also allows direct arcs so passengers or shipment can travel without visiting a hub. Our primary motivation was to allow the relative costs pressures to drive the design of the network, rather than imposing artificial constraints, restrictions, or assumptions to facilitate solution. Our formulation of this more complex hub location model does have some nice theoretical properties, detailed in [19], that allow us to solve large-scale instances to optimality. The results allow a wide range of network topologies to emerge that document the types of complexity observed in real-world hub networks, and support our efforts to allow the models to respond to the costs associated with network elements.

Extensions of this research are possible in many directions. One issue worth further examination is the role of direct arcs as parallel arcs, where a particular node pair might be linked by both a direct arc and another arc (collection, transfer or distribution). Another extension would be to incorporate service measures to better reflect passenger behavior. Also, competition is a powerful force that could be considered, as that can strongly influence optimal network designs. Finally, a case study with real-world values of the fixed and variable costs would be very useful to explore the validity of the model.

6. References

Table 1. Results for CAB25.

<table>
<thead>
<tr>
<th>$f^0$, $f^1$ &amp; $f^2$</th>
<th>$b^0$, $b^1$ &amp; $b^2$, $b^3$</th>
<th># of hubs / transfer arcs</th>
<th># of direct arcs / collection and distribution arcs</th>
<th>Hubs</th>
<th>Cpu time (secs)</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>2500, 3000, 3500</td>
<td>0.08, 0.04, 0.03</td>
<td>6* / 6</td>
<td>2 / 21</td>
<td>1,8,12,17,20,21</td>
<td>998.93</td>
<td>456,568,748</td>
</tr>
<tr>
<td>2500, 3000, 3500</td>
<td>0.08, 0.04, 0.03</td>
<td>3 / 2</td>
<td>4 / 25</td>
<td>12,20,21</td>
<td>897.79</td>
<td>475,968,673</td>
</tr>
<tr>
<td>2500, 3000, 3500</td>
<td>0.08, 0.04, 0.03</td>
<td>4 / 3</td>
<td>2 / 25</td>
<td>12,17,20,21</td>
<td>973.52</td>
<td>465,857,500</td>
</tr>
<tr>
<td>2500, 3000, 3500</td>
<td>0.08, 0.04, 0.03</td>
<td>5 / 5</td>
<td>2 / 23</td>
<td>1,12,17,20,21</td>
<td>913.21</td>
<td>458,224,214</td>
</tr>
<tr>
<td>2500, 3000, 3500</td>
<td>0.08, 0.04, 0.03</td>
<td>6 / 6</td>
<td>2 / 21</td>
<td>1,8,12,17,20,21</td>
<td>733.81</td>
<td>456,568,748</td>
</tr>
<tr>
<td>1500, 1500, 1500</td>
<td>0.08, 0.04, 0.03</td>
<td>7* / 9</td>
<td>1 / 27</td>
<td>1,4,7,8,12,17,20</td>
<td>812.39</td>
<td>413,745,571</td>
</tr>
<tr>
<td>1500, 1500, 1500</td>
<td>0.08, 0.04, 0.03</td>
<td>3 / 2</td>
<td>6 / 28</td>
<td>12,21,25</td>
<td>373.26</td>
<td>427,959,431</td>
</tr>
<tr>
<td>1500, 1500, 1500</td>
<td>0.08, 0.04, 0.03</td>
<td>4 / 3</td>
<td>3 / 29</td>
<td>12,17,20,21</td>
<td>558.80</td>
<td>420,707,234</td>
</tr>
<tr>
<td>1500, 1500, 1500</td>
<td>0.08, 0.04, 0.03</td>
<td>5 / 6</td>
<td>3 / 24</td>
<td>1,6,12,17,21</td>
<td>486.51</td>
<td>414,545,870</td>
</tr>
<tr>
<td>1500, 1500, 1500</td>
<td>0.08, 0.04, 0.03</td>
<td>6 / 7</td>
<td>3 / 28</td>
<td>1,4,11,12,17,20</td>
<td>587.79</td>
<td>414,278,410</td>
</tr>
<tr>
<td>1500, 1500, 1500</td>
<td>0.04, 0.04, 0.04</td>
<td>7* / 10</td>
<td>5 / 27</td>
<td>1,4,7,8,18,20,21</td>
<td>2343.70</td>
<td>418,569,315</td>
</tr>
<tr>
<td>1500, 1500, 1500</td>
<td>0.04, 0.04, 0.04</td>
<td>3 / 2</td>
<td>13 / 29</td>
<td>8,20,21</td>
<td>487.19</td>
<td>428,897,629</td>
</tr>
<tr>
<td>1500, 1500, 1500</td>
<td>0.04, 0.04, 0.04</td>
<td>4 / 4</td>
<td>13 / 29</td>
<td>1,8,20,21</td>
<td>346.52</td>
<td>422,932,178</td>
</tr>
<tr>
<td>1500, 1500, 1500</td>
<td>0.04, 0.04, 0.04</td>
<td>5 / 5</td>
<td>9 / 29</td>
<td>1,8,18,20,21</td>
<td>637.33</td>
<td>420,958,440</td>
</tr>
<tr>
<td>1500, 1500, 1500</td>
<td>0.04, 0.04, 0.04</td>
<td>6 / 8</td>
<td>6 / 29</td>
<td>1,2,4,6,8,13</td>
<td>929.42</td>
<td>419,568,347</td>
</tr>
</tbody>
</table>


Total cost = 456568748.1517

Figure 3: Illustration for CAB25

Total cost = 419568346.5911

Figure 4: Illustration for CAB25