Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Procedia - Social and Behavioral Sciences 96 (2013) 1611 – 1621
Procedia
Social and Behavioral Sciences

13th COTA International Conference of Transportation Professionals (CICTP 2013)

Complexity modeling and stability analysis of urban subway network: Wuhan City case study

Youjun Liu[a], Yue Tan[b] *

[a]School of Civil Enigeering and Mechanics, Huazhong Univrsity of Science and Technology,Wuhan 430074,China
[b]School of Civil Enigeering and Mechanics, Huazhong Univrsity of Science and Technology,Wuhan 430074,China

Abstract

Nowadays the rail transit is playing an increasingly important role in urban passenger trip, Meanwhile, its efficiency and reliability has been very concerned. With the tool of complex network theories, the network characteristic parameters of subway system are provided in this paper. And then the complex network models are established based on the method of space *L* and space *P*. In the case study of Wuhan city, the empirical result shows that the subway network is a scale-free network. Stability under the failure of the subway network is also analyzed. Then we define a new parameter to evaluate the important degree of one node in the subway network, which has great meaning in finding the most vital nodes (stations) and ensuring the stability of subway network. Overall, this study offers significant insights for planners in designing a more efficient and reliable system.

Selection and peer-review under responsibility of Chinese Overseas Transportation Association (COTA).

Key words:complexity network; subway system; scale-free; stability; node importance

1. Introduction

There is a growing literature on the complex networks applied to a great variety of problems following the revolutionary achievements by Watts and Strogatz (1998) and Barabasi and Albert (1999). A new class of network, *small worlds*, was defined by their studies. The small-world network is a type of mathematic graph where the typical distance L between two randomly chosen nodes (the number of steps required) grows proportionally to the logarithm of the number of nodes N in the network.

Supported by the Fundamental Research Funds for the Central Universities, HUST: 2011TS093
* Corresponding author. Tel.: +86-159-263-72293
E-mail address: lyjwuhan@163.com

Selection and peer-review under responsibility of Chinese Overseas Transportation Association (COTA).
doi:10.1016/j.sbspro.2013.08.183

Complex networks theories can be applied to an important problem in transportation systems. Latora and Marchiori (2002) studied the subway system of Boston, defined the 'efficiency' of the subway network. Sienkiewicz (2005) analyzed 22 bus networks in Poland, found out those networks can be described as small-world networks. GAO (2005) studied the scale-free properties of bus networks in China. Albert et al. (2000) studied how the properties of the Internet change when nodes are removed; they studied the error and attack tolerance of the network. Sen et al. (2003) studied the small-world properties of the Indian railway network. These studies indicate that some traffic networks have some similar characteristics with complex network.

The stability of network refers to the ability of a network maintaining its functional work without paralysis when suffering failures or attacks. With a paper titled Error and Attack Tolerance of Complex Networks by Albert et al (2000), the study of network stability has been concerned. Paolo Crucitti et al. (2004) studied the ability of a network to resist failures or attacks, which compared the results for two different network topologies: the random graph and the scale-free network. Better network can be built based on these researches by considering the occurrence of attack and failure on nodes to protect the existing network and locate the most important nodes.

Compared with the tool applied on the public transportation networks and urban roads networks, little works are reported over the complexity of subway networks. Do subway networks also have characteristics of complex network? If so, what is the general characteristic of real subway network? A static analysis of network stability under attack or failure is also presented where attack is identified by removing nodes with the highest degree and failure refers to removing nodes at random. Obviously, to explore the inner characteristics of the subway network and its stability will help us improve its efficiency and reliability.

2. Description of urban subway network

The subway network can be analyzed by defining the network topology. A real subway (Fig. 1(a)) network consists of station and tunnel (connecting couples of stations). Under normal circumstances, if there is a subway train traveling from station i to station j, there is also another one traveling from station j to station i along the same route. So the subway networks can be treated as undirected graphs. Meanwhile, the subway networks can also be treated as unweighted graphs by ignoring the difference of service frequency.

In general, the subway network can be described by two topological methods of Space L and Space P. The topological depiction based on Space L is defined in an intuitive and simple way. The nodes represent the stations and the links between two nodes only exist when they are consecutively stations. When it comes to the topological depiction based on Space P, the nodes represent the stations. However, much unlike what is in Space L, the links exist between any two stations along the same route, so that all the nodes of a specific route are linked directly. The ideas of Space L and Space P are explained in Fig. 1(b) and Fig. 1(c).

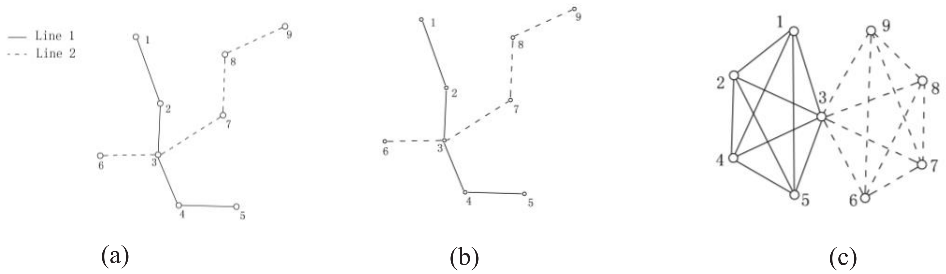

Fig. 1 A real network and its explanation of Space L and Space P

3. Properties of complex network

As its topological graph indicates, a subway network is a group of nodes connected between them by links. To use some basic ideas about complex networks, some properties that are useful when studying a subway system are introduced and explained.

3.1. Degree and Degree distribution

The degree of node *i* equals to the number of the nodes connect with it. The mathematical formulation of degree of node *i* is *k*:

$$K_i = \sum_{j=1}^{n} a_{ij} \tag{1}$$

When there is link between node *i* and *j*, a_{ij}=1. The degree distribution in the network is characterized by function *p* (*k*), which means the probability of a node has a degree of k, That is to say, *p* (*k*) is equivalent to the ratio between the number of nodes with a degree of *k* and the number of all nodes. It means the higher the degree of a note is, the more nodes connect with it. According to studies, the degree distribution of random network follows exponential distribution; and the degree distribution of scale-free network follows the power law.

3.2. Clustering coefficient

The clustering coefficient of a node can be defined as following: if the degree of node *i* is *k*, among its *k* neighbors, there can be *k**(*k*-1)/2 links at most, but the number of the links that actually exist is *E*, then the formula for clustering coefficient is C_i

$$C_i = \frac{E}{k(k-1)/2} \tag{2}$$

Average of clustering coefficients is denoted by *C*, it is used to describe the clustering degree of the network. According to the definition, the value of clustering coefficient ranges from 0 to 1. The more the value of *C* approximates to 1, the higher the clustering degree of the network will be.

3.3. Shortest path length

In an unweighted network, the shortest path length between any 2 nodes is defined with the number of links between them. It means the least number of nodes on the shortest path. A more important parameter is the average shortest path length. It is the average value of all the shortest path lengths in the network. It shows the accessibility of the network. The smaller average shortest path length a network has the better accessibility of it.

According to the properties discussed above, to calculate the parameters of Space *L* is different than that of Space *P*. In the network topology based on the idea of Space *L*, the subway network is defined in a natural way. The node degree represents the number of routes passing through the station. The shortest path length between any two nodes is the number of stations between them in the same route (plus one). This network exhibits the basic topological features of subway network.

In the network topology based on the idea of space P, the node degree means the number of stations which you can arrive from selected station without transfer. The distance between two stations can be explained as the necessary number of transfers (plus one) from one station to another. This network is used to describe the transfer of the subway network.

4. The Case of Wuhan City

Wuhan is the most populous city in Central China, having a population of more than 10 million people

(Census 2011), with about 6 million residents in its urban area. For supporting the urban sprawling and increasing of trip demand, a huge subway system has been planned and is under construction. Fig.2 shows the system map of Wuhan metro routes planned toward 2020. The network consists of N =159 stations and K = 159 tunnels extending throughout Wuhan and the other cities of the Wuhan Metropolitan Circle.

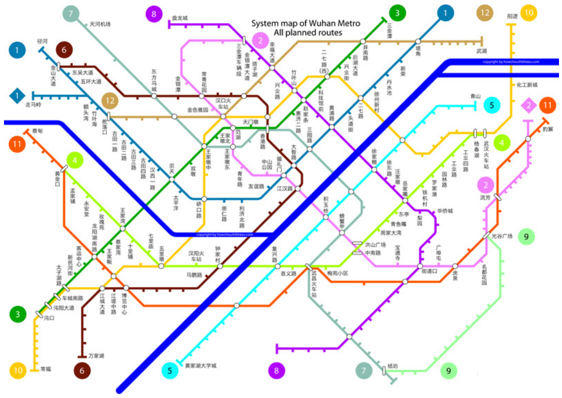

Fig. 2.System map of Wuhan metro routes planned toward 2020

We represent the subway network in the form of matrix with the tool of Matlab. Since there are 159 stations, the subway network is represented as a 159*159 matrix $A[a]$. When there is a link between node i and node j, a_{ij}=1; otherwise, a_{ij} =0.

The condition under which a_{ij} =1 in the idea of Space L differs from that of Space P. In Space L, a_{ij} =1 only if node i and node j are consecutive stations on the route. And in Space P, a_{ij} =1 on the condition that node i and node j are on the same route and there is a train traveling between them.

4.1. Space L

4.1.1. Degree and Degree distribution

Fig. 3 shows the degree distribution of the subway network in the idea of Space L where the network is very small, the value of the degree ranges from 1 to 4, and the average is 2.321. The nodes of k=1 are the ends on subway routes. 20% of the nodes have the degree of 4; these nodes represent transfer stations in the network.

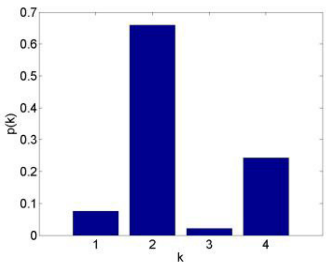

Fig. 3. Degree distribution of Wuhan subway network

Most of nodes have a degree of 2, which means these nodes are only linked to 2 adjacent nodes and there is only one route passing through the nodes of degree k=2.

4.1.2. Clustering coefficient

In the idea of Space L, the degree of most nodes are 2, the nodes are only linked with 2 adjacent nodes on the same route, which obviously have no links between each other. Thus, the clustering coefficients of most nodes are 0. The clustering degree of the subway network in the idea of Space L is too low to matter.

4.1.3. Average shortest path length

Fig. 4 shows the distribution of the shortest path length. It fits well to an asymmetric, unimodal function. The maximum of the shortest path lengths is 28. Most of shortest path lengths are 10. It means for most node pairs, there are 10 nodes between them on the shortest route. Average of the shortest path length is 10.49. It is relatively small; it means it is relatively convenient to travel in the network.

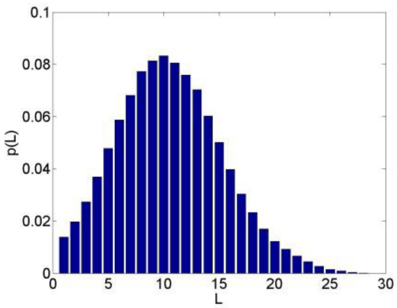

Fig. 4. Distribution of the shortest path length

Compared to the number of the network, the distance between 2 nodes are much smaller. This behavior is typical for small-world network. More obvious evidences are found in the network on Space P, which will be presented later.

Fig. 5 shows that 50% of the shortest path lengths are within 10 and 90% are within 15. The distances are very smaller compared to the number of the nodes in the network.

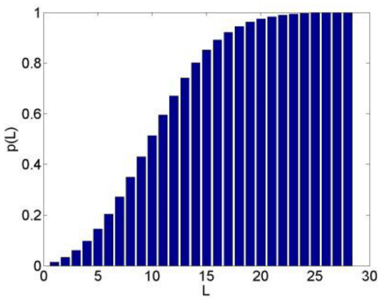

Fig. 5. Cumulative degree distribution of the shortest path lengths

Table 1 presents the parameters of the network in the idea of Space *L*. The data shows the subway network of Wuhan has a smaller average shortest path length in comparison with number of nodes in the subway network. It presents small-world networks properties.

Table 1. Parameters of subway network in Space *L*

Parameter	Value
Number of nodes	159
Number of lines	174
Average network degree	2.32
Average of the shortest path lengths	10.49
Diameter of the network	28
Clustering coefficient	≈0

4.2. Space P

4.2.1. Degree and Degree distribution

In the idea of Space *P*, the degrees of nodes in the network are much bigger than that in the idea of Space *L*. Since all nodes on the same routes are linked in this model.

In order to smooth large fluctuations, we use cumulative distribution to describe node degree distribution in Space *P*. Fig 6 shows the cumulative distribution of node degree. In Space *P*, the degrees of the nodes are much higher than that in Space *L*: the Maximum, minimum and average of the degrees are 52.16 and 25.65, respectively.

Using the log-log scale, the cumulative distribution can be presented as a power law function. It shows the network has the properties of scale-free network. This is because the nodes in the subway network are not treated equally. The nodes with bigger degree are more likely to be linked to other nodes. This preferential attachment may lead to node degree distribution follow power law function.

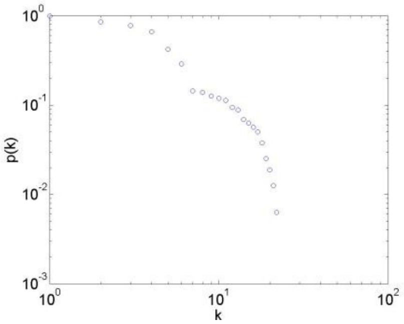

Fig. 6.Cumulative distribution of node degree

4.2.2. Clustering coefficient

In the idea of Space *P*, the average of clustering coefficients is 0.93, which is very close to 1. That 85.53% of the stations have the clustering coefficients of 1 seems to be very higher.

Here is an explanation for this phenomenon. In the idea of Space *P*, there is a link between any 2 nodes on the same route. When a station belongs to only one route, it is directly linked to all of its neighbors. Thus, most nodes have the clustering coefficients of 1, which also means for most of the stations, there is only one route passing through them.

4.2.3. Average shortest path length

In Space *P*, the network topology is used to describe the transfer of the network, the shortest path length between two stations means numbers of transfers (plus one) from one station to another.

Fig.7 shows the distribution of transfer times. Nearly 80% of the numbers of transfers are 1. Average shortest path length of the subway network of Wuhan city is 2.00, which means only one transfer is enough to travel between most of the nodes in the network. So it is convenient to travel by subway. The subway network of Wuhan city is quite effective.

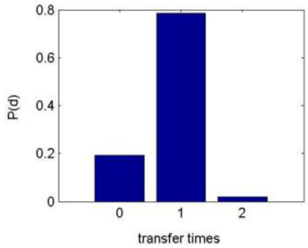

Fig. 7.distribution of transfer times

The empirical investigation result of the subway network of Wuhan city are presented both in the idea of Space *L* and Space *P*. With the empirical data, it can be proved that the subway network has some characteristics of complex network. The network in the idea of Space *L* exhibits small world effect as the distance between 2 nodes are quite smaller compared to the number of nodes in the network. In Space *P*, the node degree distribution follows power law function. It indicates scale-free network properties. As the average shortest paths are really short and the clustering coefficient is relatively big, it also presents small-world network properties. Table 2 shows the parameters of the network in the idea of Space *P*.

Table 2. Parameters of subway network in Space P

Parameter	Value
Number of nodes	159
Number of lines	2039
Average network degree	25.65
Average of the shortest path lengths	2.00
Diameter of the network	3.00
Clustering coefficient	0.93

5. Stability analysis of subway network

5.1. Network behavior

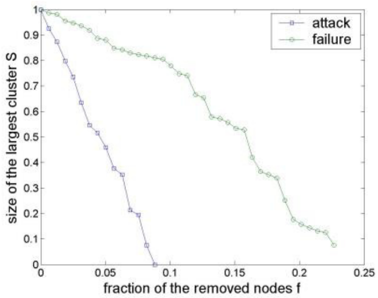

Fig. 8.Changes of S in the network vs. the fraction f of the removed nodes

As depicted in Fig. 8, the size of the fragments that breaks off from the network increases rapidly when under attack. If 4% of the nodes with highest degree are removed, the number of nodes left would be only half of the original in the network. When about 8% of the most connected nodes are removed, the network would be collapsed completely. The subway network shows great fragility when the highest degree nodes are attacked. While the nodes are removed at random, the network breaks off slowly and no threshold shows up. Even 22% of the nodes get removed randomly, ratio of the largest cluster size remains at a considerable high level, being about 10%.

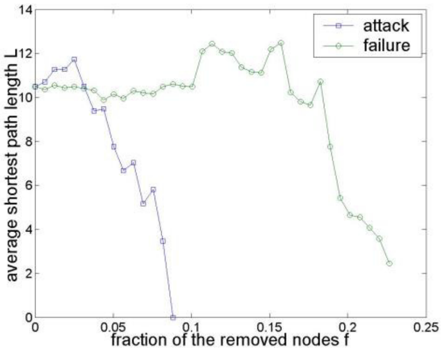

Fig. 9.changes of L in the network vs. the fraction F of the removed nodes

As Fig. 9 portrays, at the very beginning, the average shortest path length increases, which reflects the connectivity of the network being worsen. But with more nodes removed, it appears abnormal that the length value falls alarmingly. The results of attack and failure situations are similar.

The simulation results show that at the very beginning of the attack or failure a few nodes get removed and some routes become unavailable, which leads to the worsening connectivity as well as the increasement of average shortest path length. With more nodes removed, a lot of fragments break off from the network and the network itself becomes smaller and smaller. The average shortest path length decreases, which doesn't mean a better connectivity of the network.

5.2. The difference between the result of attack and failure

When being attacked, the network could be cracked into a paralyzed one with only 8.8% of the nodes being removed. But in the situation of failure, this percentage increases to 22%. The above difference indicates that the subway network shows greater robustness and is more stable under random failure. The difference between the two situations lies in the heterogeneity of the scale-free network: the degrees of most nodes are very low and uniformly distributed, while only a few nodes have high degrees. However, it is those high degree nodes that are more important for the connectivity of the network. When under attack, the nodes with high degree are removed immediately and many big fragments are thus separated from the network, as a consequence, S decreases quickly. But when the nodes are removed randomly, nodes with low degree will be selected with a higher probability because of its huge amount. Removal of those nodes doesn't damage the connectivity of the network too much, so the network breaks down much more slowly.

Generally, the subway network shows strong robustness when under random failure, and appears great fragility when under attack. This phenomenon is caused by the fact that the subway network is a scale-free network with only a few nodes responsible for the connectivity of the network.

6. New parameter for node importance evaluation

As previously illustrated, the subway network is a scale-free network and only a few nodes are more important for the connectivity of the network. So finding the most vital nodes and taking measures to protect them in the subway network would be an effective way to improve its stability.

Current researches in finding the most vital nodes in network mainly focus on two kinds of methods: one uses a single parameter to evaluate the node importance, which does not include all relevant factors and gives no precise result; and another evaluates the node importance using the change in parameters before and after the node being removed and which leads to only fuzzy result. In this section, we choose a new parameter to evaluate the node importance, overcoming the shortcomings of the two former methods.

6.1. Parameter of node centrality

For a node in the subway network, its importance is based on the following factors: a) its position in the network, the closer to the center of the network, the more important it is; b) its connectivity degree, higher, more important; and c) its neighbor nodes' degree, when its neighbors have higher degrees, it is more important.

When a node is closer to the center of a network, it will be easier to travel to other nodes because of the shortened distances.

For a better understanding, the closeness to the network centre of node i, $Ce(i)$ is defined as follows to measure the closeness of node i to the network centre,

$$Ce(i) = 1 / \sum_{j=1}^{n} L(i,j) \tag{3}$$

Where $L(i, j)$ represents the shortest path lengths from node i to node j, (j=1, 2, ⋯, n).

Also, node degree $K(i)$ is presented to evaluate the connectivity of the node i and $M(i)$ is defined as the sum of degrees of the nodes which are directly linked to node i.

$$M(i) = \sum_{j=1}^{k_i} K(j) \tag{4}$$

Then parameter $I(i)$ is defined to evaluate the importance of node i,

$$I(i) = Ce(i) * M(i) * K(i) \tag{5}$$

Put Eq. (2) into Eq. (5), parameter $I(i)$ can be calculated as follow.

$$I(i) = M(i) * K(i) / \sum_{j=1}^{n} L(i,j) \quad (6)$$

6.2. Case study

The effectiveness of this new parameter is shown with numerical results. The subway network of Wuhan city is taken for a case study. 10 nodes are randomly chosen to evaluate their importance with different methods. The comparison of the evaluating results with different methods makes it explicit that this new parameter is more effective.

Table3 shows the results of node importance evaluated with different parameters. For better observing, we magnify the term *I* by 100 times and the change of *L* expresses the change of average path lengths in the network before and after a node being removed.

Table 3.node importance evaluated by different parameters

Node number	Degree *k*	Change in *L*	*I(i)*
24	2	-3.90%	0.35
92	2	-0.41%	0.32
155	2	-1.57%	0.46
124	2	1.46%	0.42
31	2	-0.91%	0.42
8	4	3.67%	4.32
42	4	2.12%	3.22
153	2	1.52%	0.51
11	4	2.02%	4.11
152	2	0.48%	0.47

If we use node degree to evaluate node importance, these 10 nodes have only two values of degree, 4 and 2. Then it is impossible to distinguish the relative importance of the nodes with the same degree.

In order to have a clearer look at the importance of these ten nodes, we evaluate node importance by observing the change of *L* when certain node is removed from the network. Table3 above shows change of *L* with two different kinds of results, positive and negative. Here comes the problem again those two nodes which lead different directions of change can not be compared when they have the same value, such as node 24 and node 8.

As shown in Table3, the node importance can be easily evaluated in descending sort, that is, No.8> No.11> No.42> No.153> No.152> No.155> No.31> No.124> No.24> No.92.Therefore, the new parameter is much more convenient and effective than the other two methods previously in finding the most vital nodes in the network

7. Conclusion

In this paper, the topology of the subway network of in the idea of Space *P* and Space *L* is defined and the empirical results are presented, with taking Wuhan city as a case study. The simulation results indicate that the network shows small-world behavior and degree distribution in Space *P* obeys power law. The subway network can be seen as a scale-free network. The stability of the subway network is analyzed in this paper. It is found that the network is robust to random failure and fragile to attack as a scale-free network. At last a new parameter is defined to evaluate the node importance of the network. A numerical result is used to show the effectiveness of the new parameter.

Acknowledgements

The work was supported by the Fundamental Research Funds for the Central Universities (HUST: 2011TS093).

References

Watts D J, Strogatz. S H. (1998).Collective dynamics of 'small-world' networks. *Nature*, 393: 440-442
Barabsi A, Albert R. (1999). Emergence of Scaling in Random Networks. *Science*, 286: 509-513
Sienkiewicz J, Holyst J A. (2005). Statistical analysis of 22 public transport networks in Poland. *Physical Review E*, 72: 046127-1:11
Yong-Zhou Chen, Nan Li, Da-Ren He. (2007). A study on some urban bus transport networks. *Physica A* , 376: 747–754
Latora V, Marchiori M. (2002). Is the Boston subway a small-world network. *Physica A* , 314: 109 – 113
Lu Huapu , Shi Ye. (2007). Complexity of Public Transport Networks. *Tsinghua Science and Technology*, 2007, 12(2): 204-213
Angeloudis P, Fisk D. (2006). Large subway systems as complex networks. *Physica A*, 367: 553-558
Xinping Xu, Junhui Hu, Feng Liu, Lianshou Liu. (2007). Scaling and correlations in three bus-transport networks of China. *Physica A* , 374: 441-448
Albert R, Hawoong Jeong ,Barabasi A. (2000). Error and attack tolerance of complex networks. *Nature*, 406: 378-382
Latora V, Marchiori M.Lianshou Liu. (2001). Efficient Behavior of Small-World Networks. *Physical Review Letters*, 87(19): 198701-1:4
Albert R, Barabasi A. (2002). Statistical mechanics of complex networks. *Review of Modern Physics*, 2002, 74: 47-97
Li Ping, Xiong Xing, Qiao Zhong-Liang. (2006).Topological Properties of Urban Public Traffic Networks in Chinese Top-Ten Biggest Cities. *Chinese Physics Letters*, 23(12): 3384-3387
Sen P,Dasgupta S. (2003). Small-world properties of the Indian railway network. *Physical Review E*, 67: 36106-1:5
Barabsi A, Albert R. (2002). Statistical mechanics of complex networks. *REVIEWS OF MODERN PHYSICS*, 74(1):47-97
Latora V, Marchiori M.Lianshou Liu. (2001). Efficient Behavior of Small-World Networks. *Physical Review Letters*, 87(19): 198701-1:4
Liu Zhiqian, Song Rui, (2010). Reliability Analysis of Guangzhou Rail Transit with Complex Network Theory. Journal of Transportation *Systems Engineering and Information Technology*, 10(5):194-200
Chen Jing, Sun Linfu. (2009). Evaluation of Node Importance in complex Networks. *Journal of Southwest Jiaotong University*, 44(3):426-429
WANG Tao,FANG Zhigeng,WU Hui,WU Linli. (2008). An Analysis of Complexity of Subway Network in China. *Jounrnal of Academy of Military Transportation*,10(2):24-28
Tan Yuejin, Wu Jun, Deng Hongzhong. (2006). Evaluation method for node importance based on Node Contraction in Complex Networks. *System Engineering-Theory & Practice*, 11:79-84
HE Nan, LI De-Yi, GAN WenYan, ZHU Xi. (2007). Mining Vital Nodes in Complex Networks. *Computer Science*, 34(12):1-5
Paolo Crucitti, Vito Latora et al.(2004). Error and attack tolerance of complex networks. *PHYSICA A*, (340):388-394