Journal of Traffic and Transportation Engineering 12 (2024) 11-22
doi: 10.17265/2328-2142/2024.01.002

Nonlinear Relationship and Its Evolutionary Trace between Node Degree and Average Path Length of China Aviation Network Based on Complex Network

Cheng Xiangjun[1], Zhang Xiaoxuan[2] and Li Yangqi[1]
1. Beijing Jiaotong University, Beijing 100044, China
2. China Highway & Transportation Society, Beijing 100011, China

Abstract: In order to reveal the complex network characteristics and evolution principle of China aviation network, the relationship between the node degree and the average path length of China aviation network in 1988, 1994, 2001, 2008 and 2015 was studied. According to the theory and method of complex network, the network system was constructed with the city where the airport was located as the network node and the airline as the edge of the network. On the basis of the statistical data, the node average path length of China aviation network in 1988, 1994, 2001, 2008 and 2015 was calculated. Through regression analysis, it was found that the node degree had a logarithmic relationship with the average length of node path, and the two parameters of the logarithmic relationship had linear evolutionary trace.

Key word: China aviation network, complex network, node degree, average length of node path, logarithmic relationship, evolutionary trace.

1. Introduction

Aviation network is typical complex network with small world characters [1, 2]. About certain nation's aviation network, there are some unknown features in the field of complex network. This paper faces to the China aviation network through analyzing the passenger data [3] of civil aviation airlines in year 1988, 1994, 2001, 2008 and 2015 to reveal the complex network feature. According to complex network theory, network system of airports and airlines of China were constructed with airports regarded as nodes and airline regarded as edges to study the relationship between node degree and the average path length of China aviation network. Based on the statistical data, the average length of node path in China aviation network in 1988, 1994, 2001, 2008 and 2015 was calculated. Through regression analysis, it was found that the node degree had a logarithmic relationship with the average length of node path, and the two parameters of the logarithmic relationship had linear evolutionary trace.

2. Nonlinear Relationship between Node Degree and Average Path Length of China Aviation Network

For network $G=(V,E)$, where $v_i \in V$ is the node of G. V is the set of nodes. E is the set of edges [4], $(v_i, v_j) \in E$. Matrix $A=(a_{i,j})_{n\times n}$ was constructed, where:

$$a_{i,j}=\begin{cases}1,(v_i,v_j)\in E\\ 0, otherwise\end{cases} \tag{1}$$

Matrix A is called adjacent matrix of network G. The length of path $d_{i,k}$ between node v_i and v_k is defined as the quantity of edges of shortest path connecting these two nodes. The rank of v_i is defined as the serial number of descending orders by node degree [4].

Corresponding author: Cheng Xiangjun, Ph.D., lecturer, research fields: complex network, urban traffic, transportation simulation, traffic and transportation engineering.

The path length matrix $D=(d_{i,j})_{n\times n}$ was gotten from adjacent matrix A by algorithm Floyd [4].

$$d_{i,j}=\begin{cases}\Omega_{i,j}, \text{the quantity of edges of shortest path between node } v_i \text{ and } v_j\\ 0, i=j\\ \infty,\ v_i \ and\ v_j \ was\ not\ connected\end{cases} \tag{2}$$

Since the aviation network of China was fully connected network in each year [3], so the path length matrix D could be simplified as the following:

$$d_{i,j}=\begin{cases}\Omega_{i,j}, \text{the quantity of edges of shortest path between node } v_i \text{ and } v_j\\ 0, i=j\end{cases} \tag{3}$$

According to the definition, the average path length $\bar{d}_i$ of node v_i to other nodes is:

$$\bar{d}_i=\frac{1}{n-1}\sum_{j=1}^{n}\Omega_{i,j} \tag{4}$$

2.1 Nonlinear Relationship between Node Degree and Average Path Length of China Aviation Network in 1988

According to the data of China aviation network in 1988 [3], $N_{1988}=85$, the adjacency matrix A_{1988} was calculated. Using algorithm Floyd, the path length matrix D_{1988} was calculated by A_{1988}. The average path length of 85 nodes was calculated by D_{1988} and Eq. (4). The scatter points of node degree and average path length of these 85 nodes were drawn in Fig. 1.

Let the node degree be x and the average length of node path be y. Let $u=\ln x$. The points in Fig. 2 were calculated by the points in Fig. 1. The correlation coefficient r of the points in Fig. 2 was calculated by Eq. (5).

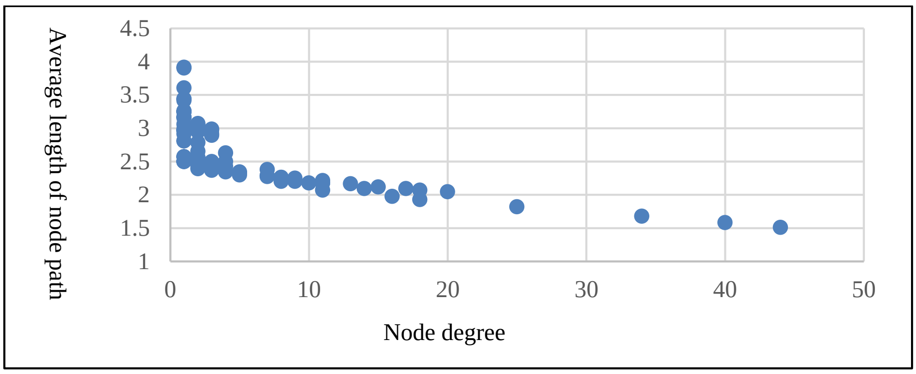

Fig. 1 Diagram of relationship between node degree and average path length of China aviation network in 1988.

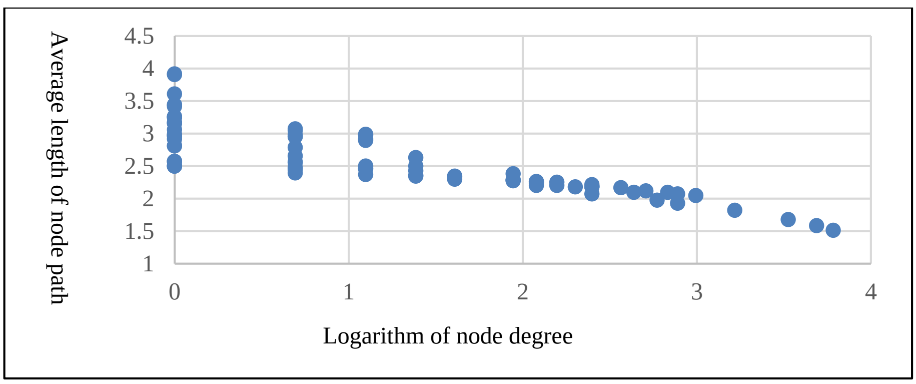

Fig. 2 Diagram of relationship between logarithm of node degree and average path length in 1988.

$$r = \frac{L_{uy}}{\sqrt{L_{uu}L_{yy}}} = \frac{\sum_{i=1}^{n}(u_i - \bar{u})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(u_i - \bar{u})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (5)$$

Here, $n = 85$. The value of correlation coefficient r was calculated, $r = -0.838$. The critical value $r_{\big|{\alpha=1\% \atop f=83}} = 0.289$ was found in critical value table [5] at degree of freedom $f = n - 2 = 83$ and level of significant α of 1%. Since $|r| = 0.838 > 0.289 = r_{\big|{\alpha=1\% \atop f=83}}$, the scattered points in Fig. 2 had significant linear correlation. Least square method [5] was used as an approach in Eq. (6) to fit the line with points in Fig. 2.

$$\begin{cases} \hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{u} = 3.035 \\ \hat{\beta}_1 = \dfrac{L_{uy}}{L_{uu}} = -0.375 \end{cases} \quad (6)$$

The linear equation:

$$\hat{y} = 3.035 - 0.375u \quad (7)$$

The fitting line Eq. (7) was drawn with the sample points in Fig. 3 with good fitting effect.

To take t test [5] of Eq. (7), test hypothesis is: $H_0 : \beta_1 = 0$. When the hypothesis was true, there is:

$$\hat{\beta}_1 \sim N(0, \frac{\sigma^2}{L_{uu}}) \quad (8)$$

Here, $\hat{\beta}_1$ fluctuates near zero, statistic t is build.

$$t = \frac{\hat{\beta}_1}{\sqrt{\dfrac{\hat{\sigma}^2}{L_{uu}}}} = \frac{\hat{\beta}_1\sqrt{L_{uu}}}{\hat{\sigma}} \quad (9)$$

where:

$$\hat{\sigma}^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (10)$$

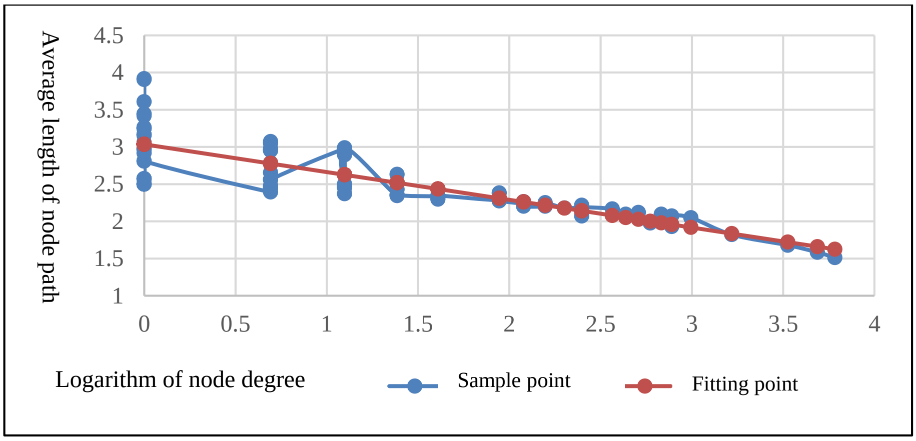

Fig. 3 Fitting linear relationship between logarithm of node degree and average path length in 1988.

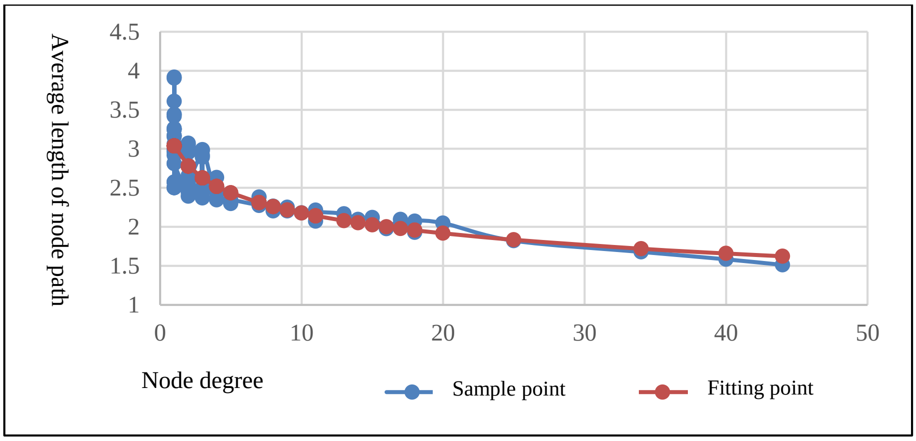

Fig. 4 Fitting nonlinear relationship between node degree and average path length in 1988.

Calculated by statistic data:

$$t = -13.95$$

To check the t distribution table [5], at significant level α of 0.01 and degree of freedom $f = n-2 = 83$, the value $t_{\substack{\alpha=0.01 \\ f=83}} = 2.379$. So, $|t| = 13.95 > 2.379 = t_{\substack{\alpha=0.01 \\ f=83}}$, null hypothesis H_0 is refused. The linear correlation of Eq. (7) is significant. The fitting curve Eq. (11) of nonlinear relationship between node degree and the nearest neighbor average degree was deduced from Eq. (7):

$$\hat{y} = 3.035 - 0.375 \ln x \qquad (11)$$

The points of the fitting curve Eq. (11) and the sample points were drawn in Fig. 4 with good fitting effect. It showed that the node degree in China aviation network in 1988 had a logarithmic relationship with the average length of node path.

2.2 Nonlinear Relationship between Node Degree and Average Path Length of China Aviation Network in 1994

According to the data of China aviation network in 1994, $N_{1994} = 122$, the adjacency matrix A_{1994} was calculated. Using algorithm Floyd, the path length matrix D_{1994} was calculated by A_{1994}. The path average length of 122 nodes was calculated by D_{1994} and Eq. (4). The scatter points of node degree and average path length of these 122 nodes were drawn in Fig. 5.

Let the node degree be x and the average length of node path be y. Let $u = \ln x$. The points in Fig. 6 were calculated by the points in Fig. 5. The correlation coefficient r of the points in Fig. 6 was calculated by Eq. (5). Here, $n = 122$. The value of correlation coefficient r was calculated, $r = -0.882$. The critical value $r_{\substack{\alpha=1\% \\ f=120}} = 0.233$ was found in critical value table at degree of freedom $f = n-2 = 120$ and level of

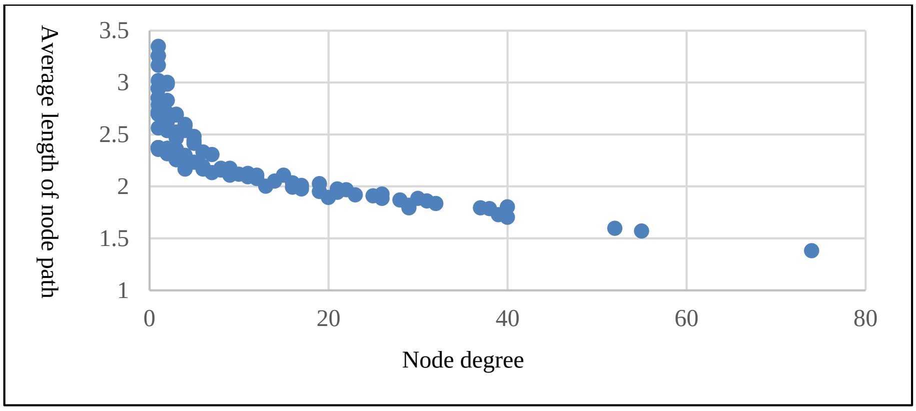

Fig. 5 Diagram of relationship between node degree and average path length in 1994.

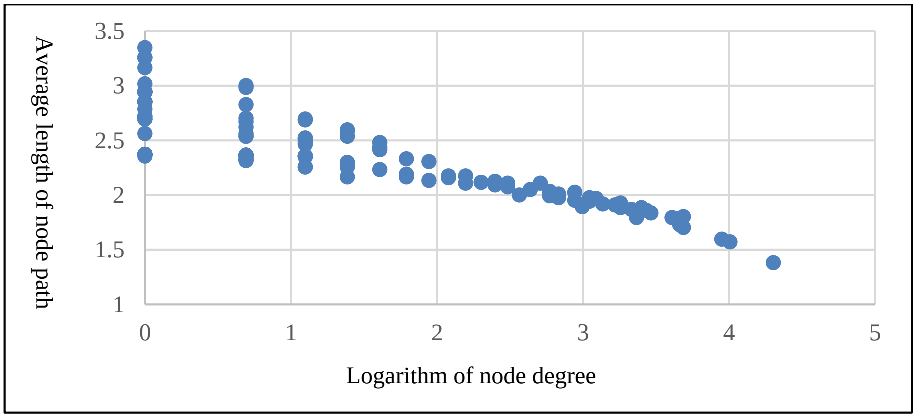

Fig. 6 Diagram of relationship between logarithm of node degree and average path length in 1994.

significant α of 1%. Since $|r| = 0.882 > 0.233 = r_{\alpha=1\%, f=120}$, the scattered points in Fig. 6 had significant linear correlation. Least square method was used as an approach in Eq. (12) to fit the line with points in Fig. 6.

$$\begin{cases} \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{u} = 2.757 \\ \hat{\beta}_1 = \dfrac{L_{uy}}{L_{uu}} = -0.275 \end{cases} \tag{12}$$

The linear equation:

$$\hat{y} = 2.757 - 0.275u \tag{13}$$

The fitting line Eq. (13) was drawn with the sample points in Fig. 7 with good fitting effect.

To take t test of Eq. (12), test hypothesis is: $H_0 : \beta_1 = 0$. When the hypothesis was true, there is Eq. (8). The statistic t was calculated by Eqs. (9) and (10), $t = -20.5$. To check the t distribution table, at significant level α of 0.01 and degree of freedom $f = n - 2 = 120$, the value $t_{\alpha=0.01, f=120} = 2.358$. So, $|t| = 20.5 > 2.358 = t_{\alpha=0.01, f=120}$, null hypothesis H_0 is refused. The linear correlation of Eq. (13) is significant. The fitting curve Eq. (14) of nonlinear relationship between node degree and the path average length was deduced from Eq. (13):

$$\hat{y} = 2.757 - 0.275 \ln x \tag{14}$$

The points of the fitting curve Eq. (14) and the sample points were drawn in Fig. 8 with good fitting effect. It showed that the node degree in China aviation network in 1994 had a logarithmic relationship with the average length of node path.

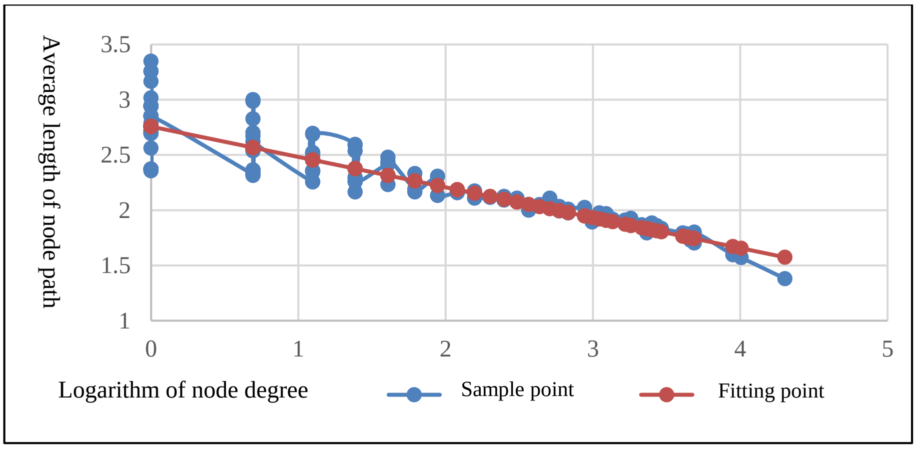

Fig. 7 Fitting linear relationship between logarithm of node degree and average path length in 1994.

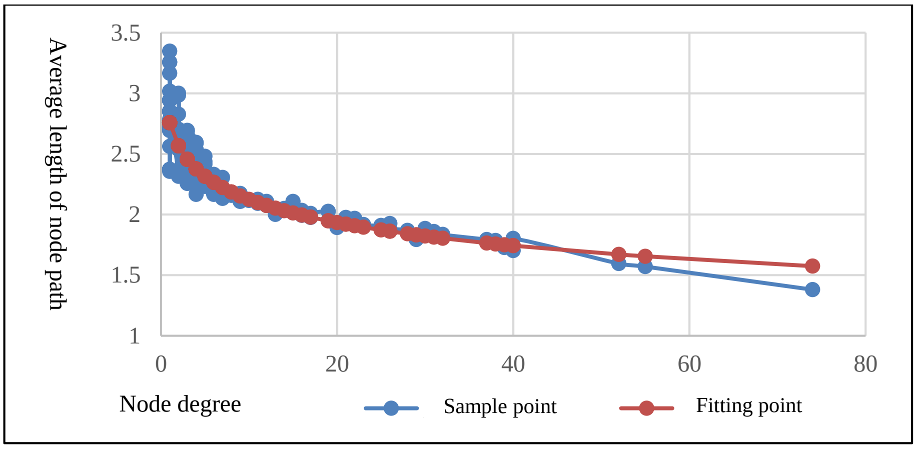

Fig. 8 Fitting nonlinear relationship between node degree and average path length in 1994.

2.3 Nonlinear Relationship between Node Degree and Average Path Length of China Aviation Network in 2001

According to the data of China aviation network in 2001, $N_{2001}=130$, the adjacency matrix A_{2001} was calculated. Using algorithm Floyd, the path length matrix D_{2001} was calculated by A_{2001}. The path average length of 130 nodes was calculated by D_{2001} and Eq. (4). The scatter points of node degree and average path length of these 130 nodes were drawn in Fig. 9.

Let the node degree be x and the average length of node path be y. Let $u=\ln x$. The points in Fig. 10 were calculated by the points in Fig. 9. The correlation coefficient r of the points in Fig. 10 was calculated by Eq. (5). Here, $n=130$. The value of correlation coefficient r was calculated, $r=-0.884$. The critical value $r_{\left|{\alpha=1\% \atop f=128}\right.}=0.226$ was found in critical value table at degree of freedom $f=n-2=128$ and level of significant α of 1%. Since $|r|=0.884>0.226=r_{\left|{\alpha=1\% \atop f=128}\right.}$, the scattered points in Fig. 10 had significant linear correlation. Least square method was used as an approach in Eq. (15) to fit the line with points in Fig. 10.

$$\begin{cases} \hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{u}=2.718 \\ \hat{\beta}_1=\dfrac{L_{uy}}{L_{uu}}=-0.267 \end{cases} \tag{15}$$

The linear equation:

$$\hat{y}=2.718-0.267u \tag{16}$$

The fitting line Eq. (16) was drawn with the sample points in Fig. 11 with good fitting effect.

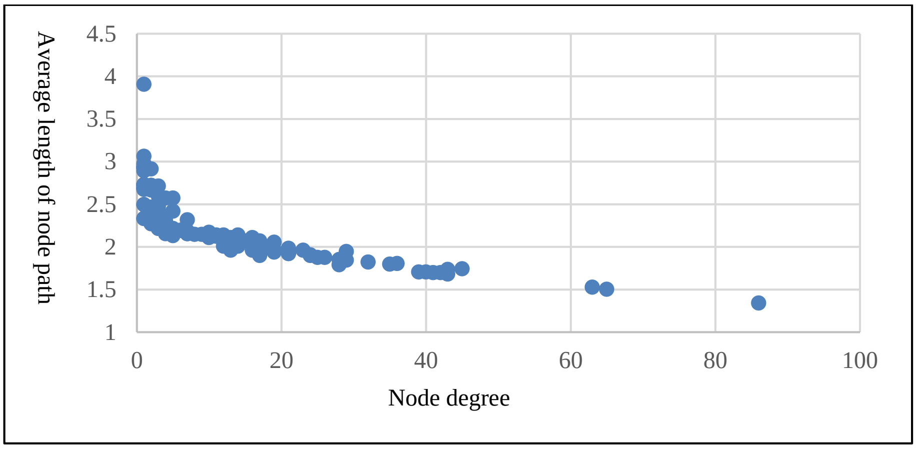

Fig. 9 Diagram of relationship between node degree and average path length in 2001.

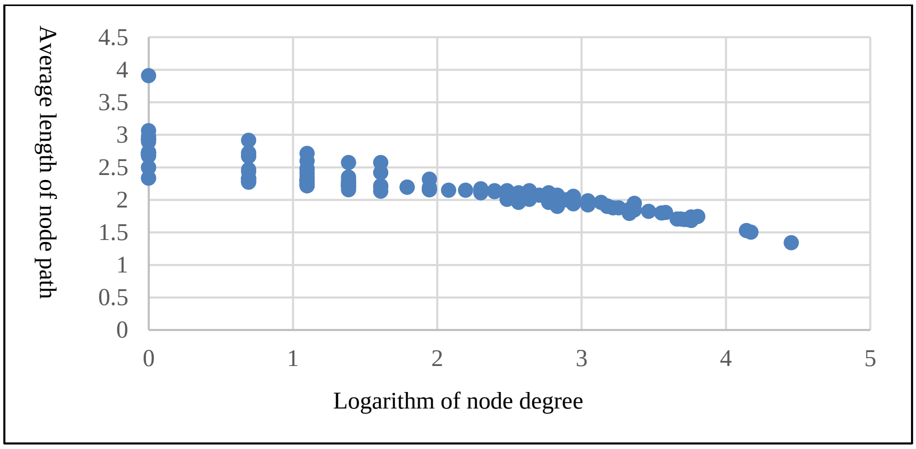

Fig. 10 Diagram of relationship between logarithm of node degree and average path length in 2001.

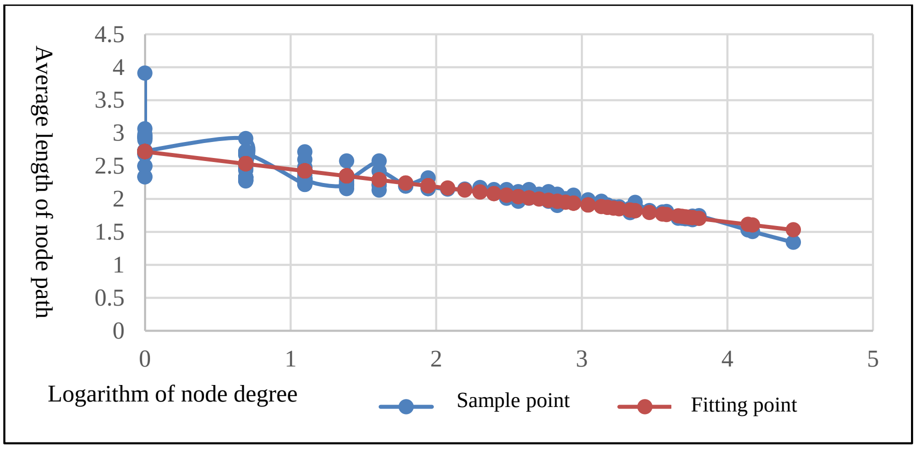

Fig. 11 Fitting linear relationship between logarithm of node degree and average path length in 2001.

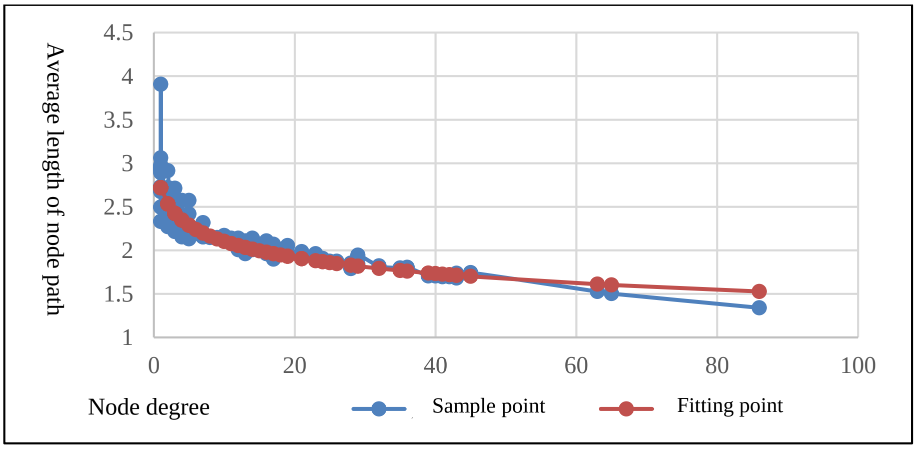

Fig. 12 Fitting nonlinear relationship between node degree and average path length in 2001.

To take t test of Eq. (16), test hypothesis is: $H_0:\beta_1=0$. When the hypothesis was true, there is Eq. (8). The statistic t was calculated by Eqs. (9) and (10), $t=-21.36$. To check the t distribution table, at significant level α of 0.01 and degree of freedom $f=n-2=128$, the value $t_{\substack{\alpha=0.01\\ f=128}}=2.357$. So, $|t|=21.36>2.357=t_{\substack{\alpha=0.01\\ f=128}}$, null hypothesis H_0 is refused. The linear correlation of Eq. (16) is significant. The fitting curve Eq. (17) of nonlinear relationship between node degree and the path average length was deduced from Eq. (16):

$$\hat{y}=2.718-0.267\ln x \tag{17}$$

The points of the fitting curve Eq. (17) and the sample points were drawn in Fig. 12 with good fitting effect. It showed that the node degree in China aviation network in 2001 had a logarithmic relationship with the average length of node path.

2.4 Nonlinear Relationship between Node Degree and Average Path Length of China Aviation Network in 2008

According to the data of China aviation network in 2008, $N_{2008}=150$, the adjacency matrix A_{2008} was calculated. Using algorithm Floyd, the path length matrix D_{2008} was calculated by A_{2008}. The path average length of 150 nodes were calculated by D_{2008} and Eq. (4). The scatter points of node degree and average path length of these 150 nodes were drawn in Fig. 13.

Let the node degree be x and the average length of node path be y. Let $u=\ln x$. The points in Fig. 14 were calculated by the points in Fig. 13. The correlation

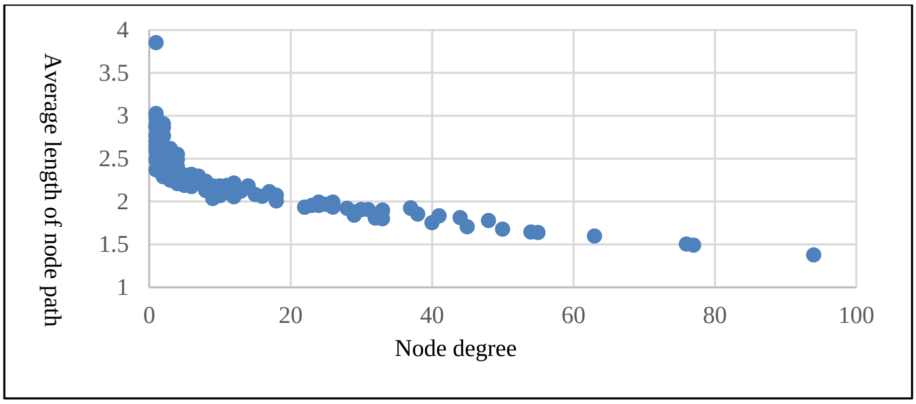

Fig. 13 Diagram of relationship between node degree and average path length in 2008.

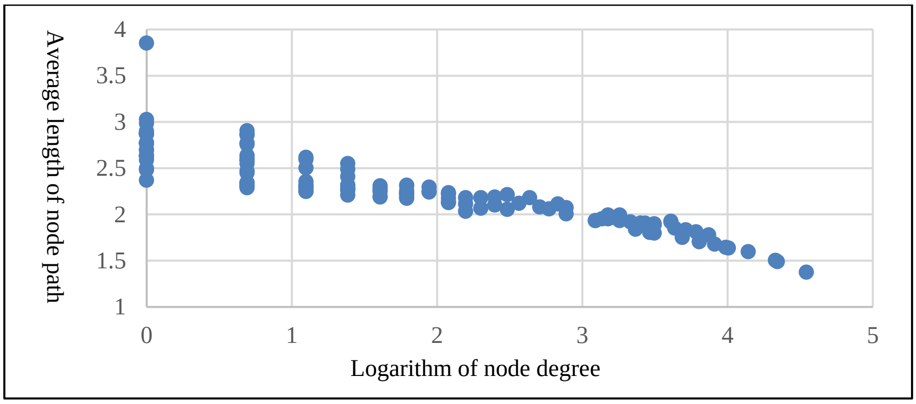

Fig. 14 Diagram of relationship between logarithm of node degree and average path length in 2008.

coefficient r of the points in Fig. 14 was calculated by Eq. (5). Here, $n=150$. The value of correlation coefficient r was calculated, $r=-0.896$. The critical value $r_{\substack{\alpha=1\% \\ f=148}}=0.21$ was found in critical value table at degree of freedom $f=n-2=148$ and level of significant α of 1%. Since $|r|=0.896>0.21=r_{\substack{\alpha=1\% \\ f=148}}$, the scattered points in Fig. 14 had significant linear correlation. Least square method was used as an approach in Eq. (18) to fit the line with points in Fig. 14.

$$\begin{cases} \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{u} = 2.746 \\ \hat{\beta}_1 = \dfrac{L_{uy}}{L_{uu}} = -0.264 \end{cases} \tag{18}$$

The linear equation:

$$\hat{y} = 2.746 - 0.264u \tag{19}$$

The fitting line Eq. (19) was drawn with the sample points in Fig. 15 with good fitting effect.

To take t test of Eq. (19), test hypothesis is: $H_0:\beta_1=0$. When the hypothesis was true, there is Eq. (8). The statistic t was calculated by Eqs. (9) and (10), $t=-26.4$. To check the t distribution table, at significant level α of 0.01 and degree of freedom $f=n-2=148$, the value $t_{\substack{\alpha=0.01 \\ f=148}}=2.354$. So, $|t|=26.4>2.354=t_{\substack{\alpha=0.01 \\ f=148}}$, null hypothesis H_0 is refused. The linear correlation of Eq. (19) is significant. The fitting curve Eq. (20) of nonlinear relationship between node degree and the path average length was deduced from Eq. (19):

$$\hat{y} = 2.746 - 0.264 \ln x \tag{20}$$

The points of the fitting curve Eq. (20) and the sample points were drawn in Fig. 16 with good fitting effect. It showed that the node degree in China aviation network in 2008 had a logarithmic relationship with the average length of node path.

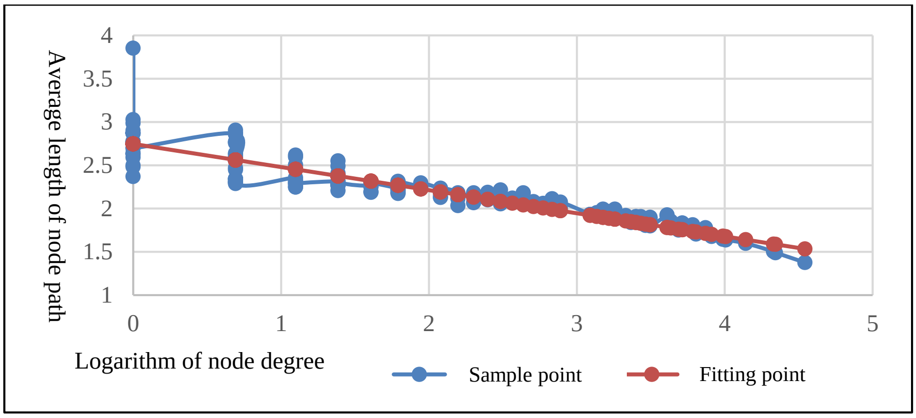

Fig. 15 Fitting linear relationship between logarithm of node degree and average path length in 2008.

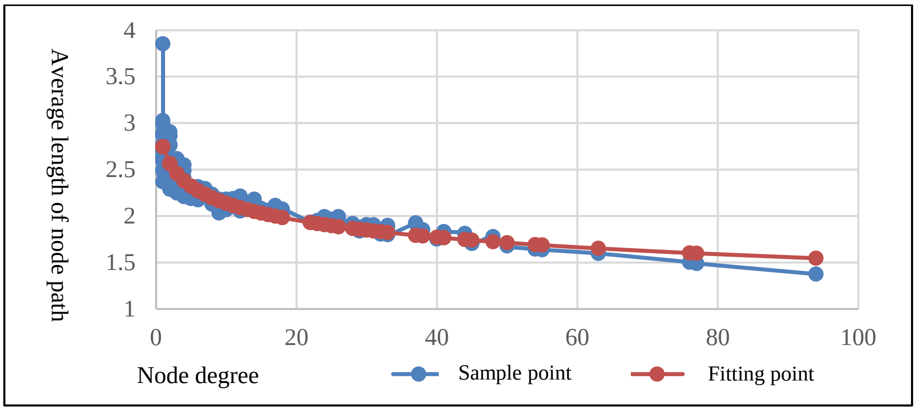

Fig. 16 Fitting nonlinear relationship between node degree and average path length in 2008.

2.5 Nonlinear Relationship between Node Degree and Average Path Length of China Aviation Network in 2015

According to the data of China aviation network in 2015, $N_{2015}=203$, the adjacency matrix A_{2015} was calculated. Using algorithm Floyd, the path length matrix D_{2015} was calculated by A_{2015}. The path average length of 203 nodes was calculated by D_{2015} and Eq. (4). The scatter points of node degree and average path length of these 203 nodes were drawn in Fig. 17.

Let the node degree be x and the average length of node path be y. Let $u=\ln x$. The points in Fig. 18 were calculated by the points in Fig. 17. The correlation coefficient r of the points in Fig. 18 was calculated by Eq. (5). Here, $n=203$. The value of correlation coefficient r was calculated, r = -0.904. The critical value $r_{\big|\substack{\alpha=1\% \\ f=201}}=0.181$ was found in critical value table at degree of freedom $f=n-2=201$ and level of significant α of 1%. Since $|r|=0.904>0.181=r_{\big|\substack{\alpha=1\% \\ f=201}}$, the scattered points in Fig. 18 had significant linear correlation. Least square method was used as an approach in Eq. (21) to fit the line with points in Fig. 18.

$$\begin{cases} \hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{u}=2.658 \\ \hat{\beta}_1=\dfrac{L_{uy}}{L_{uu}}=-0.234 \end{cases} \tag{21}$$

The linear equation:

$$\hat{y}=2.658-0.234u \tag{22}$$

The fitting line Eq. (22) was drawn with the sample points in Fig. 19 with good fitting effect.

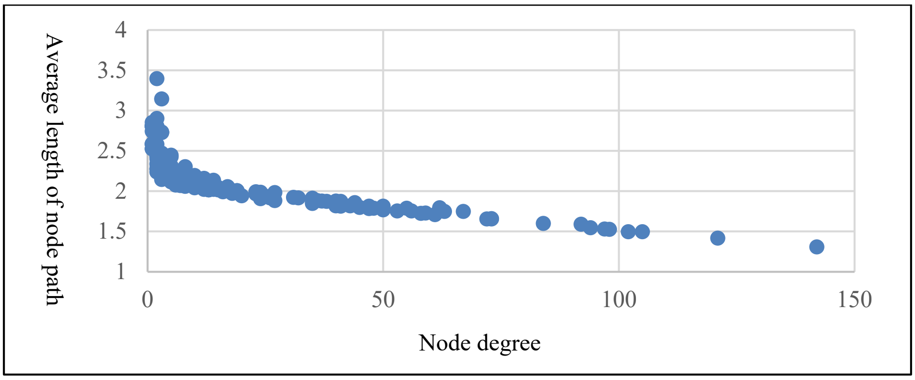

Fig. 17 Diagram of relationship between node degree and average path length in 2015.

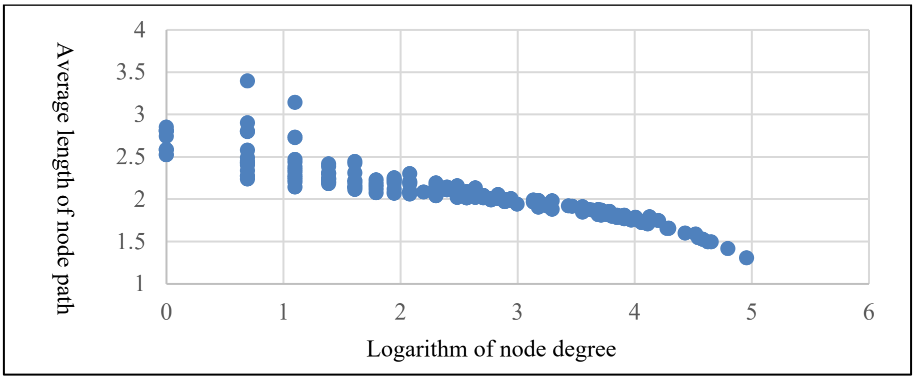

Fig. 18 Diagram of relationship between logarithm of node degree and average path length in 2015.

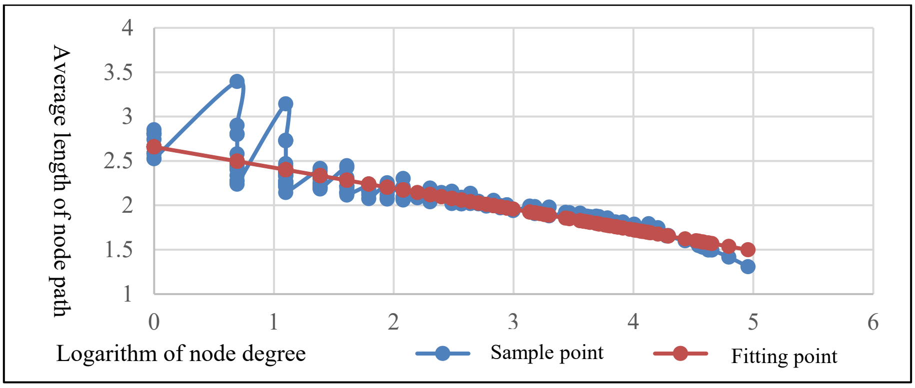

Fig. 19 Fitting linear relationship between logarithm of node degree and average path length in 2015.

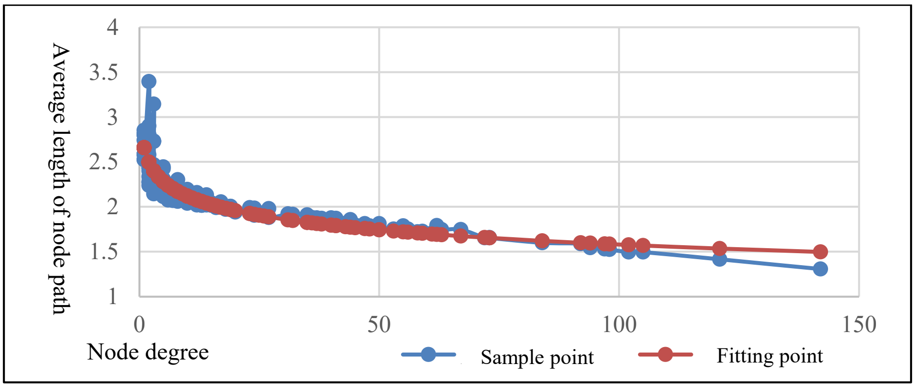

Fig. 20 Fitting nonlinear relationship between node degree and average path length in 2015.

To take t test of Eq. (22), test hypothesis is: $H_0 : \beta_1 = 0$. When the hypothesis was true, there is Eq. (8). The statistic t was calculated by Eqs. (9) and (10), $t = -29.89$. To check the t distribution table, at significant level α of 0.01 and degree of freedom $f = n - 2 = 201$, the value $t_{\substack{\alpha=0.01\\ f=201}} = 2.351$. So, $|t| = 29.89 > 2.351 = t_{\substack{\alpha=0.01\\ f=201}}$, null hypothesis H_0 is refused. The linear correlation of Eq. (22) is significant. The fitting curve Eq. (23) of nonlinear relationship between node degree and the average path length was deduced from Eq. (22):

$$\hat{y} = 2.658 - 0.234 \ln x \tag{23}$$

The points of the fitting curve Eq. (23) and the sample points were drawn in Fig. 20 with good fitting effect. It showed that the node degree in China aviation network in 2015 had a logarithmic relationship with the average length of node path.

3. Evolution of Relationship between Node Degree and Average Path Length of China Aviation Network

3.1 Logarithmic Relationship Evolution between Node Degree and Average Path Length of China Aviation Network

The fitting curves Eq. (11), Eq. (14), Eq. (17), Eq. (20) and Eq. (23) of relationship between node degree and the average path length of China aviation network in 1988, 1994, 2001, 2008 and 2015 were drawn in Fig. 21.

In Fig. 21, from 1988 to 2015, the relationship curves between the node degree and the average path length were arranged from bottom to top, and the front part intersected and the rear part diverged. With the increasing of the number of nodes, the average length of node path decreased rapidly and then slowly. The change of curve spacing was positively correlated with the change of the number of nodes in the aviation network.

3.2 Parameter Evolution of Logarithmic Relationship between Node Degree and Average Path Length of China Aviation Network

The two parameters in fitting curve equation were taken as one point. Then the parameters in five fitting curve equations could consist five points: (3.035, -0.375), (2.757, -0.275), (2.718, -0.267), (2.764, -0.264), (2.658, -0.234). These five points were drawn in Fig. 22. The correlation coefficient r of the points in Fig. 22 was calculated, $r = 0.995$. Here, $n = 5$. The critical value $r_{\substack{\alpha=1\%\\ f=3}} = 0.959$ was found in critical value table at degree of freedom $f = n - 2 = 3$ and level of significant α of 1%. Since $|r| = 0.995 > 0.959 = r_{\substack{\alpha=1\%\\ f=3}}$, the scattered points in Fig. 22 had significant linear correlation. Least square method was used as an approach in Eq. (24) to fit the line with points in Fig. 22.

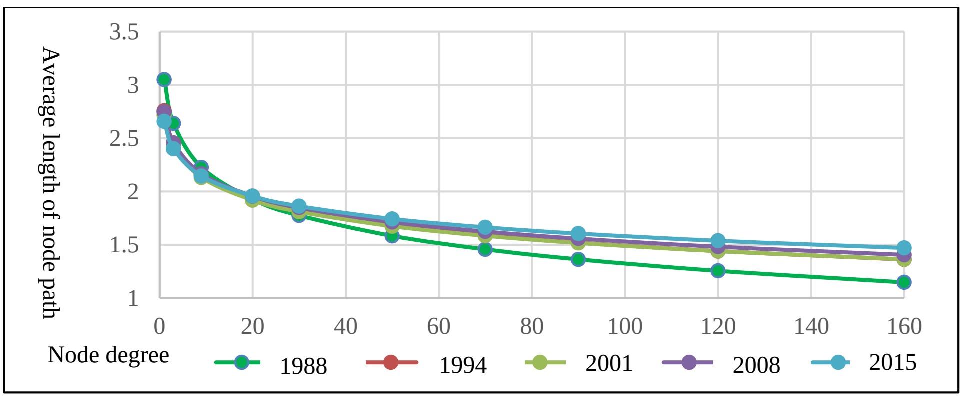

Fig. 21 Evolution diagram of logarithmic relationship between node degree and average path length of China aviation network.

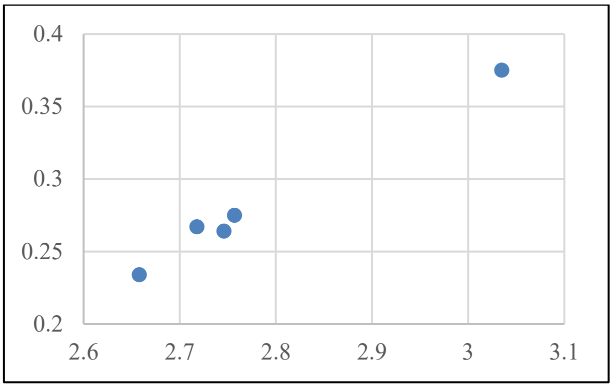

Fig. 22 Diagram of relationship between the two parameters of logarithmic curves.

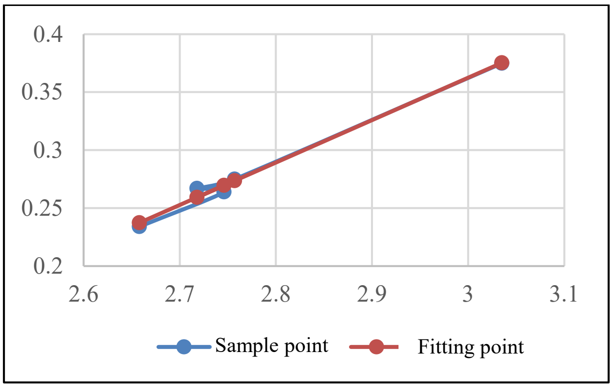

Fig. 23 Fitting linear relationship between the two parameters of logarithmic curves.

$$\begin{cases} \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = -0.7355 \\ \hat{\beta}_1 = \dfrac{L_{xy}}{L_{xx}} = 0.366 \end{cases} \tag{24}$$

The linear equation:

$$\hat{y} = 0.366x - 0.7355 \tag{25}$$

The fitting line Eq. (25) was drawn with the sample points in Fig. 23 with good fitting effect.

To take t test of Eq. (25), test hypothesis is: $H_0 : \beta_1 = 0$. When the hypothesis was true, there is Eq. (8). The statistic t was calculated by Eqs. (9) and (10), $t = 18.2$. To check the t distribution table at significant level α of 0.01 and degree of freedom $f = n - 2 = 3$, the value $t\big|_{\alpha=0.01, f=3} = 4.541$. So, $|t| = 18.2 > 4.541 = t\big|_{\alpha=0.01, f=3}$, null hypothesis H_0 is refused. The linear correlation of Eq. (25) is significant.

4. Conclusion

In order to reveal the complex network characteristics and evolution principle of China aviation network, the relationship between the node degree and the average path length of China aviation network in 1988, 1994, 2001, 2008 and 2015 was studied. According to the statistical data, it was found that the node degree had a logarithmic relationship with the average length of node path. From 1988 to 2015, the relationship curves between the node degree and the average path length were arranged from bottom to top, and the front part intersected and the rear part diverged. With the increasing of the number of nodes, the average length of node path decreased rapidly and then slowly. The two parameters of the logarithmic relationship of these five curves had linear evolutionary trace through regression analysis.

References

[1] Hossian, M. M., and Alam, S. 2017. "A Complex Network Approach towards Modeling and Analysis Australia Airport Network." *Journal of Airport Transport Management* 60: 1-9.

[2] Rocha, L. E. C. 2017. "Dynamics of Air Transport Networks: A Review from a Complex Systems Perspective." *Chinese Journal of Aeronautics* 30 (2): 469-78.

[3] Division of Development and Planning of Civil Aviation Administration of China. 1990, 1996, 2003, 2010, 2017. *Civil Aviation Statistical Yearbook of China.* Beijing: Civil Aviation Press of China. (in Chinese)

[4] Sun, X. J., and Si, S. K. 2016. *Complex Network Algorithms and Applications*. Beijing: National Defence Industry Press. (in Chinese)

[5] He, X. Q., and Liu, W. Q. 2006. *Application of Regression Analysis*. Beijing: China Renmin University Press. (in Chinese)