ISSN 1749-3889 (print), 1749-3897 (online)
International Journal of Nonlinear Science
Vol.31(2021) No.1,pp.51-57

Horizontal Visibility Graph based on Correlation Coefficient and Its Application

Jiyun Zhang*
Energy Development and Environmental Protection Strategy Research Center, Jiangsu University, Zhenjiang, Jiangsu 212013, China

(Received 3 January 2021, accepted 10 March 2021)

Abstract: In this paper, horizontal visibility graph based on correlation coefficient is proposed, and its degree distribution is evaluated by combining the existing general visibility graph and horizontal visibility graph. The results show that the improved visibility graph algorithm can effectively extract the complex network features of white noise time series, periodic time series, fractal time series and chaotic time series, and the anti noise ability is more significant. On this basis, the new visibility graph algorithm is applied to the carbon price data of the EU carbon market, and the dynamic evolution process and the linkage between them are obtained, which is convenient to make a more accurate prediction of the future EU carbon market.

Keywords: Complex networks; visibility graph; EU carbon market

1 Introduction

With the advent of the era of big data, the theory and technology of time series data mining are of great significance. The small world complex network model (WS) [1] established by Watts and Strogatz and BA scale-free network [2] studied by Barabasi and Albert lay the foundation for the research of complex networks. Subsequently, scholars have proposed many methods to map time series to complex networks [3–10], and visibility graph algorithm is one of the most effective algorithms [4].

Lacasa et al. proposed a general visibility graph (VG), whose visibility criterion is: if point $(\boldsymbol{t}_a, \boldsymbol{y}_a)$ and point $(\boldsymbol{t}_b, \boldsymbol{y}_b)$ in discrete time series are connected, then for any point $(\boldsymbol{t}_c, \boldsymbol{y}_c)$, where $\boldsymbol{t_a} < \boldsymbol{t_c} < \boldsymbol{t_b}$, satisfies $\frac{x_a - x_c}{t_c - t_a} > \frac{x_a - x_b}{t_b - t_a}$ [4]. Then, Luque et al. proposed the horizontal visibility graph (HVG), whose visibility criterion is: if the point $(\boldsymbol{t}_a, \boldsymbol{y}_a)$ and the point $(\boldsymbol{t}_b, \boldsymbol{y}_b)$ in the discrete time series are connected, then for any point $(\boldsymbol{t}_c, \boldsymbol{y}_c)$, where $\boldsymbol{t_a} < \boldsymbol{t_c} < \boldsymbol{t_b}$ satisfies $\boldsymbol{y_a}, \boldsymbol{y_b} > \boldsymbol{y_c}$ [11]. These visibility graph algorithms have been successfully implemented in many fields [12–14]. After that, Zhou et al. proposed an improved time series visibility graph network construction method, namely limited penetrable visibility graph (LPVG) [15], which has been successfully used to analyze various actual signals in different fields [16–19]. Recently, Lacasa et al. proposed some topological properties and analysis results related to visibility graph [20, 21]. Then, Wang et al. focused on a kind of general horizontal visibility graph algorithm, namely limited penetrable horizontal visibility graph (LPHVG), and extended the visibility graph to establish a directed-limited penetrable horizontal visibility graph (DLPHVG) and an image-limited penetrable horizontal visibility graph LPHVG $_n$ [22].

Although the existing visibility graph algorithms can inherit the inherent morphological characteristics of time series, the corresponding performance such as anti noise performance and algorithm complexity have some limitations. In the process of algorithm construction, this paper proposes an improved visibility graph network construction method, which combines sliding window and correlation coefficient, that is horizontal visibility graph based on correlation coefficient, and evaluates its model. Then the new visibility graph algorithm is applied to the EU carbon price data to analyze the linkage between the EU carbon markets from a new perspective, so as to make a more accurate prediction of the future EU carbon market.

The structure of this paper is as follows: the second part constructs the horizontal visibility graph based on correlation coefficient. In the third part, we evaluate the degree distribution of general visibility graph and horizontal visibility graph,

***Corresponding author**. *E-mail address*: 793451829@qq.com

IJNS.2021.03.15

and draw a conclusion. The fourth part applies the new visibility graph algorithm to the carbon price data of EU carbon market. The fifth part expounds the conclusion.

2 Model

The main steps of the visibility graph network based on correlation coefficient are as follows:

(1) Construct sliding window If the length of the time series is n, the window length is L and the sliding step is s, the number of windows is N

$$N = \frac{(n-L)}{s} + 1; \tag{1}$$

(2) Use the formula

$$\rho_{X_1X_2} = \frac{\mathrm{Cov}\left(X_1, X_2\right)}{\sigma_{X_1}\sigma_{X_2}} = \frac{E\left[\left(X_1 - \mu \mathrm{x}\right)\left(X_2 - \mu \mathrm{y}\right)\right]}{\sigma_{X_1}\sigma_{X_2}} \tag{2}$$

where

$$\begin{aligned} \mathrm{Cov}\left(X_1, X_2\right) &= \mathrm{E}\left[\left(X_1 - \mathrm{E}\left[X_1\right]\right)\left(X_2 - \mathrm{E}\left[X_2\right]\right)\right] \\ &= \mathrm{E}\left[X_1X_2\right] - 2\mathrm{E}\left[X_1\right]\mathrm{E}\left[X_2\right] + \mathrm{E}\left[X_1\right]\mathrm{E}\left[X_2\right] = \mathrm{E}\left[X_1X_2\right] - \mathrm{E}\left[X_1\right]\mathrm{E}\left[X_2\right]; \end{aligned} \tag{3}$$

$$\sigma_{X_1} = \mathrm{sqrt}\left(\frac{\left(\mathrm{x}_1 - \mathrm{E}\left[X_1\right]\right)^2 + \left(\mathrm{x}_2 - \mathrm{E}\left[X_1\right]\right)^2 + \cdots + \left(\mathrm{x}_{100} - \mathrm{E}\left[X_1\right]\right)^2}{100}\right); \tag{4}$$

$$\sigma_{X_2} = \mathrm{sqrt}\left(\frac{\left(\mathrm{x}_2 - \mathrm{E}\left[X_2\right]\right)^2 + \left(\mathrm{x}_2 - \mathrm{E}\left[X_1\right]\right)^2 + \cdots + \left(\mathrm{x}_{101} - \mathrm{E}\left[X_2\right]\right)^2}{100}\right); \tag{5}$$

Calculate the correlation coefficient between any two windows to get the correlation coefficient matrix

$$C = \begin{bmatrix} c_{11} & \cdots & c_{1N} \\ \vdots & \ddots & \vdots \\ c_{N1} & \cdots & c_{NN} \end{bmatrix}; \tag{6}$$

(3) The correlation coefficient sequence $C(t), t = 1, 2, ..., N$, is obtained by calculating the mean value in columns from the correlation coefficient matrix;

(4) The new time series $C(t)$ is applied to VG or HVG algorithm to get a new visibility graph network.

This method obtains the network which is composed of the correlation of the original data. In step (4), the network constructed by VG algorithm [4] is called CVG network; the network constructed by HVG algorithm [11] is called chvg network.

In fact, there is another way to build a visibility graph network based on correlation coefficient. On the basis of the correlation coefficient matrix C obtained in step (2), the CIVG network or CIHVG network can be obtained by processing the correlation coefficient matrix with the help of IVG or IHVG algorithm.

For example, taking 200 sample data at random, using the above steps (1-2) and combining with IHVG algorithm, the CIHVG $_4$ can be obtained. Between x $_{ij}$ and x $_{i'j'}$, if i=i', for any x $_{ip}$, where $p \in (j, j')$ meets $x_{ij}, x_{i'j'} > x_{ip}$, then x $_{ij}$ is connected with x $_{i'j'}$; if j=j', for any x $_{qj}$, where $p \in (i, i')$ meets $x_{ij}, x_{i'j'} > x_{qj}$, then x $_{ij}$ is connected with x $_{i'j'}$.

Obviously, the computational complexity of CIVG or CIHVG is much higher than that of CVG or CHVG, so we choose CHVG for application analysis.

3 Model evaluation

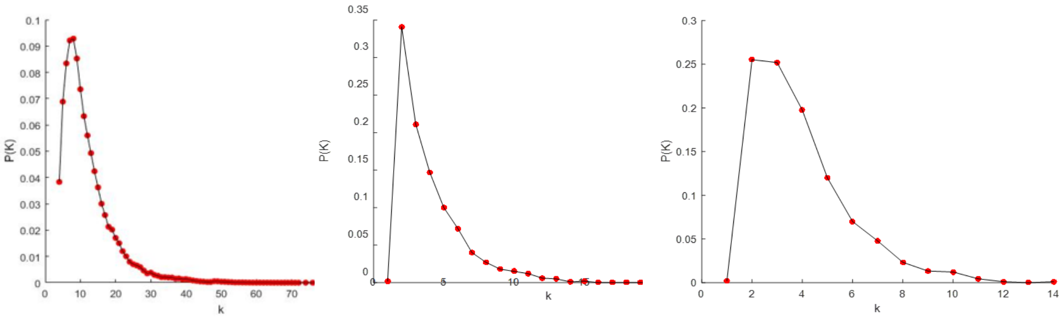
In the aspect of pattern test, this paper selects Gaussian white noise time series, periodic time series, Conway fractal time series and Lorenz chaotic time series.

3.1 White noise time series

Select Gaussian white noise signal

$$x_t \sim N_G\left(0, \sigma^2\right) \tag{7}$$

The following is a simple generation of a group of sequences. The graph of this group of sequence data obeys the positive too distribution. This is the pure random sequence we want to take. That is to say, a 1000 normal distribution time series with mean value of 0 is also called Gaussian white noise. Combined with CHVG, the degree distribution map was generated.



(a) (b) (c)

Figure 1: Degree distribution of Gaussian white noise sequence.

Figure (a) shows the degree distribution of Gaussian white noise sequence of VG; Figure (b) shows the degree distribution of Gaussian white noise sequence of HVG; Figure (c) shows the degree distribution of Gaussian white noise sequence of CHVG.

As can be seen from Figure (a), the connectivity between each node in VG network is relatively close, and the network degree distribution is also relatively scattered, which is in the form of power law. Figure (b) is the network degree distribution diagram after the transformation of HVG. It can be seen that the average degree of each node in the horizontal visibility graph complex network is smaller than that of VG, and it is generally $K \leq 10$, mainly concentrated around $K = 10$, and the network degree distribution is more concentrated. As can be seen from Figure (c), the Gaussian white noise time series transformed by CHVG algorithm has a power-law network degree distribution, and the network degree distribution is relatively scattered. The network degree values are mainly between 2 and 3 points, and the network degree distribution is also very centralized, belonging to a regular network.

3.2 Periodic time series

Select sinusoidal signal

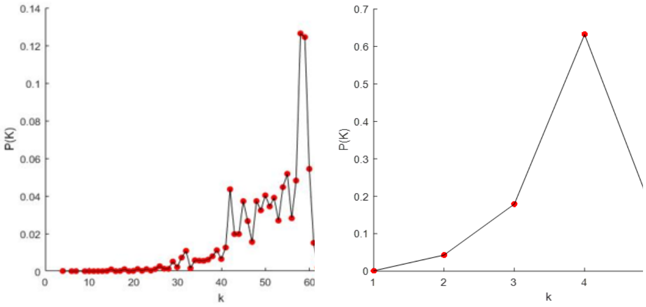
$$y = \sin(5\pi x) \tag{8}$$

The sampling interval is $\Delta x = 0.01$ and the sequence length is $n = 1000$.

The following is the degree distribution map after the period time series of these algorithms are transformed into a complex network.

Figure (a) shows the degree distribution of the periodic sequence with the sinusoidal signal of VG as $y = \sin(5\pi x)$; Figure (b) shows the degree distribution of the periodic sequence with the sinusoidal signal of HVG as $y = \sin(5\pi x)$; Figure (c) shows the degree distribution of the periodic sequence with the sinusoidal signal of CHVG as $y = \sin(5\pi x)$.

It can be seen from Figure (a) that the connectivity of the nodes in the transformed network is relatively close, and the degree distribution of the network has a power-law function. However,the degree values of the nodes are relatively scattered, belonging to a regular network with scattered degree values, and the degree distribution presents a multi peak shape as shown in the Figure above. Figure (b) is the graph drawn by the degree distribution function of the network nodes after the HVG transformation. It can be seen that the average degree of each node in the horizontally visibility graph complex network is smaller than VG, and it is generally $K \leq 10$. The connectivity between the network nodes is



(a) (b) (c)

Figure 2: Degree distribution of periodic sequence.

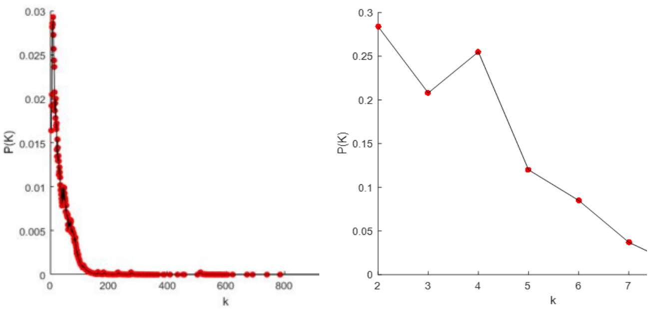
relatively sparse, and the network degree values are concentrated at the point 4 of the figure, which belongs to a regular network. It can be seen from the figure above that the number of spikes of its degree distribution function is much less than that of VG network. As can be seen from Figure (c), the degree distribution of the periodic sequence transformed by CHVG algorithm is also in the form of power law, and the connectivity between nodes is relatively sparse, and the degree value of the network is concentrated at 4.

3.3 Fractal time series

Conway fractal sequence is selected, and its recursive formula is

$$\begin{gathered} a(1) = a(2) = 1, \\ a(n) = a(a(n-1)) + a(n - a(n-1)), n > 2; \end{gathered} \tag{9}$$

The following is the degree distribution map after the fractal time series of these algorithms are transformed into complex networks.



(a) (b) (c)

Figure 3: Degree distribution of Conway fractal sequence.

Figure (a) is the degree distribution of Conway fractal sequence of VG; figure (b) is the degree distribution of Conway fractal sequence of HVG; and figure (c) is the degree distribution of Conway fractal sequence of CHVG.

Figure (a) shows that the network degree distribution of VG algorithm has a power-law form, and the network degree values are concentrated between 0 and 80. As can be seen from the HVG algorithm in Figure (b), the average degree value of the network nodes transformed by the HVG algorithm is generally smaller, the connectivity between the network nodes is relatively sparse, and the network degree value is mainly concentrated between 2 and 4. As can be seen from Figure (c), the network degree distribution of Conway time series also has a power-law form after CHVG algorithm transformation, which means that Conway can become a scale-free network after chvg algorithm transformation, that is, CHVG has certain recognition ability for Conway series. The degree value of the network is concentrated at 4.

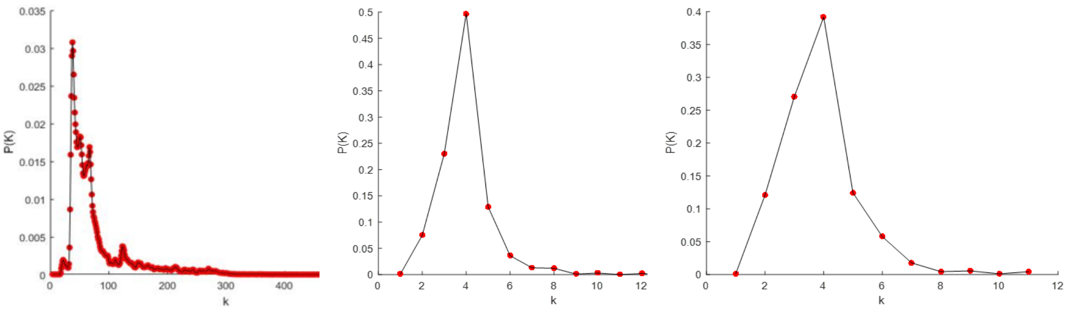
3.4 Chaotic time series

Lorenz chaotic sequence is selected

$$\begin{cases} \dot{x} = \sigma(y - x) \\ \dot{y} = x(r - z) - y \\ \dot{z} = xy - bz \end{cases} \tag{10}$$

where $\sigma = 10, \quad r = 28, \quad b = 8/3$, sampling interval $t = 0.05$.

The following is the degree distribution map after the chaotic time series of these algorithms are transformed into a complex network.



(a) (b) (c)

Figure 4: Degree distribution of Lorenz chaotic sequence.

Figure (a) shows the degree distribution of Lorenz chaotic sequence of VG; Figure (b) shows the degree distribution of Lorenz chaotic sequence of HVG; Figure (c) shows the degree distribution of Lorenz chaotic sequence of CHVG.

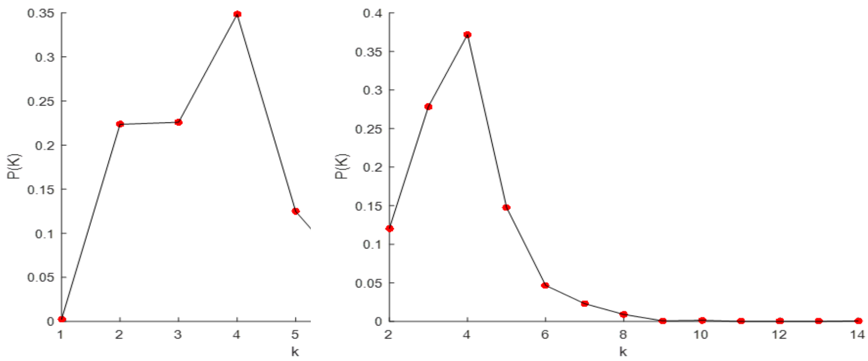
As can be seen from figure (a) above, the network degree distribution of VG algorithm has a power-law form in signal feature recognition, that is, Conway sequence can be transformed into scale-free network through VG algorithm, which shows that VG algorithm can effectively identify Conway time series. In addition, compared with its network degree value distribution from 0 to 80, it is relatively dispersed, and also presents a multi peak shape. As can be seen from the HVG algorithm in figure (b), the average degree value of the network nodes transformed by the HVG algorithm is generally smaller, the connectivity between network nodes is relatively sparse, and the network degree values are mainly concentrated between 2 and 8, which indicates that the HVG algorithm transforms the Conway time series into a regular complex network. We can find that the distribution function of the degree distribution function of the Conway time series transformed by HVG is also composed of a series of spikes, so we can judge that the HVG algorithm cannot obviously improve the connectivity the signal characteristics of Conway fractal time series are identified. Figure (c) after CHVG algorithm transformation, the network degree distribution of Lorenz chaotic time series has the same power-law form, and the degree value of the network is concentrated at 4.

4 Analysis of EU carbon price data based on CHVG

This paper selects the carbon futures price data of the second stage (2010/12/13-2012/12/31) and the third stage (2013/1/1-2018/12/27) of the EU carbon emission trading market as the sample data, and uses CHVG algorithm to obtain the network topology index (degree distribution, aggregation coefficient and average path length) for the data of the two stages.

Figure (a) shows the degree distribution of the second stage of EU carbon emission trading market; Figure (b) shows the degree distribution of the third stage of EU carbon emission trading market.

After the transformation of CHVG algorithm, the time series of the two stages have the power-law network degree distribution, and the peak value of the two stages is at 4, and the degree value is basically concentrated in 2-6. Within a certain range, the nodes with large degree are more affected by the price fluctuation in the early stage or the subsequent price waveform; the nodes with small degree have less influence. The network degree distributions of the two stages are very similar, which indicates that the fluctuation of EU carbon price has self similarity and scale invariance, that is, the fluctuation of future carbon price time series is likely to be similar to that of a certain period in the past.



(a) (b)

Figure 5: Degree distribution of carbon price data.

Table 1: Topological Characteristics of new viewable network of carbon price in EU carbon market

	Clustering coefficient	Average path length	Average degree	Network diameter
The second stage	0.4597	51.1024	3.6329	125
The third stage	0.5091	57.1142	3.8454	141

Using the obtained complex network, we use other basic concepts such as aggregation coefficient, average path length, average degree and network diameter to describe the topological properties of the complex network mapped from time series. Table 1 shows the topology measurement of the constructed CHVG network, and the observation shows that the relationship between the network structure and network dimensions is heterogeneous.

Using CHVG algorithm to explore the dynamic characteristics of carbon price, one is that CHVG can capture more accurate information of data state, extract the subtle information of each sequence segment on different scales, and keep the extraction process reasonable and simple; the other is that the time series is composed of a series of numerical values, closely related to time, and has a certain random process, usually non-stationary VG can be used to analyze this kind of nonstationary series. Third, CHVG calculates the correlation between each node, which can reflect the correlation between data more accurately and give more accurate prediction.

5 Conclusions

In this paper, we propose a correlation coefficient based horizontal visibility graph network construction method (CHVG), which is an improved algorithm based on HVG. The experimental results show that, compared with VG and HVG algorithm, the average degree of nodes in the complex network transformed by CHVG algorithm is larger, the distribution of network degree is more dispersed, the network connectivity is closer, and it can better reflect the characteristics of time series. Compared with VG and HVG, the new visibility graph method can extract complex network features of noise, period, fractal and chaotic signals more effectively, and has more significant anti noise ability. The experimental results show the effectiveness of the new visibility graph method.

The CHVG network method is applied to the carbon price data of EU carbon market, and the network degree distribution characteristic parameters such as degree distribution, average degree, aggregation coefficient and average path length are extracted. The results show that the carbon price data of the two stages of the EU carbon market have their own similarity and scale invariance, so the fluctuation of carbon price in the future is likely to be similar to that in the past.

6 Acknowledgments

This research is supported by grants from the Major Program of National Natural Science Foundation of China (No. 71690242), the National Natural Science Foundation of China (No. 11731014), the National Key Research and Development Program of China (No. 2020YFA0608601) and Jiangsu Key Lab for NSLSCS (No. 202006).

References

[1] D. J. Watts, S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature* 393 (1998) 440-442.
[2] B. Albert-Lszl, A. Rka. Emergence of Scaling in Random Networks. *Science* 286.
[3] J. Zhang, M. Small. Complex Network from Pseudoperiodic Time Series: Topology versus Dynamics. *Physical Review Letters* 96 (2006) 238701.
[4] L. Lacasa, B. Luque ,F. Ballesteros et al. From time series to complex networks: The visibility graph. *Proceedings of the National Academy of Sciences of the United States of America* 105 (2008) 4972-4975.
[5] N. Marwan, J. F. Donges, Y. Zou et al. Complex network approach for recurrence analysis of time series. *Physics Letters A* 373 (2009) 4246-4254.
[6] A. H. Shirazi, G. Reza Jafari, J. Davoudi et al. Mapping stochastic processes onto complex networks. *Journal of Statistical Mechanics: Theory and Experiment* 7 (2009) 7046.
[7] J. F. Donges, Y. Zou, N. Marwan et al. Backbone of the climate network. *Europhysics Letters* 87 (2010).
[8] Y. Zhao, T. Weng and S. Ye. Geometrical invariability of transformation between a time series and a complex network. *Physical Review E* 90 (2014).
[9] V. Stolbova, P. Martin, B. Bookhagen et al. Topology and seasonal evolution of the network of extreme precipitation over the Indian subcontinent and Sri Lanka. *Nonlinear Processes in Geophysics* 21 (2014) 901-917.
[10] M. G. Wang, L. X. Tian. From time series to complex networks: The phase space coarse graining. *Physica A: Statistical Mechanics and its Applications* 461 (2016) 456-468.
[11] B. Luque, L. Lacasa, F. Ballesteros et al. Horizontal visibility graphs: exact results for random time series. *Physical Review E* 80 (2009).
[12] L. Lacasa, B. Luque, J. Luque et al. The visibility graph: A new method for estimating the Hurst exponent of fractional Brownian motion. *Europhysics Letters* 86 (2009) 30001-30005.
[13] G. Gutin, T. Mansour, S. Severini. A characterization of horizontal visibility graphs and combinatorics on words. *Physica A: Statistical Mechanics and its Applications* 390 (2010) 2421-2428.
[14] R. Ray, M. H. Khondekar, K. Ghosh et al. Memory persistency and nonlinearity in daily mean dew point across India. *Theoretical and Applied Climatology* 124 (2016) 119-128.
[15] Z. Tingting, J. Ningde, G. Zhongke et al. Limited penetrable visibility graph for establishing complex network from time series. *Acta Physica Sinica* 61 (2012) 30506-030506.
[16] G. Zhong-Ke, H. Li-Dan, Z. Ting-Ting et al. Limited penetrable visibility graph from twophase flow for investigating flow pattern dynamics. *Acta Physica Sinica (Chinese Edition)* 62 (2013).
[17] J. Wang, C. Yang, R. Wang et al. LFunctional brain networks in Alzheimers disease: EEG analysis based on limited penetrable visibility graph and phase space method. *Physica A:Statistical Mechanics and its Applications)* 460 (2016) 174-187.
[18] W. Rongxi, G. Jianmin, G. Zhiyong et al. Complex network theory-based condition recognition of electromechanical system in process industry. *Science China(Technological Sciences)* 59 (2016) 97-110.
[19] Z. K. Gao, L. D. Hu, T. T. Zhou et al. Limited penetrable visibility graph from two-phase flow for investigating flow pattern dynamics. *Acta Physica Sinica* 62 (2013) 97-110.
[20] B. Luque, L. Lacasa, F. J. Ballesteros et al. Analytical properties of horizontal visibility graphs in the Feigenbaum scenario. *Chaos* 22 (2012) 9.
[21] Lacasa, Lucas. On the degree distribution of horizontal visibility graphs associated to Markov processes and dynamical systems: diagrammatic and variational approaches. *Nonlinearity* 27 (2014).
[22] M. G. Wang, A. L. M. Vilela, R. Du et al. Theoretical results on the topological properties of the limited penetrable horizontal visibility graph family. *Physical Review E* (2017).