A Minimax Regret Model for Hub Location under Uncertain Demand and Cost

Iman Kazemian · Samin Aref

Abstract Previous formulation methods offered for capacitated hub location problem, including deterministic models and seasonal optimization, were ineffective under uncertainty. The former might to lead to sub-optimal solutions in case of significant fluctuations in problem parameters and the latter was not practical due to offering different networks while hub location might be a strategic problem requiring one firm solution. In this paper multiple allocation capacitated hub location problem was formulated by a minimax regret model, considering uncertainty factors in setup cost and demand. A computational analysis was performed to investigate the impacts of uncertainty on location of the hubs. The results indicated the importance of the formulation technique in modeling hub location problems from a realistic viewpoint by incorporating different sources of uncertainty. Analyzing a case study derived from an industrial food production company demonstrated the efficiency of the formulation techniques to meet the challenge of companies with seasonal demands and uncertain setup cost by providing a firm strategic solution to be implemented throughout the year.

Keywords Hub location · Uncertain demand · Uncertain setup cost · Capacitated · Multiple allocation · Minimax Regret model

1 Introduction

Hub networks are one of the most common types of logistic systems serving urban transportation, airline networks, communication systems, and cargo networks. The basic characteristic of hub networks is that products are routed through a subset of the connections between nodes and not through direct connections from the origin to the destination.

Hub networks configuration suggest using a set of hubs and spokes for connecting different origins and destinations. Different industries make use of hubs to deal with logistics activities in a more productive way by reducing direct transportation paths. Drawing an analogy might be helpful to clarify the essentiality of hub networks. A complete graph with $K$ nodes has $K \times (K-1)$ edges, while all the nodes can be connected to each other by having a central node (hub) being connected to all the other peripheral nodes (spokes) which reduces the number of edges to $2 \times (K-1)$. So the connectivity can be achieved by utilizing fewer resources more productively. That is where the hub location problem originated from; the challenge of deciding upon location of the hubs for obtaining an efficient logistic network.

The communication industry seems to be the first platform for using hubs, while decades has passed since hubs networks were institutionalized in logis-
tics systems, transportation industry, air cargos, and postal services. Nowadays, hub network design is a common practice for wholesalers, distribution companies, and food production industries whose main objective is enhancing transportation and logistics activities.

The main objective in a common hub location problem is to minimize the total costs of establishing hubs and transporting products between hubs and spokes. Hub location problems are categorized into two subsets of capacitated and uncapacitated problems such that the former embodies most of the real-world problems. After hub location problem was introduced, subsequent problems like p-median, p-hub center, and hub covering problems were emerged to address different challenges of industry in terms of facility location. The principal purpose of p-median is to locate a number of hubs in the network so that the total transportation cost is minimized. The second problem, p-hub center, aims to optimize location of the hubs and allocation of nodes such that the major routes in the network are minimized. In the hub covering problem, minimizing the total cost by finding the optimal location of the nodes and their corresponding allocation shapes the question while the number of hubs is not known. Such a problem introduces limits of coverage as the number of nodes that are connected to a hub is limited. Equally relevant to the problem type, main objective, and the decision variables are the questions of single and multiple network allocation patterns i.e. whether the spokes are to be connected to one hub or multiple hubs.

The hub location problem (HLP) was first posed by OKelly in 1986 [19]. The author introduced Single-HLP concerning assignment of the appropriate location to the hub and its connection to the spokes while there was no cost for hub establishment and it had infinite capacity. He then used a mathematical model to formalise another location problem on hub airline network [20]. The basic model made progress towards P-HLP, a quadratic model, incorporating a number of hubs with direct transportation routes [3].

The proceeding linear model was then developed as p-median location problem by Campbell in 1991 in which each spoke could be connected to more than one hub [5]. A more complex mathematical model for multi allocation hub location problem was also presented by Campbell [6] embodying real world assumptions such as fixed cost for connecting spokes to hubs, minimum flow, and capacity of nodes. Although most of the hub location models developed assumed potential locations for the hubs in a discrete space, there were early research studies relaxing this assumption and considering a continuous space [11, 18].

As the logistic operations got more complicated, new problems emerged with different objective functions, mathematical modeling, and solution techniques. Costa et al. developed hub location problem with infinite capacity, a multi-objective hub network design to minimize the total costs of product flow among the nodes as well as the product processing time in the hubs [14]. The taxonomy of HLP offered four ramifications of hub location problems including Capacitated p-Median Problem [22], HLP with star network structure [23, 26], p-hub center problem [7, 8], and p-hub covering problem [16]. Different solution methods were developed to handle HLPs with distinctive computational advantages like zero-one quadratic programming by which large problems can be simply linearized [15, 23]. For a more detailed review of HLP literature one may refer to two highly cited papers [1, 13].

In the pioneering articles the parameters were assumed to be deterministic though it was not a realistic approach. Some examples of research on HLP with a more realistic approach toward the model definition are as follows. Serra and Marianov developed a model for locating hubs in the network of air transportation by formulating a M/D/c queuing system [17]. The same research field is also investigated by Yang who provided a two-stage stochastic programming model for air transportation with uncertainty in demand [27]. In the same year Sim suggested a model considering the travel time between nodes in a stochastic environment [24]. Moreover, Contreras et al. designed a model for multiple allocation hub location problem in which uncertainty in both demands and transportation costs are accounted for [9]. More recently, Alumur presented multiple allocation and single allocation hub location problem considering uncertainty factors in demand and setup costs [2].

In the real world problems capacity constraints are an indispensable component of hub location mathematical modeling. However, to the best of our knowledge there was no published research on the investigation of uncertainty impact on hub location problem with capacity constraints. In this study, a novel mathematical model was proposed to question this hypothesis whether deterministic modeling and seasonal optimization can be sound measures for obtaining the optimal location of the hubs or not.

As stated earlier hub location is an essential part
of the strategic planning for distribution companies having far-reaching effects on their operational issues and productivity. On the other hand, logistic activities are changing within time while the data used in hub network designed might be outdated by the time of network utilization. Therefore, some of the parameters required for designing the network may not be accurately determined. The most common uncertain parameters are costs, distances, demands and the like. Inobservance to the uncertainty of parameters may lead to obtaining sub-optimal network designs when the determinant input parameters are changed. The reasons behind uncertainty of such parameters are as follows. The volatility of costs for initial procurement such as land, industrial equipment, and raw materials makes the setup cost uncertain. Although, the demand can be predicted by a simple market research, the lag time between designing the network and its actual utilization makes any prediction outdated especially for the case of time-dependent demands like seasonal goods. This issue indicates that uncertainty should also be considered for demand parameters. Sometimes uncertain parameters follow a familiar probability distribution which needs stochastic optimization and there are times when the data is not fit to familiar distributions which require robust programming to take uncertainty into consideration. In both cases considering different scenarios in a discrete probabilistic space would take the uncertainty of parameters into account. This research aimed to investigate the effect of uncertainty on the solution obtained from different modeling techniques previously proposed by the contemporary researchers. It basically concerns different approaches toward the formulation of a typical hub location problem with comprehensive uncertainty factors and capacity constraints.

The structure of rest of the paper is as it follows. Section 2 presents a deterministic model to introduce the mathematical modeling foundation. Section 3 suggests a more sophisticated optimization model with uncertain parameters. The efficiency of optimization model is evaluated in the 4th Section by analyzing a numerical example. Finally, a practical case is discussed in Section 5 to demonstrate the effectiveness of the proposed model in solving real world problems.

2 Basic Model

Optimizing the location of the hubs in a logistic system has prospective impact on total cost making it a crucial strategic decision making process. As an attempt to deal with lack of precise information on the operational parameters of the logistic network, minimax regret model can be deployed.

A quick review on the current literature of hub location revealed the presence of research studies on hub location problems with capacity constraints. However, the capacitated hub location literature suffered from lack of realistic views towards uncertainty that can be obtained by developing stochastic programming and robust optimization models. This study suggested a novel modeling approach with four linear constraints to offer a more realistic view towards uncertainty in demand and setup costs. As mentioned earlier, the foundation of optimization model was first introduced by replicating a deterministic model originally developed by Campbell and then the main minimax regret model was introduced to analyze multiple allocation capacitated hub location problem in a complete graph where there was no direct links between the spokes. The total cost was structured in a way that it included setup costs for hubs and three types of transportation costs including collection costs for spoke-to-hub, distribution costs for hub-to-hub, and transfer costs for hub-to-spoke. Adapting from a deterministic model proposed in a highly cited study, the basic model notation was stated in Table 1. Accordingly, the transportation cost is formulated in Eq. 1.

$$C_{ijkm} = \beta d_{ik} + \alpha d_{km} + \delta d_{mj} \quad \forall i, j, k, m \in N \quad (1)$$

Thus, the multiple allocation capacitated hub location problem is formulated as it follows in (2) to (7):

$$\min \sum \sum \sum \sum W_{ij}C_{ijkm}x_{ijkm} + \sum F_k y_k \quad (2)$$

S.t.

$$\sum x_{ijkm} \leq y_k \quad \forall i, j, k \in N \quad (3)$$

$$\sum x_{ijkm} \leq y_m \quad \forall i, j, m \in N \quad (4)$$

$$\sum \sum \sum W_{ij}x_{ijkm} \leq \Gamma_k y_k \quad \forall k \in N \quad (5)$$
Table 1 Basic model notation

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathcal{N} = {1, 2, ..., n} )</td>
<td>Set of nodes</td>
</tr>
<tr>
<td>( \mathcal{F}_k )</td>
<td>Fixed setup cost ( k \in \mathcal{N} )</td>
</tr>
<tr>
<td>( d_{ij} )</td>
<td>Distance from node ( i \in \mathcal{N} ) to node ( j \in \mathcal{N} )</td>
</tr>
<tr>
<td>( \beta )</td>
<td>Collection cost</td>
</tr>
<tr>
<td>( \delta )</td>
<td>Distribution cost</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>Transfer cost</td>
</tr>
<tr>
<td>( W_{ij} )</td>
<td>Flow routed from node ( i \in \mathcal{N} ) to node ( j \in \mathcal{N} ) via the hubs ( k \in \mathcal{N} ) and ( m \in \mathcal{N} )</td>
</tr>
<tr>
<td>( \Gamma_k )</td>
<td>The capacity of node ( k \in \mathcal{N} ) if a hub is located there</td>
</tr>
</tbody>
</table>

\( y_k \in \{0, 1\} \quad \forall k \in \mathcal{N} \) \hspace{1cm} (6)

\( \sum_k x_{ijkm} \leq y_m \quad \forall i, j, m \in \mathcal{N} \) \hspace{1cm} (10)

\( 0 \leq x_{ijkm} \leq 1 \quad \forall i, j, k, m \in \mathcal{N} \) \hspace{1cm} (7)

In this formulation, \( x_{ijkm} \) stood for the decision variable and \( y_k \) represented a binary variable showing whether node \( k \) is a hub by taking one, or it is a spoke by taking zero. Eq. (3) and (4) ensured that the flow passes through the hubs. Eq. (5) took the capacity of the hubs into account. Finally, domain constraints were formulated in (6) and (7).

3 Minimax Regret Model

As already mentioned, network design is associated with uncertainty. It was assumed in this paper that the uncertainty in demand can be described by considering a limited number of scenarios. It was also assumed that in each scenario demand parameters are certain values. Moreover, uncertain behavior of setup costs is assumed to be interpretable by considering different scenarios. Deploying such scenarios alongside minimax regret programming as the modeling technique, the model will be able to tackle real world problems in an uncertain environment [2]. The minimax regret model notation was stated in Table 2.

\[
\min \sum_s P_s \sum_i \sum_j \sum_k \sum_m W_{ij}^s C_{ijkm} x_{ijkm} + \sum_k F_k^s y_k 
\]

S.t.

\[
\sum_m x_{ijkm} \leq y_k \quad \forall i, j, k \in \mathcal{N} \] \hspace{1cm} (8)

\[
\min \max_{s' \in S'_f} \sum_{m} P_s \sum_{i} \sum_{j} \sum_{k} \sum_{m} W_{ij}^s C_{ijkm} x_{ijkm} + \sum_{k} F_k^s y_k - Z_{s'}^* \quad \forall s' \in S'_f
\] \hspace{1cm} (18)
Table 2  Stochastic model and minimax regret model notation

| S'f | All the scenarios with different uncertain setup costs s' ∈ S'f |
| F'_k | The cost of establishing a hub at node k in scenario s' |
| S'_w | All the scenarios with different uncertain demands s ∈ S'_w |
| P_s | The probability that scenario s ∈ S'_w occurs |
| W_{ij}^s | Demand routed from node i to node j in scenario s ∈ S'_w |

Table 3  Different solutions for capacitated multiple allocation HLP

<table>
<thead>
<tr>
<th>α=0.3</th>
<th>α=0.5</th>
<th>α=0.7</th>
<th>α=1</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDM</td>
<td>Cost in K$</td>
<td>Hub</td>
<td>Cost in K$</td>
</tr>
<tr>
<td>s_f1</td>
<td>2884970</td>
<td>2.3</td>
<td>2969830</td>
</tr>
<tr>
<td>s_f2</td>
<td>2084747</td>
<td>2.3</td>
<td>2169607</td>
</tr>
<tr>
<td>s_f3</td>
<td>2461068</td>
<td>2.4</td>
<td>2547230</td>
</tr>
<tr>
<td>s_f4</td>
<td>1779440</td>
<td>1.3</td>
<td>1864380</td>
</tr>
<tr>
<td>MRM</td>
<td>-</td>
<td>3.4</td>
<td>-</td>
</tr>
</tbody>
</table>

\[ y_k \in \{0, 1\} \quad \forall k \in N \quad (19) \]

\[ 0 \leq x_{ijkm} \leq 1 \quad \forall i, j, k, m \in N \quad (20) \]

4 Computational Analysis

To test the model proposed in [3], the theoretical data related to air transportation for five different cities was used and the other parameters were considered as \( \beta = \delta = 1 \) and \( \alpha \in \{0.2, 0.4, 0.6, 0.8\} \).

Five different scenarios for uncertainty in the setup cost were designed in which \( F'_k \) was selected randomly. Moreover, for analyzing the uncertainty in demands four different scenarios were selected. Probability of each scenario occurrence was 0.25. GAMS software was used to solve the numerical example. Optimum solution in different modes was represented in Table 3 in which basic deterministic model abbreviated to BDM was shown in the first row, four scenarios of the stochastic model were presented in the next rows according to four different values assumed for cost of land represented by \( s_f \), and finally minimax regret model abbreviated to MRM was demonstrated in the lowest row. The problems were solved separately based on each scenario considering minimax regret models. The solutions were compared with the solution obtained from the basic deterministic model in which setup costs and demands were set to the mean values.

The second column form left side of Table 3 showed the total cost and the third column represented the optimal location of the hubs. Note that as the objective function in the represented the minimax regret solution, total cost function could not be compared with the costs of scenarios. Hence, the row for value of the minimax regret cost was left empty.

As was evident in Table 3, the optimum location of the hubs in the minimax regret model differed from the other scenarios. It could be concluded from the observation that it was better to use the minimax regret model instead of estimating costs and demands or using a deterministic scenario.

No relationship was observed between the costs of setting up a hub and selecting a location for the hub. For example, node 4 had the highest setup cost, but in some problems it was selected as the optimal hub. This observation indicated that in addition to the setup cost, the demand and geographical location were also determinant factors in hub location problem.

5 Real-life Example

The application of the proposed model was evaluated using the data from an Iranian industrial food production company. The case of chocolate production in Shirin Asal Tabriz Co. was a known local hub location example, analyzed and referred to by other hub location researchers [21]. Rostami and associates analyzed the impact of estimated demand on the configuration of hub networks in different scenarios. According to their model, the location of hubs should change seasonally. In contrast, they left the costs of changing role of the nodes unsaid i.e. they did not consider the costs of degrading hubs into spokes and upgrading spokes into hubs as their model required...
As already discussed, HLP is a strategic decision making process and it requires an unchangeable solution as proposed by this research.

As mentioned earlier, the fluctuation of prices makes hub establishment an uncertain activity in terms of monetary issues. Moreover, according to the data reported by the company, the demand for chocolate was seasonal and the setup cost was highly variable by time, making it an appropriate case to be analyzed by the proposed model of this study emphasizing on the fluctuations in demand and variability of the setup costs. The scenarios were designed by dividing the year into four seasons with equal length for demand and considering five scenarios for setup costs with $0 < F_k < 1.3$.

The geographical structure of Shirin Asal Tabriz Co. market was as it follows. The main factory was located in the city of Tabriz in the north-west of Iran supporting 36 distribution points with different demand across the country. The national market can be divided into three regions namely: the west, the center and the east. The data presented in this study were related to the demand of the west side of the country with 14 nodes. The company management sought to establish hubs among these cities (14 destinations). Table 4 demonstrated the demands in each scenario. Likewise, Table 4 illustrated the capacities and the fixed costs for establishing hubs in each city. The problem was to find the best location for the hubs according to nodes capacities and uncertainty in demand and setup costs. According to the studies performed for this particular example, $\alpha = 0.4$, $\beta = 1$ and $\delta = 1$ were calculated [21].

Table 4 Demands in different seasons

<table>
<thead>
<tr>
<th>City</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
<th>Winter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rasht</td>
<td>9205</td>
<td>7899</td>
<td>21848</td>
<td>26510</td>
</tr>
<tr>
<td>Kermanshah</td>
<td>10459</td>
<td>11256</td>
<td>18751</td>
<td>19628</td>
</tr>
<tr>
<td>Tabriz</td>
<td>48022</td>
<td>39529</td>
<td>65890</td>
<td>94831</td>
</tr>
<tr>
<td>Tehran</td>
<td>14412</td>
<td>12571</td>
<td>40470</td>
<td>91906</td>
</tr>
<tr>
<td>Zanjan</td>
<td>10590</td>
<td>11732</td>
<td>20402</td>
<td>21218</td>
</tr>
<tr>
<td>Ghazvin</td>
<td>4424</td>
<td>5995</td>
<td>10848</td>
<td>12502</td>
</tr>
<tr>
<td>Hamedan</td>
<td>6270</td>
<td>4802</td>
<td>9273</td>
<td>9505</td>
</tr>
<tr>
<td>Urvia</td>
<td>17022</td>
<td>16006</td>
<td>24951</td>
<td>23234</td>
</tr>
<tr>
<td>Ardabil</td>
<td>13764</td>
<td>19839</td>
<td>19281</td>
<td>16767</td>
</tr>
<tr>
<td>Sanandaj</td>
<td>7996</td>
<td>6105</td>
<td>11330</td>
<td>10807</td>
</tr>
<tr>
<td>Shahre kord</td>
<td>6142</td>
<td>5721</td>
<td>10320</td>
<td>9065</td>
</tr>
<tr>
<td>Ilam</td>
<td>4044</td>
<td>4135</td>
<td>6689</td>
<td>7881</td>
</tr>
<tr>
<td>Karaj</td>
<td>18519</td>
<td>22050</td>
<td>41018</td>
<td>48957</td>
</tr>
<tr>
<td>Arak</td>
<td>4272</td>
<td>3726</td>
<td>9402</td>
<td>10287</td>
</tr>
</tbody>
</table>

The differences in hub network designs let us argue that deterministic analysis of hub location problems suffered from capability of industrial practice in cases where there are source of uncertainty. Deterministic analysis might lead to sub-optimal hub location solutions that are counterproductive as they imposed additional costs to the company in the long term. It is also noteworthy that deploying this prospective approach leads to substantial savings by obviating the seasonal spoke-to-hub upgrades and hub-to-spoke degrades suggested by Rostami and associates.

6 Conclusion

In this paper, multi allocation capacitated hub location problem was investigated in a setting where setup costs and demands are uncertain. A generic model was developed for considering such sources of uncertainty. The modeling technique was continued by performing a computational analysis to investigate the possible changes in the optimal location of the hub in an uncertain environment.

The result showed that the optimal solution changes when the model is associated with uncertain parameters. According to the numerical example, ignoring the uncertainty may change the whole hub location solution drastically. Moreover, the industrial application of the proposed method was addressed to by discussing a case from an industrial food production company. The results obtained from solving the case study confirmed the previously drawn conclusion about significance of uncertainty and how ignoring it may lead to sub-optimal practices. There are a number of research directions that are hoped to be investigated in the near future. Firstly, one may...
Table 5 Capacity and setup cost for different cities

<table>
<thead>
<tr>
<th>City</th>
<th>Capacity</th>
<th>Avg. setup cost in K$</th>
<th>Upper bound setup cost in K$</th>
<th>Lower bound setup cost in K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rasht</td>
<td>275200</td>
<td>130800</td>
<td>1691040</td>
<td>910560</td>
</tr>
<tr>
<td>Kermanshah</td>
<td>154200</td>
<td>616800</td>
<td>801840</td>
<td>431760</td>
</tr>
<tr>
<td>Tabriz</td>
<td>321640</td>
<td>1286560</td>
<td>1672528</td>
<td>900592</td>
</tr>
<tr>
<td>Tehran</td>
<td>38528</td>
<td>1941120</td>
<td>2523456</td>
<td>1358784</td>
</tr>
<tr>
<td>Zanjan</td>
<td>201360</td>
<td>1005440</td>
<td>1307072</td>
<td>703808</td>
</tr>
<tr>
<td>Ghazvin</td>
<td>123840</td>
<td>495360</td>
<td>649368</td>
<td>346752</td>
</tr>
<tr>
<td>Hamedan</td>
<td>97600</td>
<td>470400</td>
<td>611520</td>
<td>329280</td>
</tr>
<tr>
<td>Tehran</td>
<td>110080</td>
<td>440320</td>
<td>572416</td>
<td>308224</td>
</tr>
<tr>
<td>Ardabil</td>
<td>123840</td>
<td>495360</td>
<td>649368</td>
<td>346752</td>
</tr>
<tr>
<td>Sanandaj</td>
<td>103200</td>
<td>412800</td>
<td>536640</td>
<td>288960</td>
</tr>
<tr>
<td>Shahre kord</td>
<td>137000</td>
<td>550400</td>
<td>715520</td>
<td>385280</td>
</tr>
<tr>
<td>Ilam</td>
<td>82560</td>
<td>330240</td>
<td>429312</td>
<td>231168</td>
</tr>
<tr>
<td>Karaj</td>
<td>460960</td>
<td>2143840</td>
<td>2786992</td>
<td>1500688</td>
</tr>
<tr>
<td>Arak</td>
<td>82835</td>
<td>331340</td>
<td>430742</td>
<td>231938</td>
</tr>
</tbody>
</table>

introduce a more sophisticated problem with pervasive sources of uncertainty by parameters not necessarily following a familiar distribution function to be solved by stochastic programming and compared with the technique used in this paper. It would contribute to such an idea if the research is associated with a critical industrial-sized case of an uncertain environment. Secondly, it is suggested for further research to tackle intractable hub location problems with uncertain parameters by tailored evolutionary algorithms and analyze a practical case with large data set such as international couriers.

References

5. Campbell, J.F.: Location-allocation for distribution systems with transshipments and transportation economies of scale. the Fifth International Symposium on Locational Decisions (ISOLDE V), Lake Arrowhead, CA (1990)