Complexity Analysis of Qingdao’s Public Transport Network

Li Liu, Ran Li, Fengjing Shao, and Rencheng Sun
Information and Engineering Department, Qingdao University, Qingdao, China
Email: liuli19851216@163.com, liranne623@tom.com, sfj@qdu.edu.cn, qdsunstar@163.com

Abstract—Public transport network reflects the whole city’s public transport environment, the geometrical property of which plays an important role on the public transportation. In this paper, the theory of complex network was used to study the public transport network of Qingdao and construct the transportation transfer network based on the stops. In order to study this network, some calculations are made to get topology such as degree distribution, clustering coefficient, average path length and so on. It is found that clustering coefficient couldn’t show the density of the network clearly and has negative correlation with the numbers of the bus line passed that stop. The result shows that the public transport network of Qingdao is a connected network with a rather small average path length and a big clustering coefficient and its degree distribution follows the exponential distribution. All of these prove that the Qingdao public transport network exhibits the small-word properties and Qingdao’s Public Transportation is rather convenient.

Index Terms—complex network, transportation transfer network, degree distribution, negative correlation, small-world

I. INTRODUCTION

With the rapid urbanization in the whole world, the traffic jam has become one of the main problems. The public transport, as an important transportation tool, has been given much attention from various fields. In recent years, many scholars advocate to develop the public transport, hence how to evaluate the public transport’s performance and how to improve it further becomes a crucial issue.

The urban public transport system is a typical complex network [1], so using the theory of complex networks to study its topological structure and property can be useful to learn and improve the public transport system. Recently, many studies have focused on the complexity of the urban public transport system by the use of complex networks. Latora and Marchiori [2][3] introduced the definition of efficiency coefficient and applied it to study the Boston subway. Sen et al. [4] concluded that India’s railway network exhibited small world properties. Similar properties were reported by Seaten and Hackett [5] in a study of railway networks in Boston and Vienna. Jiang and Claramunt [6] found that the topological networks of streets exhibited small world properties but were not scale-free networks. Li and Cai [7] showed that the topological structure of the air traffic network of China (ANC) had two key characteristics of small world networks, a short average path length and a high degree of clustering. Guimera Roger et al. [8] showed that the worldwide air traffic network was a scale-free and small world network. Unlike other networks, the world’s air traffic network is influenced by geography and politics. In our country, many researches have also been done on complex network to analyze traffic problems. Wu Jianjun et al. [9] concluded that the urban transit system in Beijing is a scale-free network using statistical analysis. Zhao Jinshan et al. [10] introduced three kinds of models to study the properties of the public transport in Beijing. Chen Liping et al. [11] did empirical research on public transport network of Nanjing.

Qingdao is located at the south of Shandong Peninsula, known as a summer resort with a rich historical background and natural beauty. For bus is the only public transportation tool, the layout of the buses is logical or not directly influences the citizen’s daily life. Therefore, in this paper, we investigate Qingdao’s public transport network, the stations in this network are studied by the use of complex networks. Then analyze the complexity of this network to discuss the performance of the public transport system, which can be used to give good proposal to optimize the bus network.

II. DEFINITIONS OF TRANSPORTATION TRANSFER NETWORKS

Definition 1: A bus line is an ordered set of bus stations. Denoted as: \( L(D,R) \) [12], here \( D=\{s_1, s_2, ..., s_n\} \) is a set of bus stations. \( R=\{N_1|N_1=\{s_{i-1}, s_{i}\} \mid s_{i-1}, s_{i} \in D, i=2,3,...n\} \)

Definition 2: The transportation transfer network can be seen as an undirected graph which consisted of vertices which are a set of bus stations, denoted as \( G(V,E) \), where \( V \) is a set of public bus stations. If \( s_i, s_j \in V \), then \( e_{ij} \) is the edge between \( s_i \) and \( s_j \). If only if, and only if, \( s_i, s_j \in N \), which means that there is the same bus passing through the No.\( i \) bus station and the No.\( j \) bus station.

Definition 3: The degree of vertex \( i \) in the transportation transfer network can be denoted as \( VD_i \), which is the number of bus stations that can be reached directly from the No.\( i \) bus station. In fact, the higher the degree is, the more the number of the direct linked stations is, which means that it is much more convenient to take buses from this station. So stations with higher degree are always transportation hubs.

Definition 4: The cumulative probability can be defined as follows:
\[ P_k = \sum_{k=\infty} P_k \]  

(1)

It means the probability that a vertex has more than \( k \) edges, which is equivalent to the ratio of the number of vertexes whose degree is more than \( k \) in the network to the total number of vertexes.

Definition 5: Suppose \( G(V,E) \) is a subgraph of \( G(V,E) \), here \( V_i \) is a set of bus stations including \( S_i \) and its neighbors(with a same bus). \( E_i \) is the set of the edges connecting the bus stations in \( V_i \).

Clustering coefficient of bus station \( S_i \) is defined as:

\[ \text{CC}_i = \frac{2|E_i|}{|V_i||V_i|-1} \]  

(2)

Definition 6: The shortest path length between \( S_i \) and \( S_j \) can be defined as the number of bus stations that we have past from \( S_i \) to \( S_j \). The maximum length of the shortest path is called diameter of the network.

In the transportation transfer network, the path length in a same bus line is one because the stations in the same bus are connected fully. The distance from one bus station to another bus station minus 1 equals the number of transfers between these two stations.

III. COMPLEXITY ANALYSIS

A. Dataset

Data set used in this paper comes from the web site of Qingdao information (http://www.qd.sd.cn/). Up to the year 2009, this dataset include 130 bus lines and 864 bus stations. Here those bus lines leading to the Qingdao outskirts are excluded. We process this dataset as follows:

1) We only selected dataset which come from the down line because most data in up line and down line are the same, so the network can be abstracted as an undirected graph.

2) We deleted the repeated bus stations if a bus line is ring-shaped.

3) The stations with the same name were regarded as the same station and those with different names were treated as different stations.

B. Statistic Analysis Of This Network

The topological features of the public transportation network can be studied by analyzing some basic geometric parameters, such as degree distribution, clustering coefficient, average path length and so on. Here we calculate these parameters to evaluate the public transport’s performance.

1) Analysis of vertex degree: According to definition 3, the vertex degree implies the number of bus stations which you can go directly to without transfers. The vertex degree indicates the role of the bus station in the actual public transportation network. A more convenient and efficient public transportation network should focus on the construction and management of bus stations with high degree. By calculation, the maximum degree, the minimum degree and the average degree of this network is 400, 9 and 80.

The station Fuzhou Road has the maximum degree and 32 routes pass this station. Next station is Licun with the degree of 379. Then is the Taidong station, whose degree is 347. Here the vertices we defined were stations with the same name. Fuzhou Road is a business center, Licun has a gigantic residential section and Taidong is the largest shopping center. All of these are prosperous parts in Qingdao and transportation is convenient.

In order to analyze the degree of vertex, we got the chart that illustrated the degree distribution. As illustrated in Fig.1.

![Figure 1. Distribution of degree value](image1)

Fig.1 shows the degree distribution of public transportation transfer network of Qingdao. We can see that the distribution of vertex degree is far from the Poisson distribution and slants to the right obviously. It is rather difficult to measure this curve since there is oscillation at the end. In order to analyze the distribution of the degrees better, we use the cumulative probability distribution. As illustrated in fig.2 in the semi-logarithm coordinate, we can see that, except several vertices in the middle, the change of vertex degree is rather uniform. Therefore we can fit it with a line. We used linear fitting tool of R to get a line \( y = -0.0164x + 0.02187 \). So the degree distribution of the public transportation transfer network of Qingdao approximately follows the exponential distribution.

2) Clustering coefficient: According to definition 5, the clustering coefficient is used to measure the intensity of the network. In the transportation transfer network, it reflects the connectedness between this station and the bus stations connected to it, that is to say, the clustering coefficient shows accessibility between bus stations without transfers. The clustering coefficient is 1 if, and
only if the network is full connected, otherwise, it is less than 1.

The transportation transfer network is consisted of many full connected subgraphs and all the stations along one route are linked. Fig. 3 illustrates the clustering coefficient distribution of vertex. According to it we can see that there are many stations with high clustering coefficient. To a certain degree, it shows that many stations in Qingdao transport system are passed through by only one route. These stations are either in remote place or added recently due to the urban planning. Accordingly, the remaining stations are passed through by more than one route, so it is easy to cause traffic jam. Therefore, if we make good use of the stations only passed by one route, the traffic pressure can be eased to some extent.

![Figure 3. Distribution of clustering coefficient](image)

According to the analysis fore-mentioned, the traditional clustering coefficient distribution used in the complex network is not applicable to the transportation transfer network. E.g., the clustering coefficient of stations passed by only one route is 1 since they are in the full-connected subgraph. Licun is a critical transportation hub, passed by 25 routes, but its clustering coefficient is rather small.

As illustrated in fig. 4, we can see the relationship between the clustering coefficient and the number of the bus lines passed that station. In fig. 4, x-axis stands for the number of the bus lines passed through the station, y-axis stands for the clustering coefficient of the corresponding station.

From the figure we got that the more the number of bus lines passed through the station, the less the density is. So the clustering coefficient has negative correlation with the numbers of the bus line passed that stop.

![Figure 4. Relationship between the clustering coefficient and number of lines passed that station](image)

3) Average path length: According to definition 6, the shortest path length is the number of edges in the shortest path between these two vertices. The distance from one bus station to another bus station minus 1 equals the number of transfers between these two stations. So the average network path length is an important indication of the network accessibility. The diameter of the Qingdao transportation transfer network is 4, the average network path length is 2.135, which means that we need to transfer only 1 or 2 times to reach the destination when we go out. Fig. 5 shows the probability of the number of transfers from one station to another. According to it, we can see that transportation in Qingdao is rather convenient.

![Figure 5. Distribution of number of transfers](image)

Table 1 lists the basic parameters for the Qingdao transportation transfer network and random network with the same vertices. From the table, we got that Qingdao’s public transport network has higher average weight and shorter average path, so it is a small-word network.

<table>
<thead>
<tr>
<th>Network</th>
<th>Average Path Length</th>
<th>Clustering Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transfer network</td>
<td>2.135</td>
<td>0.433</td>
</tr>
<tr>
<td>Random network</td>
<td>1.9</td>
<td>0.101</td>
</tr>
</tbody>
</table>

### IV. CONCLUSIONS

In this paper, we analyze the statistic features of public transport network of Qingdao. The results show that the public transport network of Qingdao is a connected network. The degree distribution approximately follows the exponential distribution, we used a line $y = 0.0164x + 0.021817$ to fit the cumulative degree distribution of the network in semi-logarithm coordinate. It is found that clustering coefficient couldn’t make clear the density of the network and has negative correlation with the numbers of the bus line passed that stop. Compared with the random network with the same scale, we found that the public transport network of Qingdao is a network with a rather small average path length and a bigger clustering coefficient. All of these prove that the Qingdao public transport network exhibits the small-word properties. Besides, we know that the average network path length is 2.135, which means that we need to...
transfer only 1 or 2 times to reach the destination when we go out. As a whole, Qingdao’s transport network is rather reasonable and convenient.

REFERENCES