*Journal of Information Hiding and Multimedia Signal Processing* ©2016 ISSN 2073-4212
**Ubiquitous International** Volume **7**, Number **3**, May **2016**

# Urban Road Network and Taxi Network Modeling Based on Complex Network Theory

Xin-Ling Guo
School of Aeronautics and Astronautics
Zhejiang University
Hangzhou,310027,P.R.China
969326321@qq.com

Zhe-Ming Lu*
School of Aeronautics and Astronautics
Zhejiang University
Hangzhou,310027,P.R.China
*Corresponding Author, zheminglu@zju.edu.cn

Received November, 2015; revised February, 2016

Abstract. *By introducing complex network modeling and analysis methods, researches on undirected unweighted network models for the roads inside the fifth ring of Beijing city are made in this paper based on traffic and map data. In order to construct directed weighted network models for the roads inside the fifth ring of Beijing city we come up with two modeling methords TTI and TTIS based on complex network theory and data filtering based on the trajectories data of more than 10,000 taxis from February 2nd to 8th, 2008. We find that the topologies of these undirected unweighted networks had the universal properties of random network and scale-free network, and the taxi networks also can be regarded as random networks.*
**Keywords:** Complex network, Traffic network, Network modeling, Topological characteristics.

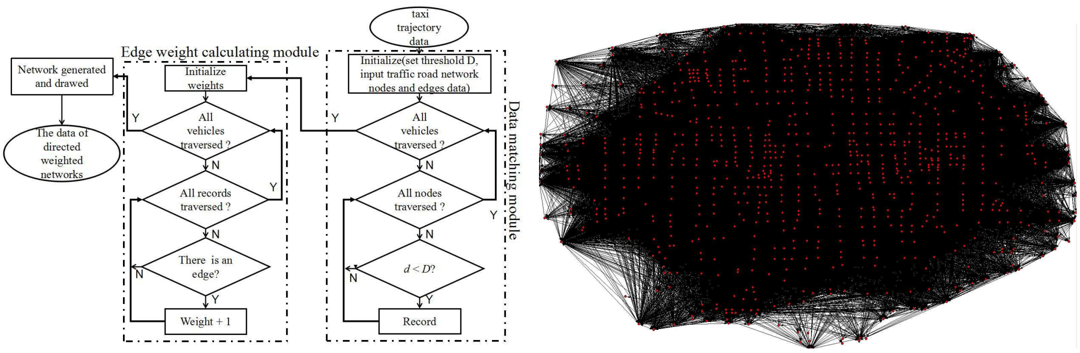
1. **Introduction.** As the application of complex network [1] in all walks of life research, such as social networks [2], the theory have gradually been introduced into the transportation field [3]. More and more scholars have carried out relative research on the characteristics of various transportation networks such as the urban road traffic network [4], rail transit network [5], railway network [6] and transit network [7].And current study in complex network has almostly been focused on modeling [8] and so on. However there is few scholars study urban road traffic network based on taxi system. In this paper, based on the network model established by the primal approach, using the theory of complex networks, taking the taxi system as the research object, we study the characteristics of urban road traffic networks and urban taxi traffic networks, in order to provid certain theoretical basis on solving city traffic problems and the development of urban traffic network. We firstly need to establish urban road network models based on complex network theory.There are two prevailing modeling methords:

Methord 1: The primal approach [9] is intuitive, where intersections become nodes and streets become edges connecting nodes. The advantages of the original method is that the raw data can be obtained directly from the geographic information system, the method is simple and the data source is convenient.

Methord 2: The dual approach [10] is a kind of transformation methods based on the primal graph, where roads as a network node, intersection is regarded as the edge of the network. Although dual method using topological distance,more suitable for network topology analysis than the primal approach,but its drawback is that it ignores the actual traffic network location information, which is difficult to reflect the actual distribution of network nodes. Therefore, most of the research on urban road network is still using the original method [11].

2. **Modeling and Analysis of Beijing Road Networks.** In this section, we establish the undirected unweighted network models of the 2nd, 3rd, 4th, 5th rings roads of Beijing on the basis of the primal approach. Then we draw these toplological graphs on the complex network analysis software Pajek. At last we study the average degree of each ring network, network cluster coefficient, the network diameter, average shortest path length and other network characteristics on Pajek, VS2010 and Matlab2014 software platform.

2.1. **Modeling Approach and Method of Beijing Road Networks.** The traffic network modeling data of this section comes from Beijing traffic map data of Baidu Maps and Google Maps, and the traditional primal approach is adopted to build these network models. We count intersections as nodes, streets as edges connecting nodes, as shown in Figure 1.



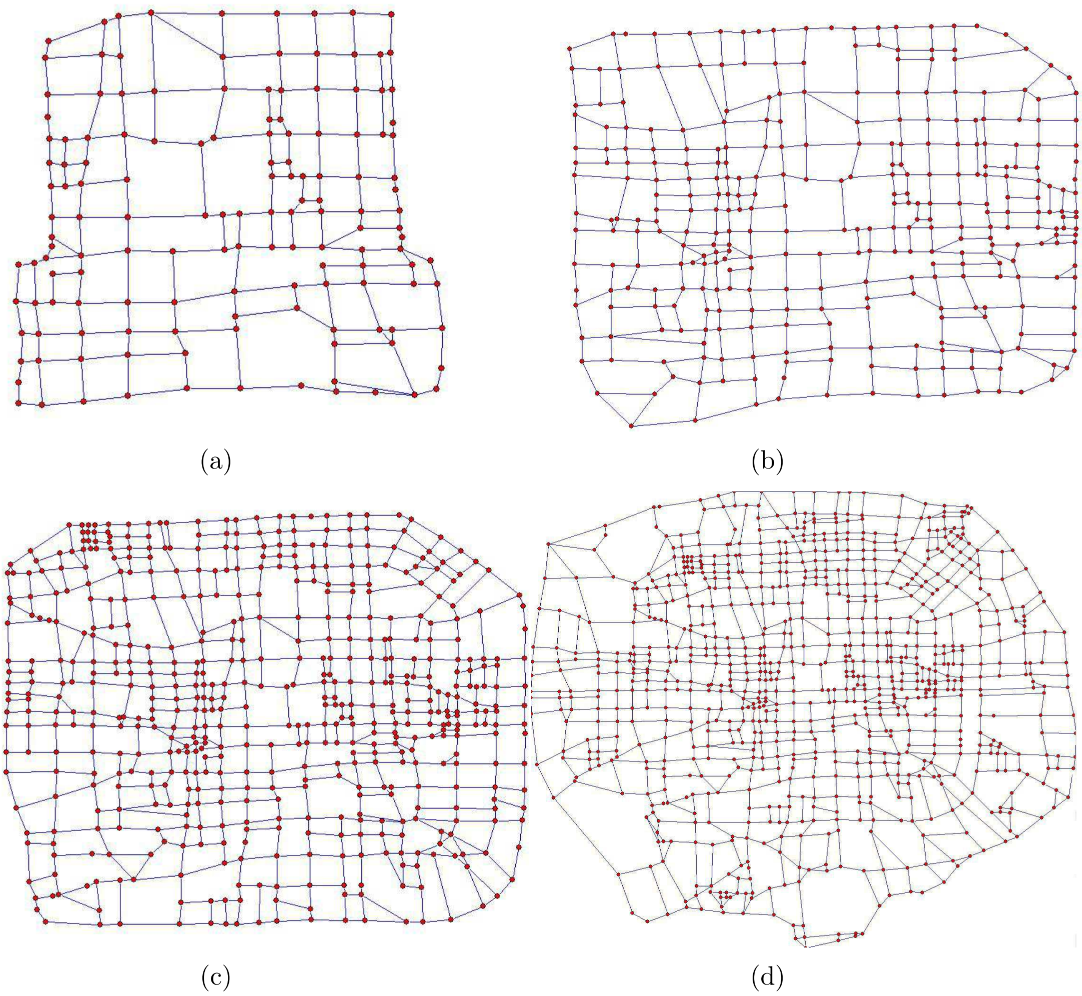

(a) (b)

FIGURE 1. Example of the primal apporach

2.2. **Modeling Result of Beijing Road Networks.** The undirected unweighted network model for the roads inside each ring of Beijing city is shown in Figure 2.

As Figure 2 shows, the undirected unweighted networks for the roads inside the each ring of Beijing are all in the form of "chessboard" networks [4], which is consistent with the actual situation: in the early urban construction, the shape of the urban road traffic network was directly influenced by urban spatial layout; with the development of the social economic, the concentration of urban population and the expansion of city size, road network also follow this trend to expand outward layer by layer according to chessboard shape. Under the influence of landform, the road network do not form regular "checkerboard", but "chessboard" network overall.

2.3. **Characteristics of Beijing Road Networks.** According to the undirected unweighted networks for the roads inside the each ring of Beijing city established by section 2.2, based on Pajek, VS2010 and Matlab2014, we get these networks static characteristics as shown in Table 1.



(a) (b) (c) (d)

Figure 2. The road network inside of the each ring of Beijing

Table 1. Characteristic values of the road network inside each ring

| Network's ring level | Nodes | Edges | The average degree | The average cluster cofficient | The average shortest path length | The network diameter |
|---|---|---|---|---|---|---|
| 2nd | 144 | 235 | 3.26 | 0.009259 | 7.743007 | 18 |
| 3rd | 322 | 548 | 3.4 | 0.020807 | 10.958495 | 25 |
| 4th | 547 | 946 | 3.46 | 0.02011 | 13.779088 | 32 |
| 5th | 871 | 1517 | 3.48 | 0.029698 | 16.517352 | 38 |

As Table 1 shows, the network size is increasingly faster and faster with the increase of ring number: Nodes inside of the 3rd ring is 178 more than inside of the 2nd ring, edges inside of the 3rd ring is 313 more than inside of the 2nd ring; Nodes inside of the 4th ring is 225 more than inside of the 3rd ring, edges inside of the 4th ring is 398 more than inside of the 3rd ring; Nodes inside of the 5th ring is 324 more than inside of the 4th ring, edges inside of the 5th ring is 571 more than inside of the 4th ring; The average degree increases with the ring number, which shows that the roads of intersections increase with the development of the traffic network.

Each of these network has a very small cluster coefficient, which is the reaction of network compact degree, and it is in conformity with the actual situation of the traffic network in Beijing. As Figure 2(a), Figure 2(b), Figure 2(c) and Figure 2(d) show, there are a lot of quadrilaterals, while few triangles, which indicate that they all have a very small network cluster coefficient; Compared with their network diameter, every network has a smaller average shortest path length.

We all know that most of the nodes' degree are very close, and all near to the average degree of $<k>$ in random networks [12]; small-world networks [13] have the characters of higher cluster coefficient and shorter average shortest path length; the degree displays a power-law distribution in scale-free networks [14]. From what has been discussed above and the characteristics of various complex networks, we figure out that none of these networks is a small-world network and each of these networks can be regarded as a random network.The randomness of the road traffic networks for inside each ring of Beijing can be attributed to these random factors in the development of the urban road traffic network, such as population relocation, the developers' location, changes in government policies, etc.

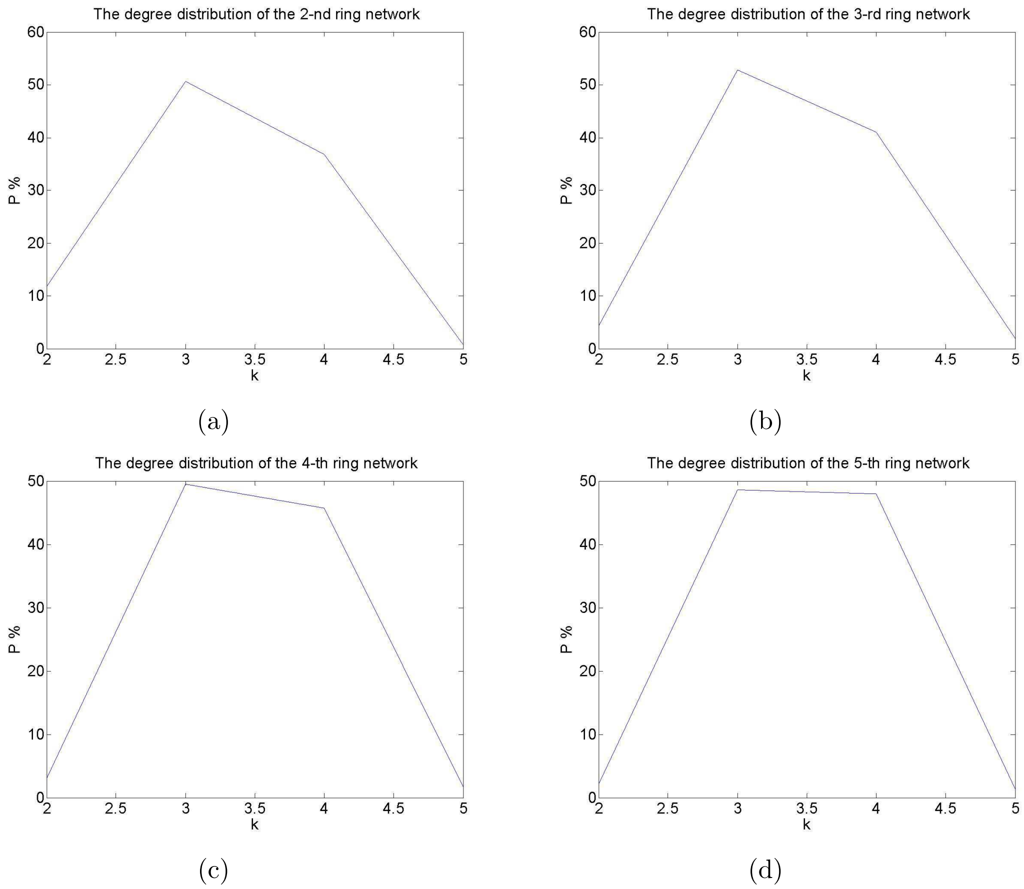



FIGURE 3. Degree distribution of the road network inside each ring

2.4. **Degree Distributions of Beijing Road Networks.** Based on the undirected unweighted network models for the roads inside the each ring of Beijing established by this section, using VS2010 and Matlab2014 software platform, we get the degree distributions of these networks (see Figure 3). It can be seen that not all nodes have a same degree value in every network, which indicates that none of these networks is a rule network; Most nodes' degree of each network are the same, according with the characteristic of random networks;These networks' maximum degree is 5 and minimum degree is 2, which shows that both in the early construction of urban road network and the later development of urban road network, the minimum path number of intersection the maximum path number are constant; After ignoring the nodes with degree value 2, these degree distributions would all present a power law distribution which we called loose scale-free feature.

3. **Modeling and Analysis of Beijing Taxi Traffic Networks.** Real road is directional. There are one-way road and two-way road according to the direction of roads. Real traffic network road has a weight, it can represent the actual length of the road, the traffic flow of the road, the passage time of vehicles, etc. We should construct a directed weighted network model to describe the real traffic network more accurately. This section based on the undirected unweighted traffic road networks established by the previous section, using Matlab2010b software platform and the method of matching traffic flow construct and analysis directed weighted Beijing taxi traffic networks. We develope two methords TTI (the modeling based on taxi trajectory and network intersections) and TTIS (the modeling based on taxi trajectory and network intersections and streets) to model Beijing taxi traffic networks.

3.1. **Modeling Approach and Method of Beijing Taxi Traffic Networks.** We choose vehicle flow as the edge weight for convenience.

There are two basis of our modeling:

Basis 1: Undirected unweighted road traffic networks of Beijing city established by section 2.3.

Basis 2: GPS trajectory data of 10,335 taxis in a week from web "Urban Computing-Microsoft Researc"
(data source: http://research.microsoft.com/apps/pubs/?id=152883). There are 15 million records and 9 million kilometers distance in this data set. We would match the latitude and longitude of every taxi trajectory record to the latitude and longitude of each node to establish the directed weighted Beijing taxi traffic network models.

The first method TTI (the modeling based on taxi trajectory and network intersections) is as follows:

We assume the initial weight(flow) of each edge is 0 at first.

Step 1: Matching the latitude and longitude of every taxi trajectory record to the latitude and longitude of each node to get the number of the node that this taxi at this moment passed through. The matching process is as follows.

Step 1.1: Set the matching threshold is $D$ m. For the convenience of matching, we assume that: if a certain moment the distance of this taxi and node $v_i$ is less than the matching threshold distance $D$, the taxi passed node $v_j$ at the moment. when there are several nodes matching the condition, we selected the nearest node.

Step 1.2: Traverse the first record of the first taxi, get the matching node number.

Step 1.3: Traverse next record using the method of Step 1.1 till every track record of every taxi is been traversed.

Step 2: Node $v_i$,$v_j$ matched to two adjacent records of the same taxi should have an edge, which presents that this taxi passed $v_i$ and $v_j$ sequentially. So we add a directed edge from $v_i$ to $v_j$ the first time we dectct that. Let this edge flow plus 1. Then Traverse next record till every record is been traversed.

Step 3: Estabilsh the directed weighted taxi traffic network models based on road traffic network data and the weight data we get, using Pajek. We only consider the directed weighted taxi traffic network model for the roads inside the 5th ring of Beijing in the flowing study.

This method does not take the actual traffic streets into account, so the directed weighted network we get are maybe unrealistic. We improve the method TTI and get the second method TTIS (the modeling based on taxi trajectory and network intersections and streets). We improve the original Step 2 and Step 3 based on the original step 1 as flows:

Step 2: Node $v_i$ and $v_j$ matched to two adjacent records of the same taxi should have an edge, which presents that this taxi passed $v_i$ and $v_j$ sequentially. If there is an edge between $v_i$ and $v_j$ in network of figure 2(d), this edge flow plus 1; Otherwise doing nothing. Then Traverse next record till every record is been traversed.

Step 3: Add the corresponding edge flow to the undirected unweighted Beijing network models to get the directed weighted taxi traffic network models. We only consider the directed weighted taxi traffic network model for the roads inside the 5th ring of Beijing in the flowing study.

The method TTI and TTIS of constructing the urban traffic network models are shown in Figure 4 and Figure 5 respectively.

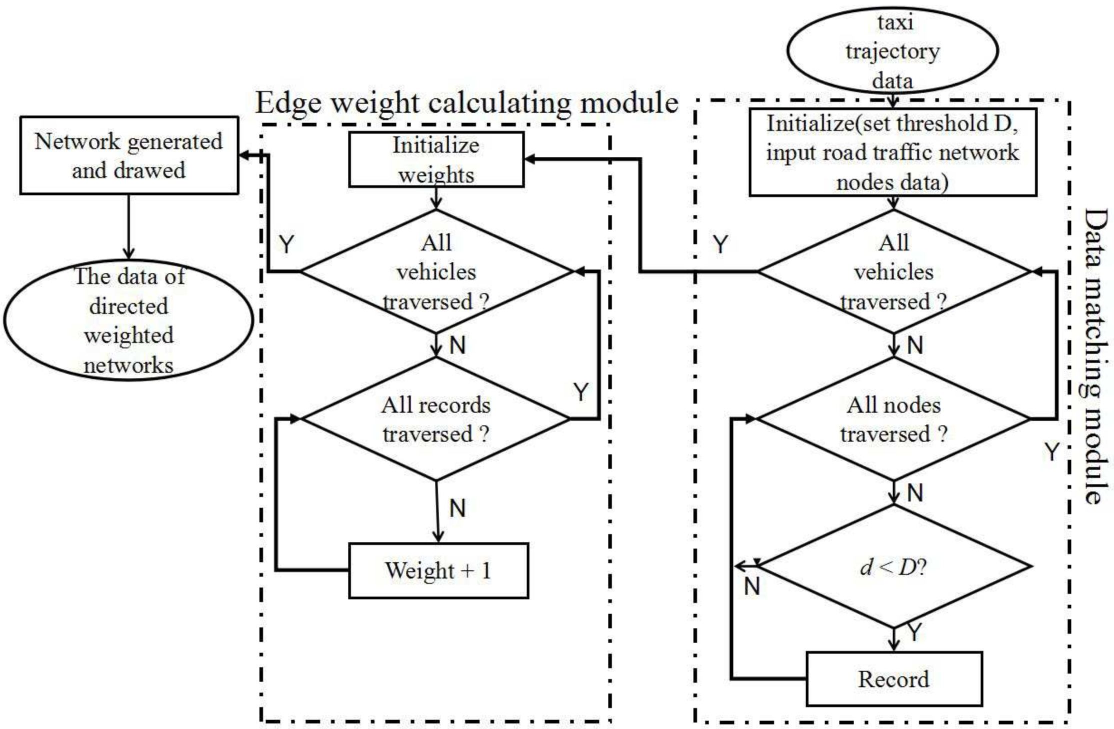



Figure 4. Flow chart of modeling method TTI

3.2. **Modeling Result of Beijing Taxi Traffic Networks.** The shortest distance between two adjacent nodes is about 190 meters approximately and distance between two adjacent nodes is in 400-600 meters mostly in Baidu maps and Google maps. So we set the matching threshold $D$ as 200 meters for a accurate description of the edge weight. The directed weighted Beijing taxi traffic network with $D$ of 200 m based on the method TTI is shown as Figure 6.And Beijing taxi traffic network with D of 200 m based on TTIS method is shown as Figure 7

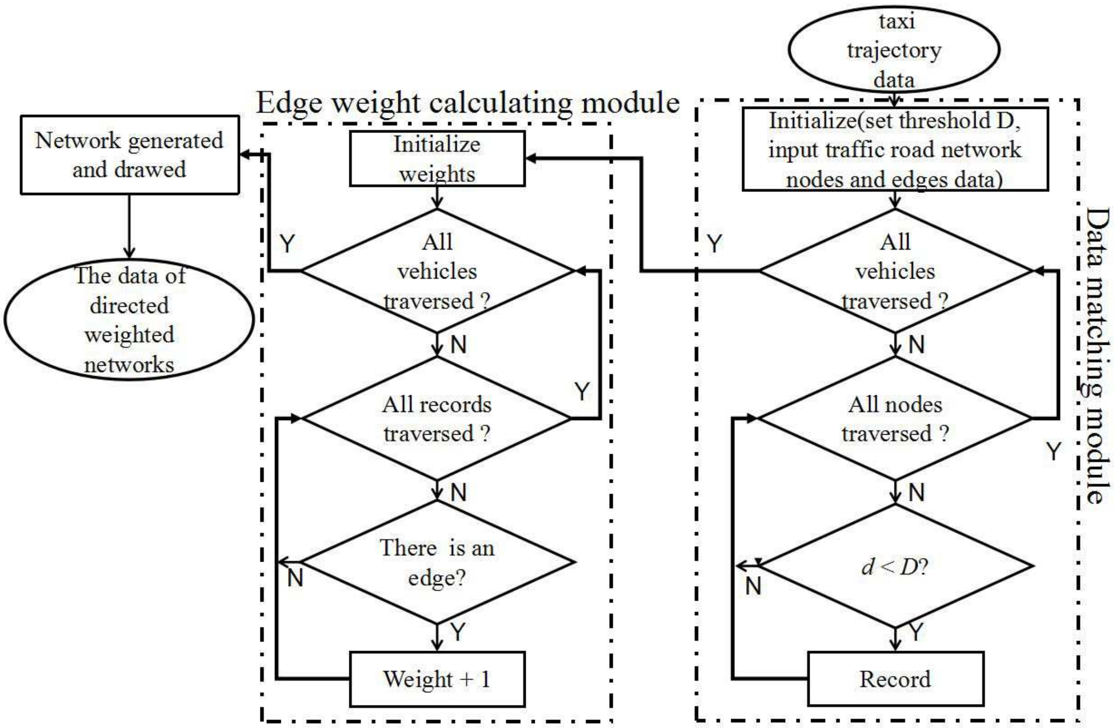



FIGURE 5. Flow chart of modeling method TTIS

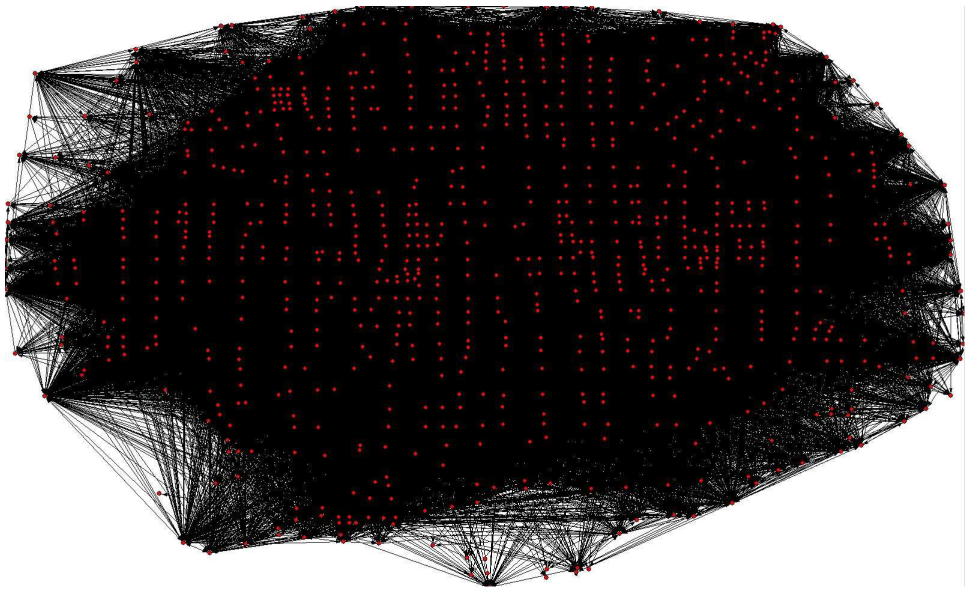


FIGURE 6. Beijing taxi traffic network with $D$ of 200 m based on TTI

From Figure 6 and Figure 7, we can see that Figure 6 is not sound to realistic Beijing road traffic network compared with Figure 7, so we select the modeling method TTIS as our modeling method.

However, the directed weighted Beijing taxi traffic network is not connected when the matching threshold is 200 m. In order to have a comparisonthe directed weighted Beijing taxi traffic network with $D$ of 300 m as shown in Figure 8. We call the directed weighted Beijing taxi traffic network with $D$ of 200 m as the directed weighted network 1 and the

FIGURE 7. Beijing taxi traffic network with $D$ of 200 m based on TTIS

directed weighted Beijing taxi traffic network with $D$ of 300 m as the directed weighted network 2. From Figure 8, we can see that the directed weighted network 2 is more connected than the directed weighted network 1.

3.3. **Characteristics of Beijing Taxi Traffic Networks.** Based on Pajek and VS2010 software platform, we get these networks' static characteristic values as shown in Table 2.

TABLE 2. Characteristic values of the Beijing taxi traffic networks

| Network | Nodes | Edges | Number of largest components | Size of the component |
|---|---|---|---|---|
| 1 | 871 | 2529 | 39 | 815(93.5%) |
| 2 | 871 | 2729 | 12 | 859(98.6%) |

From Table 2, we find that the size of the largest component of the directed weighted network 1 is 815, and it's relative size is 93.571%. The disconnectedness of the directed weighted network 1 due to the abnormal situation. In the traffic flow database,the change of nodes number matched to two adjacent records of the same taxi indicates that this

Figure 8. Beijing taxi traffic network with $D$ of 300 m

taxi pass the pair of nodes in a row, however the time interval of record is too long or the running speed of taxi is too fast to detect some nodes between the pair of nodes, which we called the abnormal situation. The only thing we can do is ignore the abnormal situation to optimize the result.There are 269267 pairs of nodes direct attached and 1180060 pairs of nodes indirect attached, so the ratio of abnormal data is 81.421%.

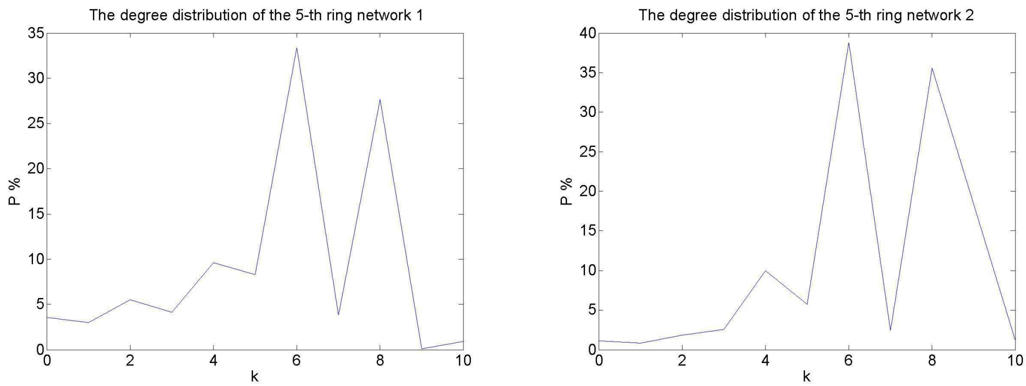
The size of the largest component of the directed weighted network 2 is 859, and it's relative size is 98.622%. There are 658980 pairs of nodes direct attached and 1981308 pairs of nodes indirect attached, so the ratio of abnormal data is 75.039%. The directed weighted network 2 is more connected than the directed weighted network 1 apparently.

From Table 3, we can also find that the average degree of the directed weighted network 2 is bigger, this is because the ratio of abnormal data of the network 2 is smaller and it's network traffic flow is more huge overall; The average shortest path length of the largest component and The network diameter of the largest component of the two networks are close, and they both have a short average shortest path length of the largest component; The average cluster coefficient of the largest component of two networks are close to 0. From the above analysis, we finally konw that neither of them is a small-world network and both of the two networks have the random network characteristics.

Table 3. Characteristic values of the Beijing taxi traffic networks

| Network | The average degree | The average cluster coefficient of the largest component | The average shortest path length of the largest component | The network diameter of the largest component |
|---|---|---|---|---|
| 1 | in:2.90<br>out:2.90<br>all:5.80 | in:0.03888<br>out:0.0402<br>all:0.0317 | 17.2674 | 48 |
| 2 | in:3.13<br>out:3.13<br>all:6.27 | in:0.05001<br>out:0.0517<br>all:0.04196 | 17.4935 | 46 |

3.4. **Degree Distribution of Beijing Taxi Traffic Networks.** Based on VS2010 and Matlab2014 software platform, the degree distributions of the two networks are as shown in Figure 9.




(a) (b)

Figure 9. Degree distribution of the Beijing taxi traffic networks

As Figure 9 shows: The two networks both have two peaks in the degree of 6 and 8, it shows that most of the nodes' degree value are 6 and 8, while not all nodes have the same degree value indicate that neither of the two networks is a rule network; Most nodes of each of the two networks have a similar degree conforms to the characteristics of random networks.

4. **Conclusions.** This article established the undirected unweighted networks for the roads inside each ring of Beijing city based on the Beijing traffic map data of Baidu maps and the primal approachand found these networks have the features of random networks and scall-free networks. Then, we made Beijing taxi traffic network models based on Beijing road networks and the taxi trajectory database of web "Urban Computing-Microsoft Research", using the modeling method TTIS after compareing the two modeling method TTI and TTIS we proposedand figured out that Beijing taxi traffic networks can be regarded as random networks. The randomness of Beijing road traffic networks and Beijing taxi traffic networks indicates that the randomness of the urban road traffic development ,whose city has a long history represented by Beijing.

Further research should adopt the dual approach [8] to modeling traffic networks, and take the artificial planning dominated emerging cities for research object.

## REFERENCES


[1] R. Albert, A. L. Barabsi, Statistical mechanics of complex networks, *Reviews of modern physics*, vol.74, no.5, pp.47, 2002.

[2] Z. M. Lu, Z. Wu, S. Z. Guo, a New Dynamic Community Model for Social Networks *International Journal of Modern Physics C*, vol.25, no.2, pp.215-244, 2014.

[3] Y.Zhao, W. Du , S.Chen, Application of Complex Network Theory to Urban Transportation Network Analysis, *Urban Transport of China*, vol.7, no.1, pp.57-65, 2009.

[4] C. L. Liu,D. Z Duan, Y. U Rui-Lin , J. Luo, Spatial correlation and evolution of complex urban-rural road network of wuhan metropolitan area, *Human Geography*, vol.129, no.1, pp.113-128, 2014.

[5] V. Latora ,M. Marchiori, Is the Boston subway a small-world network, *Physica A Statistical Mechanics and Its Applications*, vol.314, no.1, pp.109-113, 2002.

[6] W. Wang, J. Liu, X. Jiang, Y. Wang, Topology properties on chinese railway network, *Journal of Beijing Jiaotong University*, vol.34, no.3, pp.148-152, 2010.

[7] L. L. Qi, D. D.Zhao, Urban bus transport network optimization from complex network *Modelling, Identification and Control (ICMIC),Proceedings of International Conference onIEEE*, pp.1159-1164, 2012.

[8] X. F. Li, Z. M. Lu, H. Li, Controllability of deterministic complex networks *International Journal of Modern Physics C*, vol.26, no.3, pp.1159-1164, 2015.

[9] P. Y. Ye, Complex network characteristics of urban road network topology *Journal of Transportation Engineering and Information*, vol.10, no.1, pp.13-30, 2012.

[10] S. Porta ,P. Crucitti,V. Latora, The network analysis of urban streets: a dual approach, *Physica A: Statistical Mechanics and its Applications*, vol.369, no.2, pp.853-866, 2006.

[11] S. Porta ,P. Crucitti,V. Latora, The network analysis of urban streets: a primal approach, *Environment and Planning B Planning and Design*, vol.33, no.5, pp.705-725, 2006.

[12] Z. M. Lu,Y. X. Su,S. Z. Guo, Deterministic scale-free small-world networks of arbitrary order, *Physica A Statistical Mechanics* & *Its Applications* , vol.392, no.17, pp.3555-3562, 2013.

[13] Z. M. Lu, S. Z. Guo, A small-world network derived from the deterministic uniform recursive tree, *EPhysica A Statistical Mechanics* & *Its Applications*, vol.391, no.1-2, pp.87-92, 2012.

[14] M. Lv, X. L. Guo, J. Q. Chen, Second-order centrality correlation in scale-free networks, *International Journal of Modern Physics C*, vol.26, 2015.