

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

Cheng Xiangjun¹, Zhang Chunyue² and Zhang Xiaoxuan³

1. Beijing Jiaotong University, Beijing 100044, China

2. Central University of Finance and Economics, Beijing 100081, China

3. 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 nearest neighbor average degree and its evolution trace of China aviation network in 1988, 1994, 2001, 2008 and 2015 were 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. According to the statistical data, the node nearest neighbor average degree 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 negative exponential relationship with the nearest neighbor average degree, and the two parameters of the negative exponential relationship had linear evolution trace.

Key words: China aviation network, complex network, node degree, nearest neighbor average degree, negative exponential relationship, evolution 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 was constructed with airports regarded as nodes and airline regarded as edges to study the relationship between node degree and the nearest neighbor average degree and the relationship evolution trace of China aviation network. Based on the statistical data, the nearest

neighbor average degrees of nodes in China aviation network in 1988, 1994, 2001, 2008 and 2015 were calculated. Through regression analysis, it was found that the node degree had a negative exponential relationship with the nearest neighbor average degree, and the two parameters of the negative exponential relationship had linear evolution trace.

2. Nonlinear Relationship between Node Degree and Nearest Neighbor Average Degree 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:

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

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

Matrix A was called adjacent matrix of network G . The degree of node v_i was defined as the number of edges connected to v_i . The nearest neighbor average degree $k_{nn,i}$ of node v_i was defined as following [4]:

$$k_{nn,i} = \frac{\left[\sum_j (a_{i,j} \cdot k_j) \right]}{k_i} \quad (2)$$

In Eq. (2), k_i —the degree of node v_i ; k_j —the degree of other node v_j in the network; $a_{i,j}$ —the element of adjacent matrix.

Then the nearest neighbor average degree $k_{nn}(k)$ of all nodes with degree k was defined as following [4]:

$$k_{nn}(k) = \frac{\sum_{i|k_i=k} k_{nn,i}}{[N \cdot P(k)]} \quad (3)$$

In Eq. (3), N —amount number of node; $P(k)$ —distribution function of node degree.

2.1 Nonlinear Relationship between Node Degree and Nearest Neighbor Average Degree of China Aviation Network in 1988

According to the data of China aviation network in 1988, $N_{1988} = 85$. Using the adjacency matrix A_{1988} and Eq. (2), there were $n = 21$ different degrees of 85 nodes in 1988 China aviation network. The nearest

neighbor average degree corresponding to each degree could be obtained by using Eq. (3). Thus, there were 21 pairs of data, forming functional relationships, as shown in Fig. 1.

Let the node degree be x axis and the nearest neighbor average degree be y axis in Fig. 1. Let $v = \ln y$ in Fig. 2. 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. (4).

$$r = \frac{L_{xv}}{\sqrt{L_{xx}L_{vv}}} = \frac{\sum_{i=1}^n (x_i - \bar{x})(v_i - \bar{v})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (v_i - \bar{v})^2}} \quad (4)$$

Here, $n = 21$. The value of correlation coefficient r was calculated, $r = -0.868$. The critical value of

$r_{\alpha=1\%, f=19}$ was 0.549 found in critical value table [5] at degree of freedom $f = n - 2 = 19$ and level of significant α of 1%. Since $|r| = 0.868 > 0.549 = r_{\alpha=1\%, f=19}$, the scattered points in

Fig. 2 had significant linear correlation. Least square method [5] was used as an approach in Eq. (5) to fit the line with points in Fig. 2.

$$\begin{cases} \hat{\beta}_0 = \bar{v} - \hat{\beta}_1 \bar{x} = 3.17 \\ \hat{\beta}_1 = \frac{L_{xv}}{L_{xx}} = -0.0213 \end{cases} \quad (5)$$

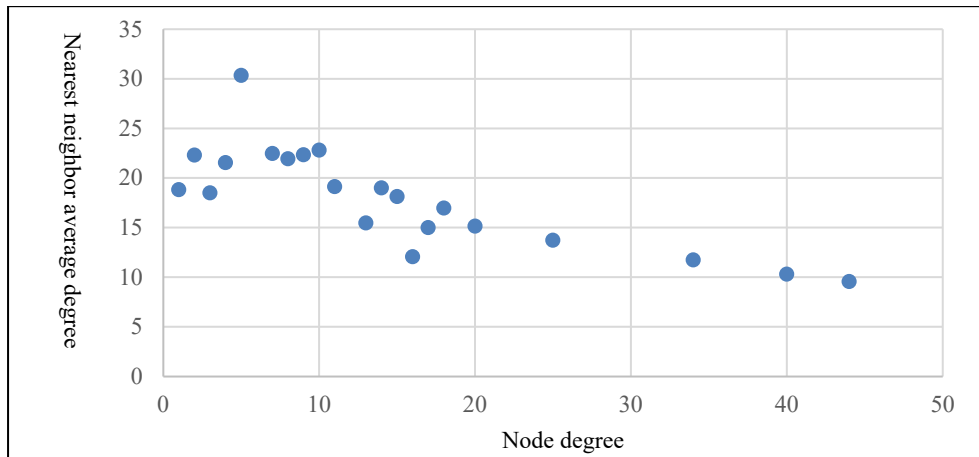


Fig. 1 Diagram of relationship between node degree and nearest neighbor average degree of China aviation network in 1988.

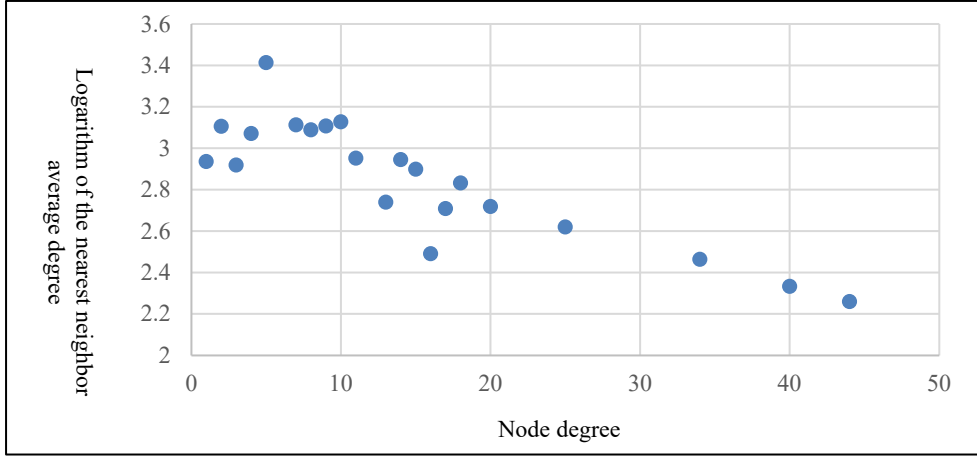


Fig. 2 Diagram of relationship between node degree and logarithm of nearest neighbor average degree in 1988.

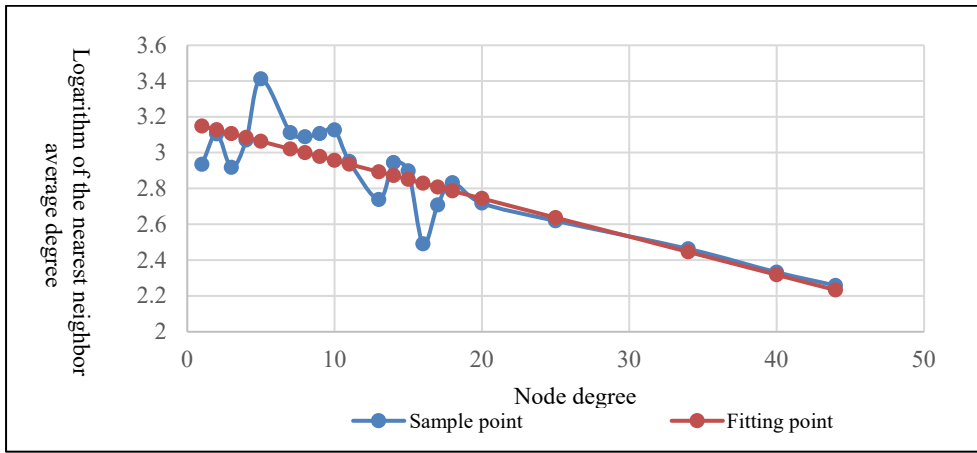


Fig. 3 Fitting linear relationship between node degree and logarithm of nearest neighbor average degree in 1988.

The linear equation:

$$\hat{v} = -0.0213x + 3.17 \quad (6)$$

The fitting line Eq. (6) was drawn with the sample points in Fig. 3 with good fitting effect.

To take t test [5] of Eq. (6), test hypothesis is: $H_0: \beta_1 = 0$. When the hypothesis was true, there is:

$$\hat{\beta}_1 \sim N(0, \frac{\sigma^2}{L_{xx}}) \quad (7)$$

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

$$t = \frac{\hat{\beta}_1}{\sqrt{\frac{\hat{\sigma}^2}{L_{xx}}}} = \frac{\hat{\beta}_1 \sqrt{L_{xx}}}{\hat{\sigma}} \quad (8)$$

where:

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^n (v_i - \hat{v}_i)^2 \quad (9)$$

and is calculated by statistic data:

$$t = -7.65$$

To check the t distribution table [5], at significant level α of 0.01 and degree of freedom $f = n - 2 = 19$, the value of $t_{\alpha=0.01, f=19}$ in table is 2.539. So,

$|t| = 7.65 > 2.539 = t_{\alpha=0.01, f=19}$, null hypothesis H_0 is refused.

The linear correlation of Eq. (6) is significant. The fitting curve Eq. (10) of nonlinear relationship between node degree and the nearest neighbor average degree was deduced from Eq. (6):

$$\hat{y} = 23.807e^{-0.0213x} \quad (10)$$

The points of the fitting curve Eq. (10) 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 negative exponential relationship

with the nearest neighbor average degree.

2.2 Nonlinear Relationship between Node Degree and Nearest Neighbor Average Degree of China Aviation Network in 1994

According to the data of China aviation network in 1994, $N_{1994} = 122$. Using the adjacency matrix A_{1994} and Eq. (2), there were $n = 36$ different degrees of 122 nodes in 1994 China aviation network. The nearest neighbor average degree corresponding to each degree could be obtained by using Eq. (3). Thus, there were 36 pairs of data, forming functional relationships, as shown in Fig. 5.

Let the node degree be x axis and the nearest neighbor average degree be y axis in Fig. 5. Let $v = \ln y$ in Fig. 6. 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. (4), $r = -0.878$. Here,

$n = 36$. The critical value of $r_{\alpha=1\%, f=34}$ was 0.424 found in

critical value table at degree of freedom $f = n - 2 = 34$ and level of significant α of 1%.

Since $|r| = 0.878 > 0.424 = r_{\alpha=1\%, f=34}$, the scattered points

in Fig. 6 had significant linear correlation. Least square method was used as an approach in Eq. (11) to fit the line with points in Fig. 6.

$$\begin{cases} \hat{\beta}_0 = \bar{v} - \hat{\beta}_1 \bar{x} = 3.669 \\ \hat{\beta}_1 = \frac{L_{xy}}{L_{xx}} = -0.0144 \end{cases} \quad (11)$$

The linear equation:

$$\hat{v} = -0.0144x + 3.669 \quad (12)$$

The fitting line Eq. (12) was drawn with the sample points in Fig. 7 with good fitting effect.

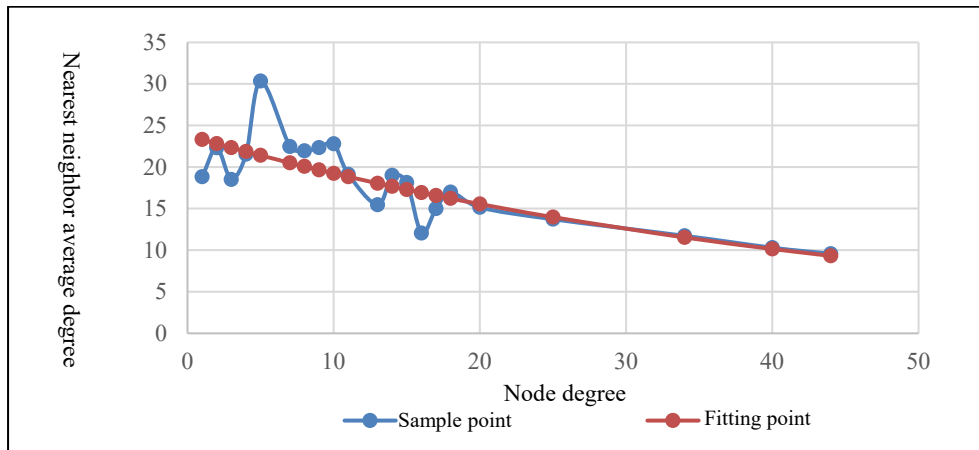


Fig. 4 Fitting nonlinear relationship between node degree and nearest neighbor average degree in 1988.

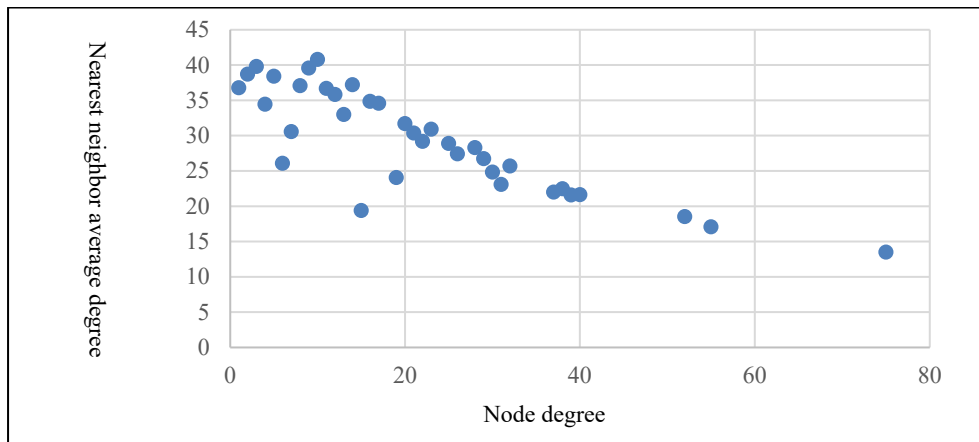


Fig. 5 Diagram of relationship between node degree and nearest neighbor average degree in 1994.

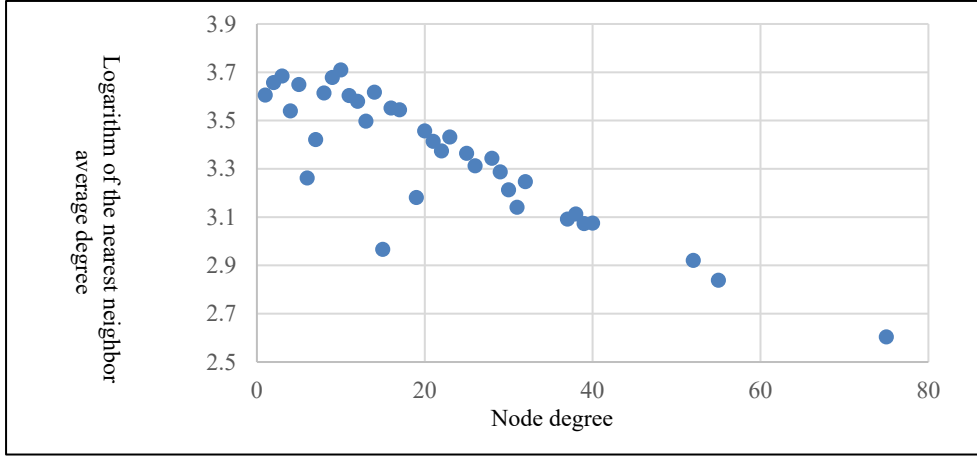


Fig. 6 Diagram of relationship between node degree and logarithm of nearest neighbor average degree in 1994.

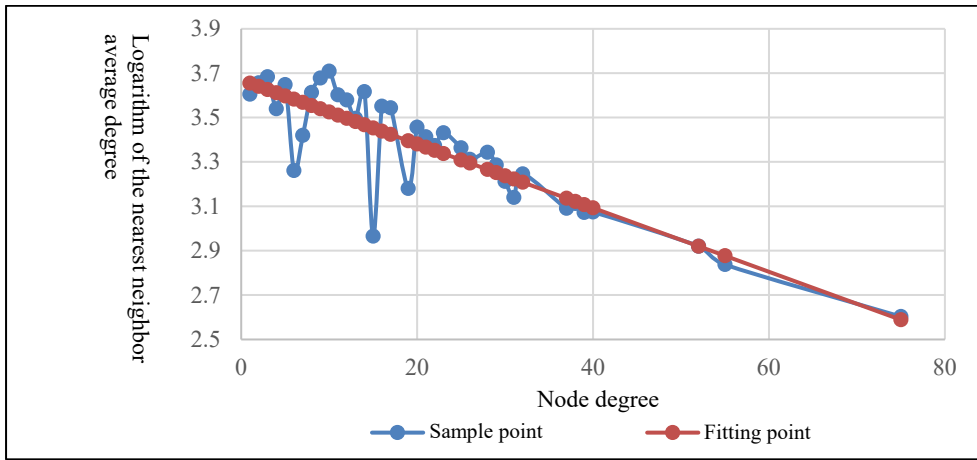


Fig. 7 Fitting linear relationship between node degree and logarithm of nearest neighbor average degree in 1994.

To take t test of Eq. (12), test hypothesis is: $H_0: \beta_1=0$. When the hypothesis was true, there is Eq. (7). The statistic t was calculated by Eqs. (8) and (9), $t = -10.74$. To check the t distribution table, at significant level α of 0.01 and degree of freedom $f = n - 2 = 34$, the value of $t_{\alpha=0.01, f=34}$ in table is 2.443. So, $|t| = 10.74 > 2.443 = t_{\alpha=0.01, f=34}$, null hypothesis H_0

is refused. The linear correlation of Eq. (12) is significant. The fitting curve Eq. (13) of nonlinear relationship between node degree and the nearest neighbor average degree was deduced from Eq. (12):

$$\hat{y} = 39.212e^{-0.0144x} \quad (13)$$

The points of the fitting curve Eq. (13) 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 negative exponential

relationship with the nearest neighbor average degree.

2.3 Nonlinear Relationship between Node Degree and Nearest Neighbor Average Degree of China Aviation Network in 2001

According to the data of China aviation network in 2001, $N_{2001}=130$. Using the adjacency matrix A_{2001} and Eq. (2), there were $n=38$ different degrees of 130 nodes in 2001 China aviation network. The nearest neighbor average degree corresponding to each degree could be obtained by using Eq. (3). Thus, there were 38 pairs of data, forming functional relationships, as shown in Fig. 9.

Let the node degree be x axis and the nearest neighbor average degree be y axis in Fig. 9. Let $v = \ln y$ in Fig. 10. 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. (4), $r = -0.952$. Here, $n = 38$. The critical value of $r_{\alpha=1\%, f=36}$ was 0.413 found in critical value table at degree of freedom $f = n - 2 = 36$ and level of significant α of 1%. Since $|r| = 0.952 > 0.413 = r_{\alpha=1\%, f=36}$, the scattered points in Fig. 10 had significant linear correlation. Least square method was used as an approach in Eq. (14) to fit the line with points in Fig. 10.

$$\begin{cases} \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 3.819 \\ \hat{\beta}_1 = \frac{L_{xy}}{L_{xx}} = -0.0143 \end{cases} \quad (14)$$

The linear equation:

$$\hat{y} = -0.0143x + 3.819 \quad (15)$$

The fitting line Eq. (15) was drawn with the sample

points in Fig. 11 with good fitting effect.

To take t test of Eq. (15), test hypothesis is: $H_0: \beta_1 = 0$. When the hypothesis was true, there is Eq. (7). The statistic t was calculated by Eqs. (8) and (9), $t = -18.62$. To check the t distribution table, at significant level α of 0.01 and degree of freedom $f = n - 2 = 36$, the value of $t_{\alpha=0.01, f=36}$ in table is 2.437.

So, $|t| = 18.62 > 2.437 = t_{\alpha=0.01, f=36}$, null hypothesis

H_0 is refused. The linear correlation of Eq. (15) is significant. The fitting curve Eq. (16) of nonlinear relationship between node degree and the nearest neighbor average degree was deduced from Eq. (15):

$$\hat{y} = 45.559e^{-0.0143x} \quad (16)$$

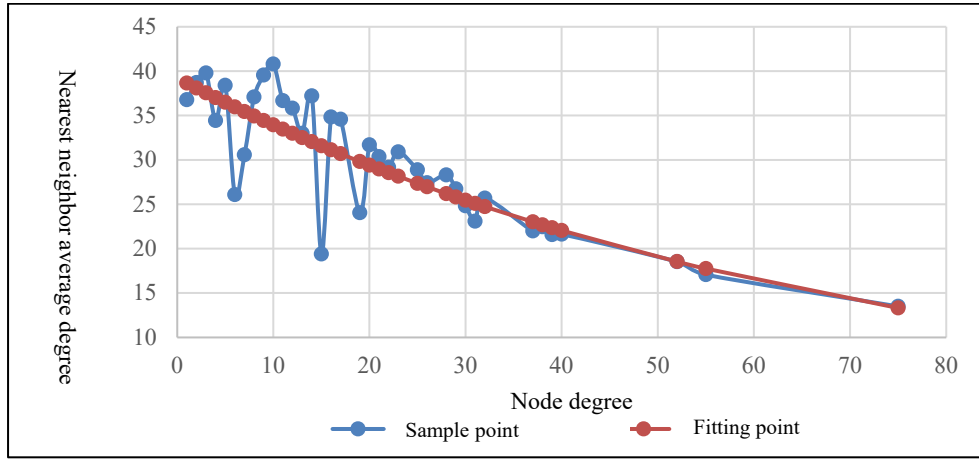


Fig. 8 Fitting nonlinear relationship between node degree and nearest neighbor average degree in 1994.

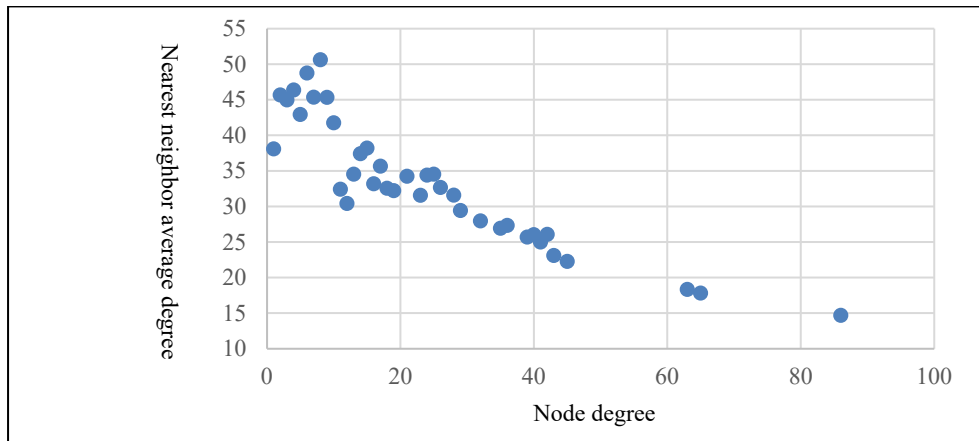


Fig. 9 Diagram of relationship between node degree and nearest neighbor average degree in 2001.

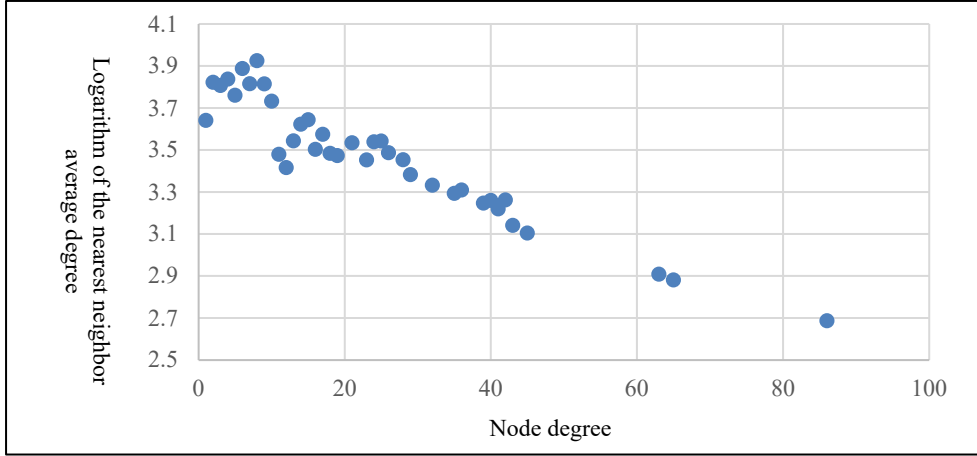


Fig. 10 Diagram of relationship between node degree and logarithm of nearest neighbor average degree in 2001.

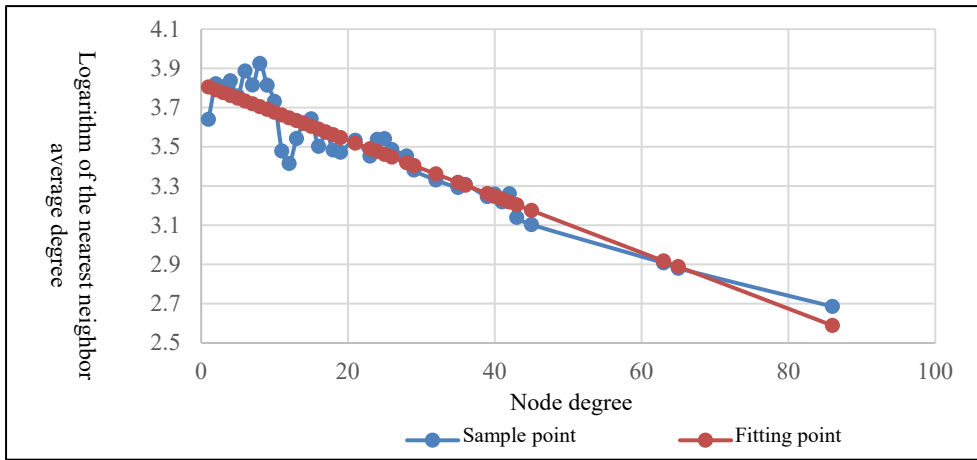


Fig. 11 Fitting linear relationship between node degree and logarithm of nearest neighbor average degree in 2001.

The points of the fitting curve Eq. (16) 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 negative exponential relationship with the nearest neighbor average degree.

2.4 Nonlinear Relationship between Node Degree and Nearest Neighbor Average Degree of China Aviation Network in 2008

According to the data of China aviation network in 2008, $N_{2008}=150$. Using the adjacency matrix A_{2008} and Eq. (2), there were $n=42$ different degrees of 150 nodes in 2008 China aviation network. The nearest neighbor average degree corresponding to each degree could be obtained by using Eq. (3). Thus, there were 42 pairs of data, forming functional relationships, as shown in Fig. 13.

Let the node degree be x axis and the nearest neighbor average degree be y axis in Fig. 13. Let $v = \ln y$ in Fig. 14. The points in Fig. 14 were calculated by the points in Fig. 13. The correlation coefficient r of the points in Fig. 14 was calculated by Eq. (4), $r = -0.963$. Here, $n = 42$. The critical value of $r_{\alpha=1\%, f=40}$ was 0.393 found in critical value table at degree of freedom $f = n - 2 = 40$ and level of significant α of 1%. Since $|r| = 0.963 > 0.393 = r_{\alpha=1\%, f=40}$, the scattered points in

Fig. 14 had significant linear correlation. Least square method was used as an approach in Eq. (17) to fit the line with points in Fig. 14.

$$\begin{cases} \hat{\beta}_0 = \bar{v} - \hat{\beta}_1 \bar{x} = 4.002 \\ \hat{\beta}_1 = \frac{L_{xv}}{L_{xx}} = -0.0133 \end{cases} \quad (17)$$

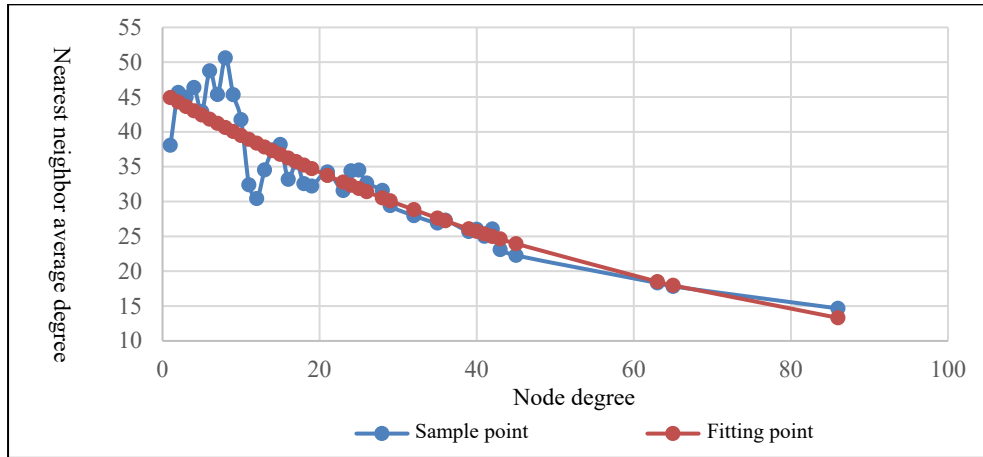


Fig. 12 Fitting nonlinear relationship between node degree and nearest neighbor average degree in 2001.

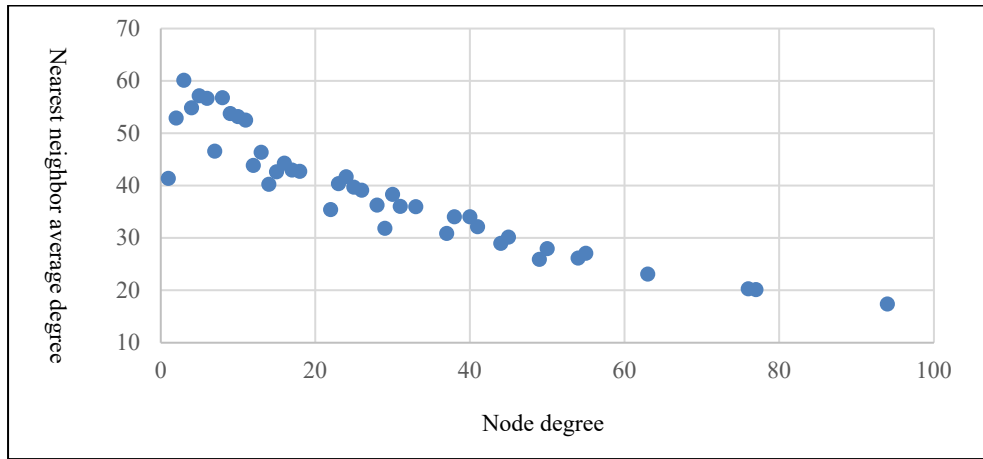


Fig. 13 Diagram of relationship between node degree and nearest neighbor average degree in 2008.

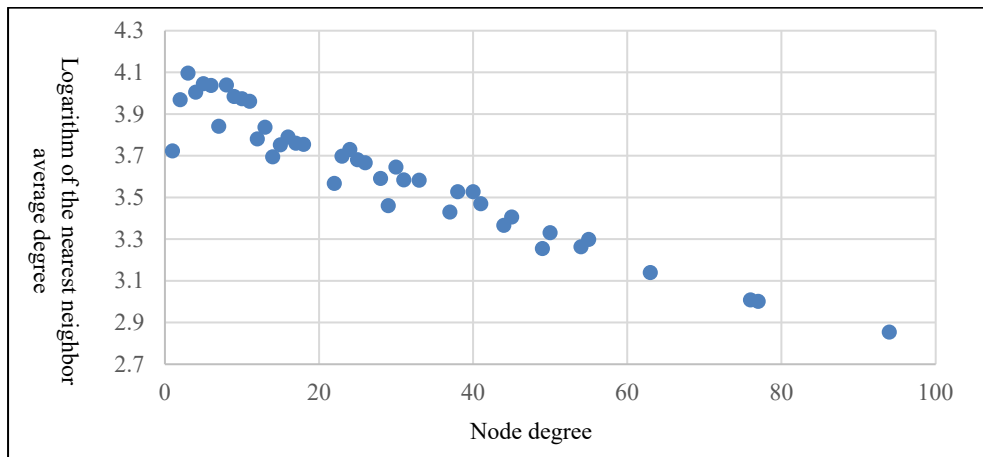


Fig. 14 Diagram of relationship between node degree and logarithm of nearest neighbor average degree in 2008.

The linear equation:

$$\hat{v} = -0.0133x + 4.002 \quad (18)$$

The fitting line Eq. (18) was drawn with the sample points in Fig. 15 with good fitting effect.

To take t test of Eq. (18), test hypothesis is: $H_0: \beta_1 = 0$. When the hypothesis was true, there is Eq. (7). The statistic t was calculated by Eqs. (8) and (9), $t = -22.62$. To check the t distribution table, at

significant level α of 0.01 and degree of freedom $f = n - 2 = 40$, the value of $t_{\alpha=0.01, f=40}$ in table is

2.423. So, $|t| = 22.62 > 2.423 = t_{\alpha=0.01, f=40}$, null hypothesis H_0 is refused. The linear correlation of Eq. (18) is significant. The fitting curve Eq. (19) of nonlinear relationship between node degree and the nearest neighbor average degree was deduced from Eq. (18):

$$\hat{y} = 54.707e^{-0.0133x} \quad (19)$$

The points of the fitting curve Eq. (19) 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 negative exponential relationship with the nearest neighbor average degree.

2.5 Nonlinear Relationship between Node Degree and Nearest Neighbor Average Degree of China Aviation Network in 2015

According to the data of China aviation network in 2015, $N_{2015} = 203$. Using the adjacency matrix A_{2015} and Eq. (2), there were $n = 59$ different degrees of 203 nodes in 2015 China aviation network. The nearest neighbor average degree corresponding to each degree could be obtained by using Eq. (3). Thus, there were 59 pairs of data, forming functional relationships, as shown in Fig. 17.

Let the node degree be x axis and the nearest neighbor average degree be y axis in Fig. 17. Let $v = \ln y$ in Fig. 18. The points in Fig. 18 were calculated by the points in Fig. 17. The correlation coefficient r of the points in Fig. 14 was calculated by Eq. (4), $r = -0.968$. Here, $n = 59$. The critical

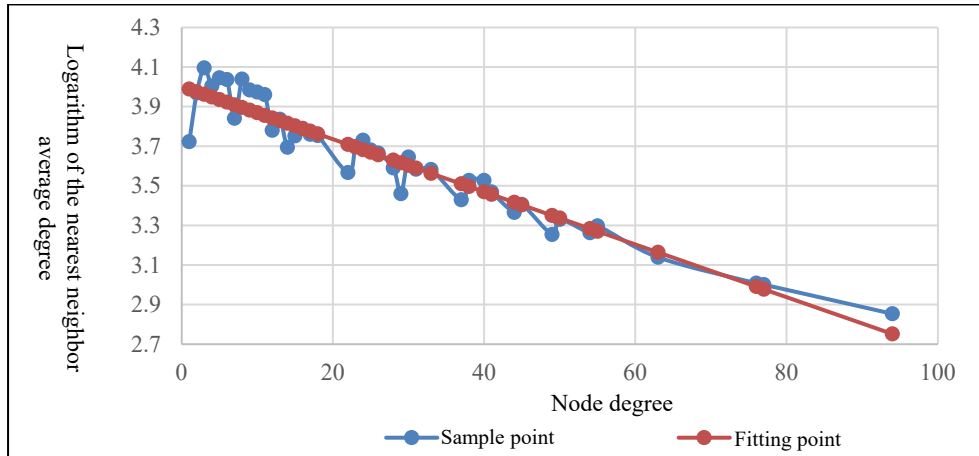


Fig. 15 Fitting linear relationship between node degree and logarithm of nearest neighbor average degree in 2008.

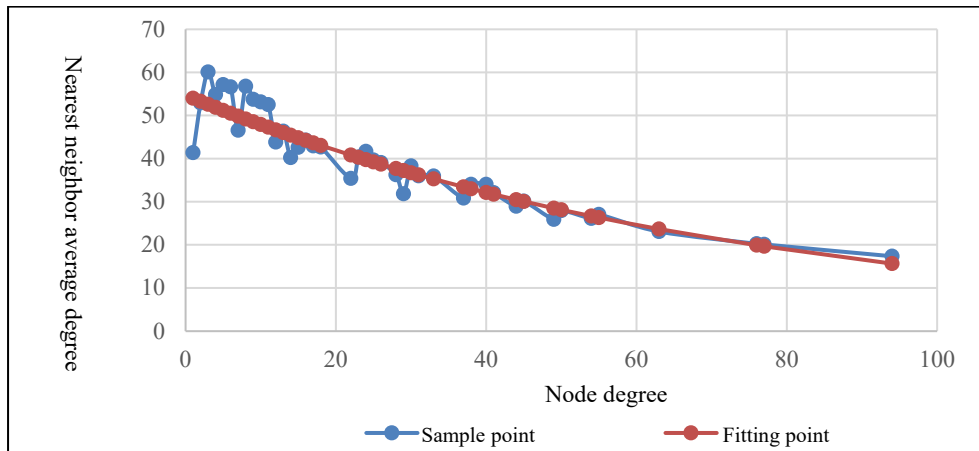


Fig. 16 Fitting nonlinear relationship between node degree and nearest neighbor average degree in 2008.

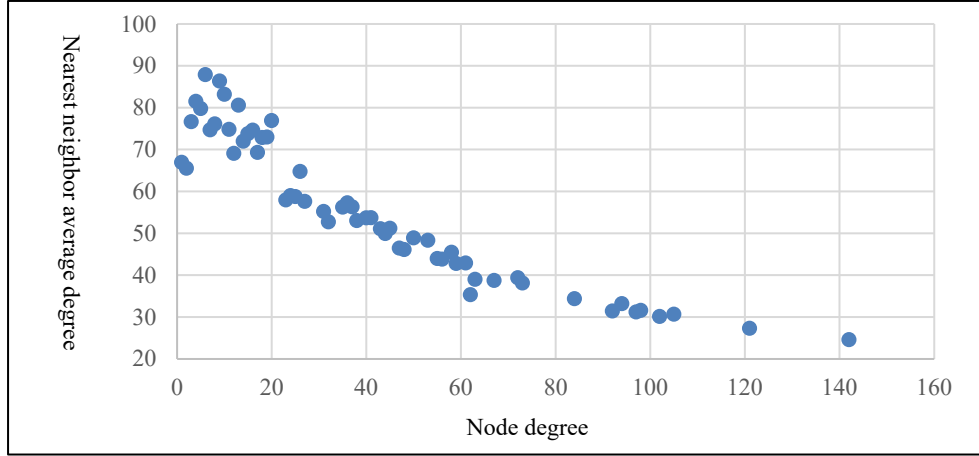


Fig. 17 Diagram of relationship between node degree and nearest neighbor average degree in 2015.

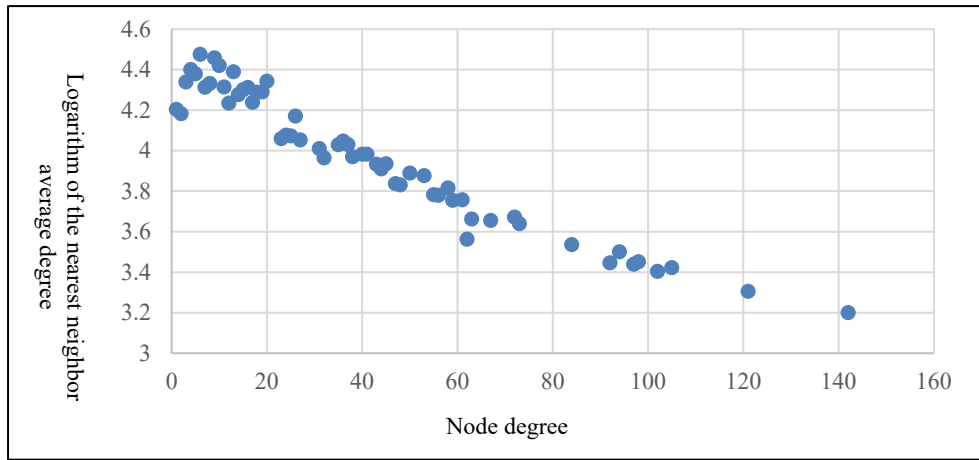


Fig. 18 Diagram of relationship between node degree and logarithm of nearest neighbor average degree in 2015.

value of $r_{\alpha=1\%, f=57}$ was 0.334 found in critical value table at degree of freedom $f = n - 2 = 57$ and level of significant α of 1%. Since $|r| = 0.968 > 0.334 = r_{\alpha=1\%, f=57}$, the scattered points in Fig. 18 had significant linear correlation. Least square method was used as an approach in Eq. (20) to fit the line with points in Fig. 18.

$$\begin{cases} \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 4.378 \\ \hat{\beta}_1 = \frac{L_{xy}}{L_{xx}} = -0.00968 \end{cases} \quad (20)$$

The linear equation:

$$\hat{y} = -0.00968x + 4.378 \quad (21)$$

The fitting line Eq. (21) was drawn with the sample points in Fig. 19 with good fitting effect.

To take t test of Eq. (21), test hypothesis is: $H_0: \beta_1 = 0$. When the hypothesis was true, there is Eq. (7). The statistic t was calculated by Eqs. (8) and (9), $t = -29.1$. To check the t distribution table, at significant level α of 0.01 and degree of freedom $f = n - 2 = 57$, the value of $t_{\alpha=0.01, f=57}$ in table is 2.395. So, $|t| = 29.1 > 2.395 = t_{\alpha=0.01, f=57}$, null hypothesis

H_0 is refused. The linear correlation of Eq. (21) is significant. The fitting curve Eq. (22) of nonlinear relationship between node degree and the nearest neighbor average degree was deduced from Eq. (21):

$$\hat{y} = 79.679e^{-0.00968x} \quad (22)$$

The points of the fitting curve Eq. (22) and the sample points were drawn in Fig. 20 with good fitting

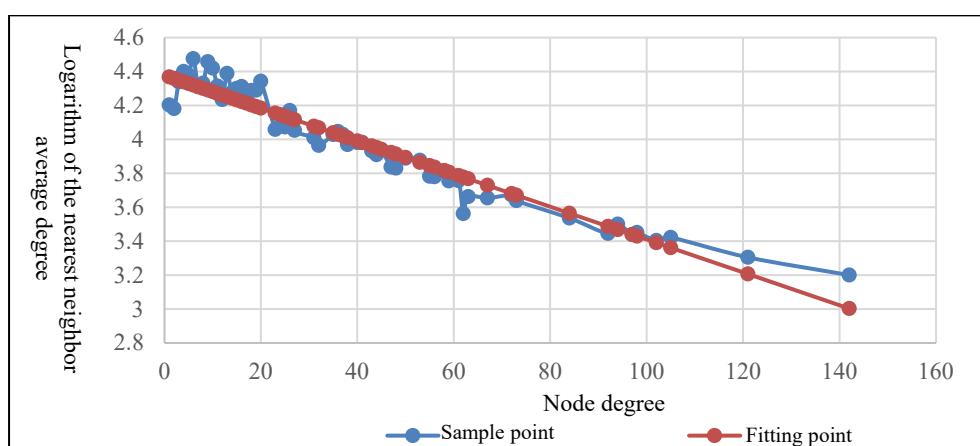


Fig. 19 Fitting linear relationship between node degree and logarithm of nearest neighbor average degree in 2015.

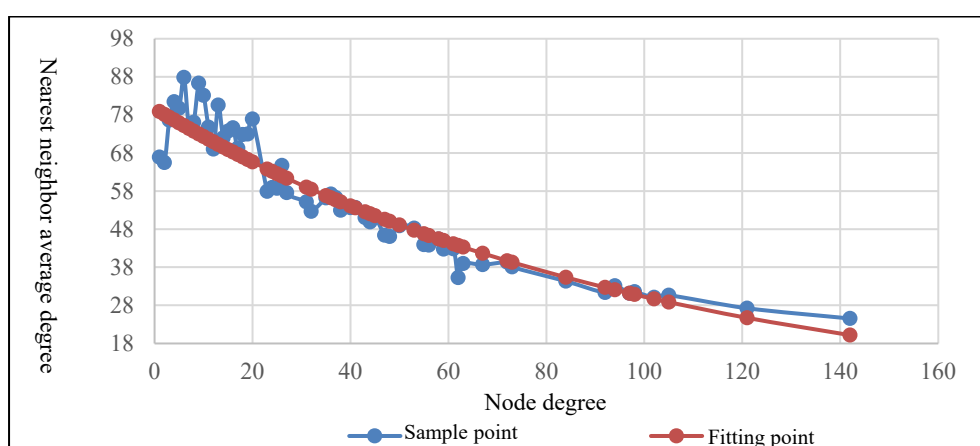


Fig. 20 Fitting nonlinear relationship between node degree and nearest neighbor average degree in 2015.

effect. It showed that the node degree in China aviation network in 2015 had a negative exponential relationship with the nearest neighbor average degree.

3. Evolution of Relationship between Node Degree and Nearest Neighbor Average Degree of China Aviation Network

3.1 Negative Exponential Relationship Evolution between Node Degree and Nearest Neighbor Average Degree of China Aviation Network

The fitting curve Eq. (10), Eq. (13), Eq. (16), Eq. (19) and Eq. (22) of relationship between node degree and the nearest neighbor average degree of China aviation network in 1988, 1994, 2001, 2008 and 2015 were drawn in Fig. 21.

In Fig. 21, the relationship curves between the node degree and the nearest neighbor average degree from

1988 to 2015 were arranged in parallel from bottom to top. They all had a negative exponential relationship, and as node degree increased, the average degree of the nearest neighbor average degree slowly decreased. The change of curve spacing was positively correlated with the change in the number of nodes in the aviation network.

3.2 Parameter Evolution of Negative Exponential Relationship between Node Degree and Nearest Neighbor Average Degree 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: (23.807, -0.0213), (39.212, -0.0144), (45.559, -0.0143), (54.707, -0.0133), (79.679, -0.00968). These five points were drawn in Fig. 22. The correlation coefficient r of the points in

Fig. 22 was calculated, $r = 0.929$. Here, $n = 5$. The critical value of $r_{\alpha=5\%, f=3}$ was 0.878 found in critical value table at degree of freedom $f = n - 2 = 3$ and level of significant α of 5%. Since $|r| = 0.929 > 0.878 = r_{\alpha=5\%, f=3}$, the scattered points in Fig. 22 had relatively significant linear correlation. Least square method was used as an approach in Eq. (23) to fit the line with points in Fig. 22.

$$\begin{cases} \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = -0.0238 \\ \hat{\beta}_1 = \frac{L_{xy}}{L_{xx}} = 0.000189 \end{cases} \quad (23)$$

The linear equation:

$$\hat{y} = 0.000189x - 0.0238 \quad (24)$$

The fitting line Eq. (24) was drawn with the sample points in Fig. 23 with good fitting effect.

To take t test of Eq. (24), test hypothesis is: $H_0: \beta_1 = 0$. When the hypothesis was true, there is Eq. (7). The statistic t was calculated by Eqs. (8) and (9), $t = 4.37$. To check the t distribution table, at significant level α of 0.05 and degree of freedom $f = n - 2 = 3$, the value of $t_{\alpha=0.05, f=3}$ in table is 2.353.

So, $|t| = 4.37 > 2.353 = t_{\alpha=0.05, f=3}$, null hypothesis H_0

is refused. The linear correlation of Eq. (24) is relatively significant.

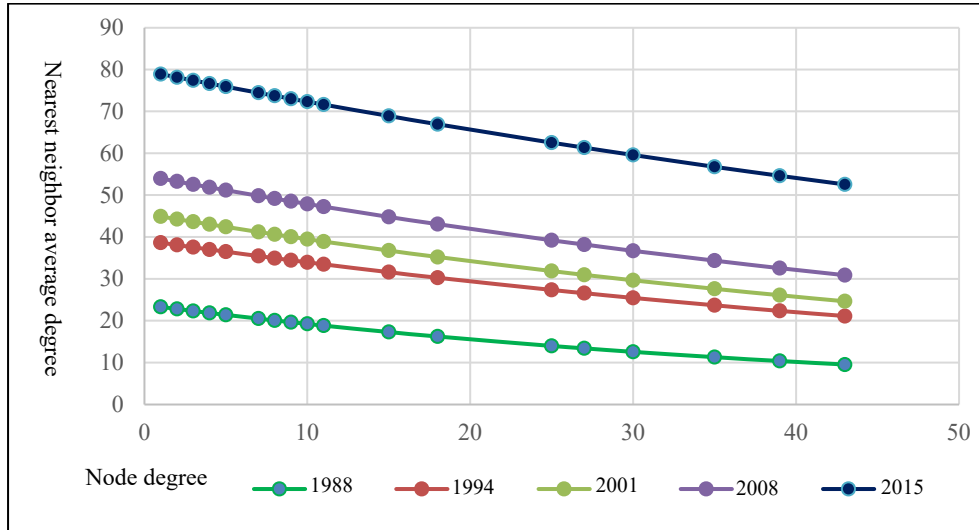


Fig. 21 Evolution diagram of negative exponential relationship between node degree and nearest neighbor average degree of China aviation network.

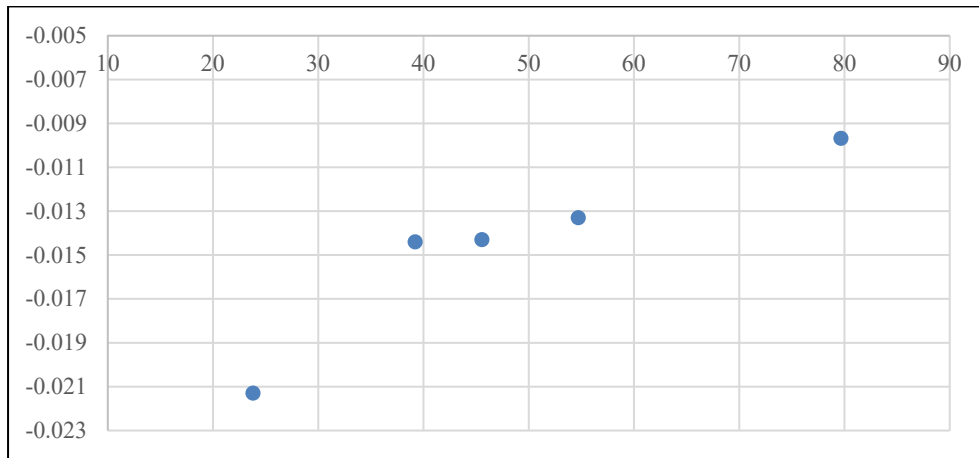


Fig. 22 Diagram of relationship between the two parameters of negative exponential curve.

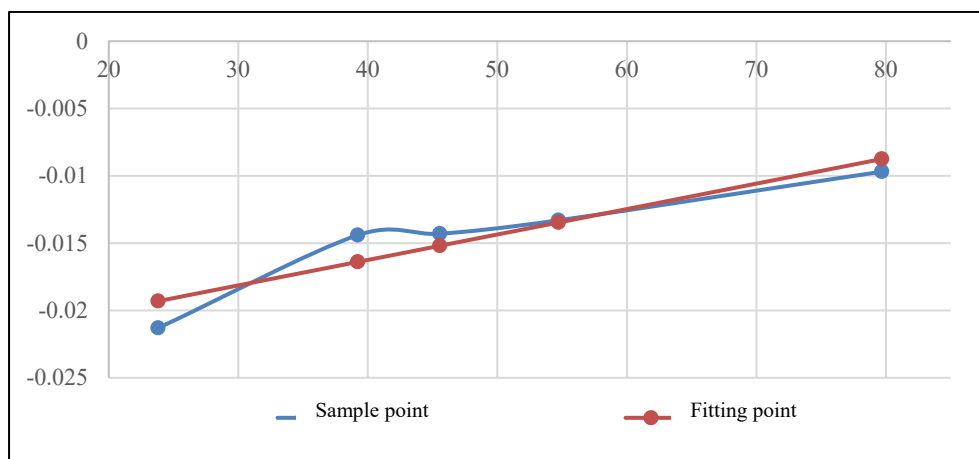


Fig. 23 Fitting linear relationship between the two parameters of negative exponential curve.

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 nearest neighbor average degree and relationship evolution trace of China aviation network in 1988, 1994, 2001, 2008 and 2015 were studied. According to the statistical data, it was found that the node degree had a negative exponential relationship with the nearest neighbor average degree and the two parameters of the negative exponential relationship of these five years had linear evolution 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)