Connectivity analysis method of comprehensive transport network based on spanning tree

Qiao XIONG*
School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, (CHINA)
E-mail : xiongqiao@126.com

ABSTRACT

The network layout is the basis of transportation operation and the connectivity degree is one of important indexes to evaluate the network structure. When mapping out and selecting the comprehensive transport network planning scheme, in order to judge the connectivity of network, for its topological structure being composed of different modes of transport lines with the characteristics of complex graph, the counting of spanning tree in graph theory is applied to analyze the network connection quality. At the same time, to examine and evaluate the influence of the edges between nodes on the entire network connectivity, the calculation model of edge influence degree is proposed in terms of the changes in the number of spanning trees after removing one of edge as well as considering the traffic capacity, function and administrative status of links. The method based on spanning tree combined with traffic attributes is simple and easy to operate, and the example illustrates the effectiveness. It can provide decision-making basis for the evaluation and comparison of connectivity performance of comprehensive transport network layout schemes.

KEYWORDS

Comprehensive transportation; Transport network layout; Complex graph; Connectivity; Spanning tree; Edge influence degree;
INTRODUCTION

Comprehensive transportation system is composed of a variety of transport modes in a certain region, being a function system set of mutual coordination, reasonable use, completing passengers and goods transportation\[1\]. The formation of different modes of transportation network is the material basis for the establishment of comprehensive transportation system. The connectivity of transportation network is an important issue to be considered in the network performance evaluation during transportation network layout planning, which illustrates whether exists connection between each pair of nodes, number of connections, and going through steps of realizing the connection\[2,3\]. The existing literatures have researched the network connectivity or connectivity reliability of single traffic mode from the point of view of network topology, running situation and connectivity capability\[4-8\]. Ref\[9\] described the important degree of each vertex in the network through introducing vertex influencing degree and network influencing degree, with an simple algorithm to solve the node connectivity of network. In the environment of multi-modal transportation system, networks of different modes stacking together, the transportation process can be realized through spatial link structure of lines and equipments of various means of transport connected in series, parallel and series-parallel. In addition, because of all kinds of random factors, when natural disasters, equipment failure or congestion caused due to traffic accident or events leading to the line of one of transport modes is interrupted, the original network structure may be destroyed at any time. If there are other complementary and alternative modes or substitute circuitous path, it can effectively support the normal social economy activities. Therefore, it is very important to consider network connection and substitutability in single mode and mutual modes as well when comparing and evaluating comprehensive transportation network layout schemes.

DESCRIPTION OF COMPREHENSIVE TRANSPORT NETWORK AND CONNECTIVITY

The comprehensive physical network includes lines, stations, hubs and other infrastructures of five modes of transportation – railway, highway, waterway, aviation and pipeline. The traffic lines of various modes are as connecting lines, transportation hubs, stations and ports as the connection points and transition points, and they have a certain hierarchical structure. The comprehensive transportation network is represented as undirected connected graph of not containing ring, \( G = (V, E) \), where \( V \) denotes the nodes set of the network \( G \), including towns, hubs, stations and cross nodes, \( V = \{v_1, v_2, ..., v_n\} \), \( n \) is the number of nodes; \( E \) denotes edge set of \( G \), including railway, highway, waterway, air lines and pipelines, \( E = \{e_1, e_2, ..., e_m\} \), \( m \) is the number of connected edges.

Whether each node in the network can be communicated with others lies on the physical connections among nodes firstly. Between any two nodes if there are multiple paths can be accessed, then it shows the higher connectivity of the network is, and the stronger survivability. The transportation network connectivity is usually defined as\[3\]:

\[
K = \frac{L / \zeta}{nH} = \frac{L / \zeta}{\sqrt{nA}}
\]  

(1)

where, \( K \) is the connectivity of transport network; \( L \) is the total length of network in planning region; \( H \) is the average beeline between nodes; \( A \) is the area of planning region; \( \zeta \) is the deformation coefficient of network, which is related to the curve degree of lines and the geometric shape of nodes distribution. Ideally, \( \zeta = 1 \), so it can be approximately represented as:

\[
K \approx m / n
\]  

(2)

According to above equation, the connectivity is only related to the number of nodes and edges. If the networks have the equal number of nodes and equal number of connected lines, it will obtain the equal connectivity degree by approximate algorithm although they are different in connecting structure. There is also exact measurement to the connectivity of network\[2,10,11\]:

\[
\kappa(G) = \min \{|V_c|\}
\]  

(3)

\[
\lambda(G) = \min \{|E_c|\}
\]  

(4)

where, \( \kappa \) is the vertex connectivity degree, \( V_c \) is the vertex cut set; \( \lambda \) is the edge connectivity degree, \( E_c \) is the edge cut set.

To evaluate the exact connectivity condition of network layout, according to the definition of edge connectivity, it needs acquire the minimum edge cut number of network\[6\], through applying the maximum flow and minimum cut algorithm firstly to convert the graph into capacity network, and then calculate separately for each pair of nodes as in the single source and single sink network. The solution process is very complicated and the workload of calculation is great, even for large networks. Furthermore, in the comprehensive transport network there is usually more than one traffic line of various
transportation modes linking two nodes, railway, road or others. So it means that the network connection graph is not a simple graph, as having duplicated edges between nodes that is difficult to judge the overall connectivity. It also can be layered or adding virtual nodes, but this increases the number of nodes and edges, which further increases the calculation workload. During the evaluation of different layout schemes, if can quickly and accurately find out the pros and cons of connectivity quality of schemes, it will assist to improve the efficiency of comparison and selection of planning.

**EVALUATION OF COMPREHENSIVE TRANSPORT NETWORK CONNECTIVITY**

(1) Number of spanning tree of comprehensive transport network.

Spanning tree $T$ is a connected spanning subgraph of graph $G$, which connects all of the $n$ vertices together with $n-1$ edges to form a tree. The analysis of the spanning tree can be applied to research the connectivity of network$^{[12-15]}$. Because in spanning tree any two vertices are connected, different spanning trees of a network illustrate the connection in different ways. Therefore, the number of spanning trees$^{[16, 17]}$ can be used as a measurement to determine the network connectivity of comprehensive transportation.

If a network has at least one spanning tree, it indicates that the network is connected. The spanning tree of a connected graph is not unique in general, and all of its different spanning trees usually are a very large numbers. The more spanning trees the more possible paths choices between any pairs of nodes. Spanning trees are different indicating that the nodes in the network can choose different paths of transport lines, so it can determine the connection circumstances of the comprehensive transportation network from the point of macro view in planning. The better network connectivity can guarantee the higher network reliability, robustness and path flexibility, and this layout is the optimum scheme.

According to the definition and theorem of the spanning tree in graph theory, the method of determine the number of spanning tree any two vertices are connected, different spanning trees of a network illustrate the connection in different ways. Therefore, the number of spanning trees$^{[16, 17]}$ can be used as a measurement to determine the network connectivity of comprehensive transportation.

According to the definition and theorem of the spanning tree in graph theory, the method of determine the number of spanning tree for a undirected graph $G$ of comprehensive transportation network is as follows$^{[18]}$. Firstly, define the degree diagonal matrix of graph

$$F(G) = (f_{ij})_{n \times n}$$

where, $f_{ij} = \begin{cases} d(v_i), & i = j \\ 0, & i \neq j \end{cases}$

In the formula, $d(v_i)$ is the degree of node $v_i$, i.e. the number of each node connected with different modes of transportation lines.

Let $A = (a_{ij})_{n \times n}$, which denotes the extended adjacency matrix of $G$, and $a_{ij}$ is the number of edges connected between $v_i$ and $v_j$. $B = F(G) - A$. If let $v_i$ as a reference vertex, $B'$ is the matrix after deleting row $i$ and column $i$ from $B$, then the number of spanning tree of $G$, $\tau(G)$ is the determinant of matrix $B'$:

$$\tau(G) = |B'|$$

When the number of nodes is fixed, the more number of connected edges, the more paths between any two nodes. If the number of edges is same, but the layout ways of connecting is different, the counting of the spanning trees is not always the same. The connectivity analysis method based on spanning tree is can be applied to simple graph, and so is to complex graph.

(2) Determination of all spanning tree. Using the incidence matrix of a directed graph can easily obtain all the spanning trees of a network. Suppose $M = (m_{ij})_{n \times m}$ is the incidence matrix of a no-ring $n$-orders connected graph $D$, delete a row corresponding to any node $v_i$ in $M$, and can obtain a $(n-1) \times m$ matrix $M'$ whose rank equals $n-1$, where $M'$ is called the basic incidence matrix of $D$ corresponding to the reference node $v_i$. For directed connected graph there is the following theorem$^{[18]}$: take square matrix $C$ of $(n-1)$ order from $M'$, called the principal submatrix, $D_C$ is a generated subgraph with edges corresponded to columns of $C$, then the sufficient and necessary condition of $C$ being non-singular is that $D_C$ is a spanning tree of $D$. That is, if $|C|=1$ or $|C|=-1$, $D_C$ is a spanning tree of graph $D$.

Therefore, convert a undirected graph of transport network into a directed graph, using above method to obtain all the spanning trees of network, so as to further analyze the connected paths and quantities between each pairs of nodes in comprehensive transportation network. The flow chart of analyzing the spanning-tree-based connectivity of comprehensive transport network is shown in Figure 1.

**ANALYSIS OF EDGE INFLUENCE ON COMPREHENSIVE TRANSPORT NETWORK CONNECTIVITY**

In the regional comprehensive transportation network, any mode of transportation lines destroyed due to various random factors will influence the overall network connection structure and the operation of whole transport network. So the
connecting edge between nodes is the important factor to be considered for reliable network connectivity and alternative paths analysis [9, 20].

In order to study the influence of links on the whole transportation network connectivity, here put forward the concept of edge influence degree, which means the influence magnitude on comprehensive transport network connectivity of failure of one of interrupted edge. For a certain section of lines in network it also reflects its importance in the whole transport network structure. The edge influence degree is defined as the rate of reducing the number of spanning trees of comprehensive transport network graph after deleting one edge:

$$\frac{\Delta \tau(G_i)}{\tau(G)} = \frac{\tau(G) - \tau(G_i)}{\tau(G)}$$

(7)

where, $\tau(G_i)$ denotes the count of spanning trees of network after removing edge $e_i$. If the remaining graph is disconnected, let $\tau(G_i) = 0$.

Each edge linking up two nodes in itself is different in function, position and status in comprehensive transportation network. In general, the higher capacity and higher administrative rank of link the more important in network [7]. So, revise the edge influence degree with consideration of capability of passenger and freight transportation flow, and the administrative hierarchies of section of lines. Therefore,

$$IE_i = \alpha_i \cdot \frac{C_i \Delta \tau(G_i)}{\sum C_i \tau(G)}$$

(8)

where, $IE_i$ denotes the influence degree of edge $i$; $\alpha_i$ denotes the administrative rank coefficient of edge; $C_i$ denotes the traffic capacity of each edge, expressed by the capable of carrying passenger and freight flow volume.

**EXAMPLE**

Figure 2 shows two different connection planning layouts of a certain regional comprehensive transportation network that consists of 6 nodes to be connected. They both have 11 edges. Numbering of all vertices and edges is shown in the Figure. In the network, the connecting line between the nodes includes railway, highway and waterway. If calculate connectivity according to Eq. (2), the result is the same, but in fact, the connection effect is different.
In Figure 2 (a), the extended adjacency matrix of graph $G_a$ and the diagonal matrix of degree respectively are:

$$A_a = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 2 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 2 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 2 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad F(G_a) = \begin{pmatrix} 6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

Therefore,

$$B_a = F(G_a) - A_a = \begin{pmatrix} 6 & -1 & -1 & -1 & -1 & -2 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ -1 & -1 & 4 & -2 & 0 & 0 \\ -1 & 0 & -2 & 4 & -1 & 0 \\ -1 & 0 & 0 & -1 & 3 & -1 \\ -2 & 0 & 0 & 0 & -1 & 3 \end{pmatrix}$$

If let $v_6$ as a reference node, according to Eq. (6) it can be obtained the number of spanning tree of graph $G_a$ is

$$\tau(G_a) = |B_a^*| = \begin{vmatrix} 6 & -1 & -1 & -1 & -1 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 4 & -2 & 0 \\ -1 & 0 & -2 & 4 & -1 \\ -1 & 0 & 0 & -1 & 3 \end{vmatrix} = 139$$

Similarly, to the connection layout mode $G_b$ in figure 2 (b), there is

$$B_b = F(G_b) - A_b = \begin{pmatrix} 6 & -1 & -1 & 0 & -2 & -2 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 3 & -2 & 0 & 0 \\ 0 & 0 & -2 & 4 & -2 & 0 \\ -2 & 0 & 0 & -2 & 5 & -1 \\ -2 & 0 & 0 & 0 & -1 & 3 \end{pmatrix}$$

Then it can be obtained the number of spanning trees of $G_b$ is $\tau(G_b) = |B_b^*| = 76 < \tau(G_a)$.

As can be seen, if only from the point view of network connectivity, the layout of $G_a$ is better than $G_b$, which has the better network connection structure, more reliability and stronger survivable. Further analyzing the graph $G_a$, convert it into any directed graph, and the incidence matrix is

$$M = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & -1 & -1 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Still make $v_6$ as a reference node, and the sixth row is deleted from $M$ then obtain $M'$. The number of the principal submatrix of $M'$ is $C_{11} = 462$, of them 139 being non-singular. For example,
are three principal submatrixes whose determinant value is not equal to 0. They respectively denotes taking \( \{e_1, e_2, e_4, e_7, e_9\} \), \( \{e_1, e_3, e_6, e_8, e_{11}\} \) and \( \{e_2, e_3, e_6, e_7, e_{10}\} \) as edges to form 3 different connected mode of spanning tree, which can be seen the transportation paths between nodes.

Interrupted lines mean to delete one or more links, which will reduce the connectivity of transport network. If the traffic on a certain disrupted line can be evacuated by other lines, and each transportation mode line can mutual securing, then the network connectivity is less affected. Suppose the capacity and level of administration of each edge as shown in TABLE 1. Calculate the changes of network spanning trees after removing each edge according to Eq. (8), and the edge influence degree is shown in TABLE 2.

### TABLE 1 : Capacity and administrative rank coefficient of edge

<table>
<thead>
<tr>
<th>Edge</th>
<th>Classification</th>
<th>( \alpha_i )</th>
<th>( C_i ) (million tons/year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e_1 )</td>
<td>Freeway</td>
<td>1</td>
<td>5000</td>
</tr>
<tr>
<td>( e_2 )</td>
<td>Second class highway</td>
<td>0.2</td>
<td>950</td>
</tr>
<tr>
<td>( e_3 )</td>
<td>Freeway</td>
<td>1</td>
<td>3000</td>
</tr>
<tr>
<td>( e_4 )</td>
<td>First class highway</td>
<td>0.3</td>
<td>2800</td>
</tr>
<tr>
<td>( e_5 )</td>
<td>Class IV channel</td>
<td>0.2</td>
<td>2000</td>
</tr>
<tr>
<td>( e_6 )</td>
<td>Class I railway</td>
<td>1</td>
<td>4000</td>
</tr>
<tr>
<td>( e_7 )</td>
<td>Freeway</td>
<td>1</td>
<td>7000</td>
</tr>
<tr>
<td>( e_8 )</td>
<td>First class highway</td>
<td>0.5</td>
<td>3500</td>
</tr>
<tr>
<td>( e_9 )</td>
<td>Freeway</td>
<td>1</td>
<td>7300</td>
</tr>
<tr>
<td>( e_{10} )</td>
<td>Freeway</td>
<td>1</td>
<td>5000</td>
</tr>
<tr>
<td>( e_{11} )</td>
<td>Class I railway</td>
<td>1</td>
<td>4000</td>
</tr>
</tbody>
</table>

### TABLE 2 : Changes in the number of spanning tree and calculation of edge influence degree

<table>
<thead>
<tr>
<th>Delete edge</th>
<th>( T (G_0) )</th>
<th>( \Delta \tau (G_0)/\tau (G_0) )</th>
<th>( IE_{ui} ) (Normalized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e_1 )</td>
<td>55</td>
<td>0.604</td>
<td>0.166</td>
</tr>
<tr>
<td>( e_2 )</td>
<td>55</td>
<td>0.604</td>
<td>0.006</td>
</tr>
<tr>
<td>( e_3 )</td>
<td>81</td>
<td>0.417</td>
<td>0.069</td>
</tr>
<tr>
<td>( e_4 )</td>
<td>89</td>
<td>0.360</td>
<td>0.017</td>
</tr>
<tr>
<td>( e_5 )</td>
<td>89</td>
<td>0.360</td>
<td>0.008</td>
</tr>
<tr>
<td>( e_6 )</td>
<td>83</td>
<td>0.403</td>
<td>0.089</td>
</tr>
<tr>
<td>( e_7 )</td>
<td>65</td>
<td>0.532</td>
<td>0.205</td>
</tr>
<tr>
<td>( e_8 )</td>
<td>79</td>
<td>0.432</td>
<td>0.042</td>
</tr>
<tr>
<td>( e_9 )</td>
<td>66</td>
<td>0.525</td>
<td>0.211</td>
</tr>
<tr>
<td>( e_{10} )</td>
<td>86</td>
<td>0.381</td>
<td>0.105</td>
</tr>
<tr>
<td>( e_{11} )</td>
<td>86</td>
<td>0.381</td>
<td>0.084</td>
</tr>
</tbody>
</table>

The line with larger edge influence degree impacts on the entire network connectivity and reliability greatly, which is the network vulnerability, so to these sections it should strengthen maintenance and management. It can be seen from the TABLE that for layout \( G_0 \), the failure of the ninth edge and the seventh edge both greatly influence the connectivity of the network. Take \( e_9 \) for example, along with the change rate of its capacity, the change of edge influence degree is shown in Figure 3. If the traffic capacity of other edges change, the impact on the sorting of \( IE \) is shown in Figure 4, where the change rate is 30%.
Figure 3: Change of $IE_i$ with traffic capacity change of $e_9$

Figure 4: Sorting of $IE_i$ with traffic capacity change

It can be seen from the figure, the higher change rate of capacity the greater change of edge influence degree. To $e_9$, increase of capacity of $e_1$ and $e_7$ impact on the sorting of its edge influence degree is relatively great.

CONCLUSIONS

A variety of modes of transportation lines form parallel and circuitous connections between nodes, so that different modes of regional comprehensive transport network can substitute and complement for each other, and provide more path selections for users. And also can make up for the interrupt of some nodes or the whole network when some sections are disrupted, so as to guarantee the operation of the whole network and the normal passenger and freight transportation. The connectivity of network is the foundation of the unblocked transportation flow and the reliability analysis. To construct a rational comprehensive transportation network, in the evaluation of the connectivity of network layout schemes, applying the method based on spanning tree analysis in graph theory can generally grasp the pros and cons of comprehensive transport network layout in connectivity with considering the characteristics of lines connection of various modes. At the same time, it can through the analysis of the changes of number of spanning trees combined with the line capacity and administrative rank to evaluate the influence of edge on the entire network connectivity. The given example shows that the method is effective, and can provide some assistance for the accurate and efficient evaluation and decision-making on the comprehensive transportation network layout schemes. On this basis, the routes between nodes in the network and selection, reachability, modes selection, reliability and so on can be further researched. The connectivity is only one of attributes of network, so it need combine with other indexes to carry out comprehensive evaluation and selection of transport network layout planning.

REFERENCES

[18] Zhang Xiandi, Li Zhengliang; Graph Theory and its Application, High education press (2005).