Research on layout optimization of logistics network of metropolitan area based on hub-and-spoke theory

Zhao Fang

Zhao Fang, Beijing, Beijing Jiaotong University

ABSTRACT: In order to give full play to the role of regional logistics in the regional economic development of the metropolitan area, a kind of hub-and-spoke regional logistics network based on hub-and-spoke model was established, which has the characteristics of hub-and-spoke network scale economy, scope economy, density economy and structure economy. Seventeen technical and economic indicators are sorted out from regional economy, industrial and commercial development, transportation infrastructure development and logistics industry development, and 35 counties and districts in the Capital metropolitan area are used as the research objects in 2020. The model is used to determine the network layout among node cities and to optimize the logistics network in the Capital metropolitan area.

1. INTRODUCTION

Since the Chinese government issued the "Guidance on Fostering the Development of Modernized Metropolitan Areas" in 2019, it has become a key task to promote the new urbanization strategy in China to enhance the capacity of central cities themselves and their radiation-driving ability, optimize the allocation of factors and resources with metropolitan areas as carriers, optimize the spatial development pattern, and lead the high-quality development of China's urbanization. As a key element of the integrated development of logistics in the metropolitan area, the optimal layout of logistics network nodes, as the core of logistics network optimization, is essential to enhance the operational efficiency and stability of the entire logistics network in the metropolitan area.

2. THE CONSTRUCTION OF THE HUB-AND-SPOKE LOGISTICS NETWORK

2.1. The concept and characteristics of the hub-and-spoke logistics network

Logistics network is driven by the decentralization of demand and multi-directionality of logistics activities in order to reduce logistics organization cost and realize logistics scale operation[1]. From the perspective of geographical movement form, logistics network is the intermediate form of spatial organization of logistics economic activities formed by logistics center city and logistics economic belt and logistics channel[2].

The spatial pattern of the "hub-and-spoke" network is similar to that of a bicycle wheel, which is a node-path system. In short, goods arriving from different origins (spoke) to different destinations (spoke), or from the same origin to different destinations, must first arrive at an intermediate location (hub) in the hub-and-spoke network, where they are redirected and then enjoy[3]. The purpose is to concentrate traffic flows and achieve economies of scale. The "hub-and-spoke" network originated in the aviation field, and is a product of air transport deregulation and airlines' pursuit of network economies of scale and airports' (logistics hubs) pursuit of economies of scale, and is a special network form of logistics network[4].

The hub-and-spoke logistics network is a transportation network hub node composed of transport flows and hub nodes are divided into several levels, the integrated hub node in a dominant position is the hub, and other nodes are the spoke points[5]. The characteristics of the hub-and-spoke logistics network are that the traffic is gathered to the hub city on the trunk line through the spoke network, and the unit transportation cost is reduced through the conversion and resource gathering of various means of transportation in the node city, forming the scale effect and spatial effect, and improving the timeliness, convenience and economy of logistics activities in the region. Therefore, the key to build a hub-and-spoke regional logistics network lies in the selection of hub nodes and the analysis of radiation range[6].

2.2. Node selection of the hub-and-spoke logistics network

The selection of hub nodes requires the establishment of a complete evaluation index system and qualitative and quantitative analysis to discern the comprehensive logistics strength of each node city in the region[7]. Hub cities should have relatively high level of economic and social development, sufficient logistics demand, relevant talents, broad hinterland and corresponding infrastructure.
equipment, strong comprehensive logistics strength and can provide services for the spoke cities, the spoke cities should be in the corresponding hub city radiation range and provide resources to support the hub city. The selection of the hub cities should meet the qualitative conditions such as minimum number and reasonable distance between cities. Ideally, 4-5 hub nodes are the most reasonable number to achieve the lowest total cost of regional logistics[8]. According to the above idea, Table 1 gives a total of 17 indicators in four aspects to build a city logistics development potential evaluation index system.

<table>
<thead>
<tr>
<th>No.</th>
<th>Level 1 Indicators</th>
<th>Level 2 Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Regional Economic Development</td>
<td>Regional GDP</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>GDP per capita</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>Growth rate of social fixed asset investment</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>Number of resident population</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>Per capita disposable income of residents</td>
</tr>
<tr>
<td>6</td>
<td>Industrial and Commercial Development</td>
<td>Industrial added value</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>Number of industrial enterprises above the scale</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>Total retail sales of social consumer goods</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>Total Import and Export</td>
</tr>
<tr>
<td>10</td>
<td>Transportation Infrastructure Development</td>
<td>Density of highway network</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>Density of highway network</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>Number of railroad logistics nodes</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>Number of civil transport airports</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>Road freight volume</td>
</tr>
<tr>
<td>15</td>
<td>Logistics Industry Development</td>
<td>Number of A-class logistics enterprises</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>Number of national outstanding logistics parks</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>Number of model logistics parks</td>
</tr>
</tbody>
</table>

The level of regional economic development, as an important external factor reflecting whether the city has developed into a logistics hub in the metropolitan area, can reflect the degree of regional logistics demand. Among them, regional GDP reflects the degree of demand of the logistics node city, per capita GDP reflects the level of economic development per capita of the city, the growth rate of total social fixed asset investment is used to measure the level of social reproduction of the city, and the number of resident population and per capita disposable income reflect the vigorous degree of consumer demand and development potential of the city to a certain extent.

As the main target of logistics service, the good or bad development of industry and commerce directly affects the future operation efficiency of logistics node cities. Among them, the industrial added value and the number of industrial enterprises above the scale reflect the industrial logistics demand to a certain extent, the total retail sales of social consumer goods reflect the purchasing power of social goods and the scale of retail market, and the total import and export reflects the development level of foreign trade of the city.

The level of development of transportation infrastructure, as an important factor influencing the classification of logistics node city level, can reflect the logistics development capacity of the region, and the density of road network, highway network density, the number of railroad logistics nodes and the number of civil transport airports are selected as evaluation indicators.

The development level of logistics industry as one of the necessary conditions to measure the potential of the axis city, it is an important indicator to reflect the supply capacity of logistics services in the region. The transportation, road freight volume, the number of A-class logistics enterprises, the number of national excellent logistics parks and the number of model logistics parks are selected as evaluation indicators.

To determine the logistics hub node cities in the metropolitan area, it is necessary to evaluate the logistics development potential of each city in the metropolitan area and determine the selection weights in its metropolitan area logistics node network. In order to make the evaluation results objective and comprehensive, the selection of evaluation indicators involves four aspects: regional economy, industrial development, transportation facilities and logistics industry development, and there is a certain overlap of information reflected by these indicators, and there are also problems such as non-uniformity of the scale and how to determine the weight coefficients when accumulating, which makes the evaluation work relatively complicated[9]. Using principal component analysis method, the comprehensive information of the sample is reflected by finding the comprehensive indicators of the sample, which achieves the effect of condensing the information and solving the problem of determining the weights, making the problem simpler[10]. Therefore, the principal component analysis method is used to quantify the advantages of the conditions for each city in the metropolitan area to become a logistics node city in the metropolitan area, and the basic steps of the analysis are as follows:
Performing normalized transformations:

$$Z_{ij} = \left( x_{ij} - \bar{x}_i \right) / S_j$$  \hspace{1cm} (1)

Finding the correlation matrix:

$$r_{ij} = \frac{1}{n-1} \sum_{n=1}^{n} Z_{ij} Z_{jj}, r_{ii} = 1, r_{ij} = r_{ji}$$  \hspace{1cm} (2)

Finding the eigenroot $$\lambda_g$$ and the eigenvector $$L_g$$:

$$\lambda_g - R = 0, L_g = l_{g1}, l_{g2}, ..., l_{gn}$$  \hspace{1cm} (3)

Finding the principal component vector:

$$F_g = l_{g1} Z_{i1} + l_{g2} Z_{i2} + \cdots + l_{gn} Z_{in}$$  \hspace{1cm} (4)

Determining the number of principal components:

$$\sum_{g=1}^{G} \lambda_g / \sum_{g=1}^{G} \lambda_g \geq 85\%$$  \hspace{1cm} (5)

Receiving a composite rating:

$$F = \sum_{g=1}^{G} \left( \lambda_g / \sum_{g=1}^{G} \lambda_g \right) F_g$$  \hspace{1cm} (6)

2.3. The layout of the hub-and-spoke logistics network node

After determining the location of the hub city, the radiation range of the hub city needs to be determined. In the hub-and-spoke network, the connection between hub cities and non-hub cities should meet the low transportation cost between cities on the one hand, while the transportation cost between cities depends on the geographical location of the cities; on the other hand, it should also make the transportation efficiency improved by achieving the distribution of transportation flow through the transit of hub node cities, while the improvement of transportation efficiency brought by the distribution depends on the intensity of logistics activities between cities.

The urban spatial gravity model is an extension of Newton's gravity formula in the field of economics, and can be used to measure the spatial linkage between two economies. In this paper, the gravity model is used to initially measure the strength of logistics linkages between cities, and to measure the logistics affiliation of non-hub nodes by the proportion of the strength of logistics linkages between each non-hub node and different hub nodes, which can determine the range of the hub cities in the region. The formulae of inter-city logistics linkage strength and logistics affiliation degree constructed by using the gravity model are shown below.

The logistics linkage affiliation formula is as follows:

$$L_{ij} = M_i \ast M_j / D_{ij}^2$$  \hspace{1cm} (7)

The logistics linkage affiliation formula is as follows:

$$P_{ij} = L_{ij} / \sum_{j=1}^{N} L_{ij}$$  \hspace{1cm} (8)

In the formula of logistics linkage strength, $$L_{ij}$$ indicates the strength of logistics linkage between two cities; $$M_i$$ and $$M_j$$ indicate the respective quality of two cities, and the logistics development potential score of logistics node cities in metropolitan area is used to measure the logistics quality of cities in this thesis; $$D_{ij}$$ indicates the distance between two cities. In the logistics linkage affiliation formula, $$P_{ij}$$ is the city's logistics linkage affiliation, which indicates the probability that city i belongs to the axis city j; $$L_{ij}$$ indicates the strength of logistics linkage between two cities; the denominator indicates the sum of the strength of logistics linkage between non-hub node cities and different hub node cities. The city spatial gravity model mainly measures the strength of logistics connection and logistics affiliation between cities, and the results can visualize whether the logistics activities between two cities are frequent or not, which can be used as a method to initially judge the range of hub cities' fullscope, but the connection situation when the cost of different connection methods is minimized is not considered comprehensively from the perspective of cost.

The economy of scale is the biggest advantage of the hub-and-spoke model, therefore, determining the connection between hub nodes and non-hub nodes requires a comprehensive consideration of all connection cases and selecting the least costly connection. The factors affecting the connection cost between different cities are mainly related to the distance between cities and the OD flow between cities, therefore, the estimation of OD flow should be performed first when doing the cost optimization model.

The inter-city OD flow estimation is also measured based on the urban spatial gravity model, but since the flow of goods in different directions is different, the gravity model is slightly modified here by using the urban logistics quality share as the weight. The inter-city OD flow equation is as follows:

$$F_{ij} = M_i / M_i + M_j - D_{ij}^2$$  \hspace{1cm} (9)

$$M_i$$ is the comprehensive capacity of logistics development of node city i; $$D_{ij}$$ is the distance between cities i and j. In the model, $$F_{ij} \neq F_{ji}$$, indicating that the OD flows in different directions are different. The comprehensive capacity of logistics development adopts the score of logistics development potential of logistics node cities in the metropolitan area. The distance matrix between cities is substituted to obtain the logistics linkage strength, and then it is multiplied with the average freight volume of each city to obtain the approximate material flow between node cities.

After the inter-city OD flows are estimated, a single-assignment multi-axis hub spoke model is constructed.

Assumptions.

(i) Assume that the hub cities have been identified and the set of hub cities is $$\{1, 2, ..., m\}$$, where m is the number of hub cities.

(ii) Each non-hub node city is connected to only one hub city, and the set of non-hub node cities is $$\{m + 1, m + 2, ..., n\}$$, where n is the number of all logistics nodes in the region.

(iii) Assume that the non-hub node city i is connected to the hub city s(i), $$i = m + 1, ..., n$$, and the function s(i) determines how the hub-and-spoke network is connected.

Under the hub-and-spoke network structure, there is no direct connection between non-hub nodes, so there are
only the following three cases of transportation paths. (i) Connection path between hub $k$ and hub $l$: $L(k, l)$.  
(ii) Connection path between hub $k$ and node $i$: $L(s(i)) + L(s(i), k)$.  
(iii) Connection path between node $i$ and node $j$: $L(i, s(i)) + L(s(i), s(j)) + L(s(j), j)$.  
Here, assume that $L(i, s(i)) = 0$ so that the 2nd case becomes $L(i, s(i))$ if $k = s(i)$; and the 3rd case becomes $L(i, s(i)) + L(s(j), j)$.

Because the transportation cost is proportional to the inter-city distance and OD flow, the cost per unit distance per unit flow is assumed to be 1. Although this is different from the actual one, the effect of the hypothetical case on the results can be ignored given that the minimum cost here only considers the connection method between nodes. Also, since there are economies of scale for logistics transportation between hub nodes, a discount factor $\alpha$ ($0 < \alpha < 1$) is assumed to exist for transportation costs between hubs. The single-assignment multi-hub problem is that the required function $s(i)$ makes the total transportation cost minimum.  

Under this assumption, the cost can be calculated as shown below.  
(i) Cost between hub nodes.  
\[ C_1 = L(k, l) \times F_{ik} \]  
(ii) Cost between hub and non-hub nodes.  
\[ C_2 = [L(i, s(i)) + \alpha L(s(i), k)] \times F_{ik} \]  
(iii) Cost between two non-hub nodes.  
\[ C_3 = [L(i, s(i)) + \alpha L(s(i), s(j)) + L(s(i), j)] \times F_{ij} \]  

In summary, the single-assignment multi-hub least-cost problem can be expressed as follows.  
Objective function:  
\[ \min \sum (C_1 + C_2 + C_3) \]  
Constraints:  
\[ k, l = \{1, 2, \ldots, m\} \]  
\[ i, j = \{m + 1, m + 2, \ldots, n\} \]  
\[ s(i) \in \{1, 2, \ldots, m\} \text{ and } s(i) \text{ has only unique solutions} \]  

2.4. Algorithm design  
Genetic algorithms started from computer simulation experiments on biological systems, simulating biological evolution developed stochastic full-range search and optimization studies, which can automatically find and reserve relevant data about the search range in the process of search and automatically control the search process to find the best solution\[11\].  

(1) Chromosome coding  
Before running the program, it is necessary to encode the relevant data for the solution and construct chromosomes composed of different strings.  

(2) Generating initial populations  
During the running of the program, a certain number of initial chromosomes will be randomly generated, and the later operations will iterate with the populations composed of these initial chromosomes as the starting point.  

(3) Adaptation calculation  
The higher the fitness of the chromosome, the higher the success rate of further inheritance will be, and the algorithm can run to find the optimal solution.  

(4) Three processes of selection, crossover and mutation  
The three processes of selection, crossover and mutation are the three major operators in the genetic algorithm. Selection is to select the chromosomes of good quality from the population and eliminate the chromosomes of poor quality, in order to inherit the chromosomes of good quality to the next generation and produce new chromosomes with high quality, repeat the above process and inherit them to the next generation, and so on, until the chromosomes of the highest quality are produced; crossover converts part of the structure of two chromosomes to reconstitute new chromosomes, and through crossover this arithmetic process, the search power has been improved to a great extent.  

Variation is when the genetic algorithm finds the neighborhood of the optimal solution by the crossover operation, the neighborhood can converge most quickly and find the optimal solution by the random search ability of part of the space of the variation operator.  

(5) Generating new populations  
After the three operations of selection, crossover and mutation, the new chromosomes generated will constitute a new population, and the new population will re-compute the fitness and repeat the above operation process, and the operation process will be terminated when the fitness of the optimal individual and the population fitness are stable, or when the number of iterations reaches the optimal state.  

3. CASE STUDY  
The metropolitan area of the capital is a region including Beijing, Tianjin and 13 cities in Hebei Province, including Shijiazhuang, Tangshan, Baoding, Langfang, Handan, Xingtai, Qinhuangdao, Cangzhou, Hengshui, Chengde and Zhangjiakou, which is an important core area of the economy in northern China. The level of economic development and logistics industry in the region is high, and the distribution by city shows certain characteristics. The cross-sectional data of 13 cities in the metropolitan area of the capital in 2020 is used as the research object for example analysis.  

3.1. Data analysis by principal component analysis  
Table 2 shows the logistics development potential scores and rankings of 13 prefecture-level cities and above in the capital metropolitan area. As can be seen from Table 2, the logistics development potential scores of Beijing and Tianjin are much higher than those of other cities, and the logistics development potential of Shijiazhuang and Tangshan in Hebei Province is different compared with that of Beijing and Tianjin, but also higher than that of other cities in Hebei Province. Therefore, the four cities of
Beijing, Tianjin, Shijiazhuang and Tangshan can be used as logistics axis cities to undertake the hub function of the regional axis-spoke logistics network. According to the four logistics hub cities in the capital metropolitan area obtained from the empirical study, the logistics hubs meet the requirements of large urban logistics demand and high infrastructure supply level, and the layout of logistics resources according to the hub and spoke network of multiple logistics hub cities can alleviate the congestion problem of single hub cities.

### 3.2. Data analysis of gravity model

The road transportation distance data is selected to calculate the strength of logistics linkage between cities. In addition, since there are negative values in the city logistics comprehensive capability scores obtained by principal component analysis, the logistics comprehensive capability scores of each city are subtracted from the lowest score and shifted up one unit as the adjusted city logistics quality data here. The distances between the four axis cities of Beijing, Tianjin, Shijiazhuang and Shijiazhuang in the capital metropolitan area and the nine non-hub node cities in Hebei province were obtained by consulting the road information network.

The logistic affiliation degree and logistic linkage strength of non-hub node cities and axis cities are shown in Table 3. According to the results of logistics affiliation data of non-hub node cities and hub node cities obtained from the city space citation as model, it can be seen that Beijing has the largest radiation influence, while the four hub cities of Tianjin, Shijiazhuang and Tangshan have one to two radiation influence cities, if the logistics network is laid out according to this, it will inevitably intensify the imbalance of infrastructure layout in the capital metropolitan area, therefore, it is necessary to further optimize the logistics network based on the least cost model. Therefore, it is necessary to further optimize the connection methods between nodes based on the minimum cost model.

### 3.3. Data analysis of Minimum cost model

According to the single-allocation multi-hub hub-and-spoke model, the costs of the connection methods of four different logistics nodes are solved by using genetic algorithm, and the minimum cost solutions obtained are selected as shown in Table 4. The optimal connection method obtained from the single-allocation axis and spoke model is basically consistent with the results of the radiation range of the hub cities measured by the gravity model, with slight differences in the connection methods of Baoding and Chengde. In the gravity model, Baoding is the most influenced by Beijing, with a logistics affiliation of 0.433, while Shijiazhuang is the second most influenced by Baoding, with a logistics affiliation of 0.272. However, in the optimal results obtained from the single-allocation spoke model, the logistics transportation cost of the whole network is the lowest when Baoding is connected to the Shijiazhuang hub. In the gravity model to measure the strength of inter-city logistics links, Chengde is also most influenced by Beijing, with a logistics affiliation of 0.490, and Tangshan has the second highest radiation influence on Baoding, with a logistics affiliation of 0.272. However, in the optimal results obtained from the single-allocation axial spoke model, Chengde has the lowest logistics transportation cost for the entire network when it is connected to the Tangshan hub.
4. CONCLUSION

The capital metropolitan area can build a hub-and-spoke logistics network with four cities of Beijing, Tianjin, Shijiazhuang and Tangshan as the axis cities. Among them, Beijing relies on its larger economic strength, perfect transportation infrastructure network and superior geographical location to drive the logistics development of Langfang and Zhangjiakou, forming Beijing Logistics Circle; Tianjin relies on its logistics hub function to drive the development of logistics in Cangzhou, forming Tianjin Logistics Circle; Shijiazhuang accelerates the construction of modern trade logistics base to undertake the logistics of Handan, Xingtai, Hengshui and Baoding, forming Shijiazhuang Logistics Circle; Tangshan accelerates the construction of modern trade logistics base, undertaking the logistics of four cities Shijiazhuang to accelerate the construction of modern trade logistics base, to undertake the logistics of Handan, Xingtai, Hengshui, Baoding four cities, and radiation drive the development of logistics industry in the surrounding cities, the formation of Shijiazhuang logistics circle; Tangshan to accelerate the improvement of transport infrastructure network, and actively undertake the capital city logistics hub function, to undertake the logistics of Qinhuangdao, Chengde two cities transit, and radiation drive the development of logistics industry in the surrounding cities, the formation of Tangshan logistics circle.

The scale and capacity of logistics resources such as transportation infrastructure network, logistics parks and warehousing and transfer equipment should be allocated differently between hub and non-hub node cities to avoid the waste of resources caused by the duplicated allocation of logistics resources; in the same level of logistics hub cities, the level of logistics infrastructure construction in cities with weaker logistics supply capacity such as Shijiazhuang and Tangshan should be strengthened to weaken the imbalance of infrastructure allocation in the same level of hub cities. Based on the network layout model among the nodes, it is possible to determine the transfer mode of transportation between different cities in the metropolitan area, so as to minimize the overall transportation cost of the network and finally realize the optimal allocation of logistics resources in the metropolitan area.

REFERENCES