Network and Complex Systems
ISSN 2224-610X (Paper) ISSN 2225-0603 (Online)
Vol.4, No.7, 2014
www.iiste.org

Threshold grouping method to derive complex networks from time series

Meenatchidevi Murugesan* R. I. Sujith
Indian Institute of Technology Madras, Chennai, 600036, India
* E-mail of the corresponding author: mena.murugesan90@gmail.com

Abstract

We propose a new method to derive complex networks from time series data. Each data point in the time series is treated as a single node and nodes are connected if their values differ by a value less than a threshold ε. The method is easy to implement and transforms periodic time series into regular networks, and random time series into random networks. Network specific properties such as scale freeness, small world effects and the presence of hubs are captured by this method. The method converts a chaotic time series into a scale-free network. The method is applied on a model time series (a chaotic time series of Henon map, a periodic time series and a random time series) and an experimental time series (fluctuating pressure time series measured from the combustor involving turbulent flow). The complex network derived from the Henon map obeys a power law degree distribution, highlighting the scale free behavior of the associated chaotic state. The motivation for proposing the present method is that the method is able to convert specific patterns in the dynamics of a time series into spatial structures in the complex network. We show this specialty of the present method by constructing a complex network from a time series of acoustic pressure measured from a combustor involving turbulent flow that exhibits intermittency. The intermittent bursts in the considered time series are converted into clusters in the corresponding complex network.

Keywords: Time series analysis; Complex networks; Threshold grouping method

1. Introduction

The study of complex networks is a fast growing field with applications in several areas such as biomedical applications (Alm & Arkin 2003; Barabasi & Oltvai 2004), worldwide web networks (Albert *et al.* 1993; Albert *et al.* 2000), power grid networks (Pagani & Aiello 2013; Arianos *et al.* 2009; chen *et al.* 2010) etc. The connectivity among the elements of the network plays an important role in deciding the vulnerability of the network (Arianos *et al.* 2009; chen *et al.* 2010). Often, complex networks are used in modeling the behavior of dynamical systems (Winfree 1980; Kuramoto 1984; Strogatz & Stewart 1993).

The behavior of a dynamical system can be understood by analyzing the time series data of its evolution. Linear time series analysis is very helpful in understanding the nature of a linear system. The nonlinear aspects such as fractality, multifractality, self organization and presence of multiple scales etc. arising out of complex systems are studied using the tools from nonlinear time series analysis (Tan *et al.* 2009; Zorich & Mandelkern 2013; Eser *et al.* 2014).

The representation of a time series from such complex systems as a complex network can provide a better way to visualize hidden patterns in the time series. The presence of multiple scales, fractal behavior etc. can be visual in terms of connections between the nodes in the network representation. The features of the network can be characterized using the properties of the network. The network measures can be useful in quantifying the information present in the time series.

Zhang & Small (2006) proposed a method for constructing complex networks from pseudo-periodic time series. Each cycle in the time series is considered as a single node. For an experimental time series, the connection between any two nodes are formed when the distance between the corresponding two cycles in phase space is less than a threshold. For simple toy models, connectivity is based on the temporal correlation between the two cycles. This method is applicable only for time series with dominant peaks in the power spectra. The partition of a time series into sequence of cycles depends on the dominant frequency in power spectra. Further, in this method, the process of finding connectivity requires either reconstruction of the phase space or a complex calculation of the temporal correlation coefficient between different cycles (Zhang *et al.* 2008).

In an alternative scheme, complex networks called recurrence networks are constructed from time series data based on the recurrence behavior of state points in the phase space (Donner *et al.* 2010). The phase space is reconstructed from the time series and each state point in the phase space is considered as a node. The nodes are connected if the state space distance between them is less than a certain threshold. This method is also complicated because it requires embedding and reconstruction of the state space.

Algorithms for constructing complex networks called visibility graphs and horizontal visibility graphs, based on the visibility of nodes are presented by Lacasa *et al.* (2008) and Luque *et al.* (2009) respectively. In both these methods, data points in the time series are converted into nodes. In a visibility graph, two nodes are connected if a straight line can be drawn between these nodes without intersecting the data points lying between them (Lacasa *et al.* 2008). Horizontal visibility graph is a special case of visibility graph (Luque *et al.* 2009). If a horizontal line can be drawn between the two nodes without intersecting any node between them, the nodes are treated as connected.

Complex calculations are involved in the construction of a network from time series using the methods proposed in references (Zhang & Small 2006; Zhang *et al.* 2008; Donner *et al.* 2010). In contrast, the visibility graph and horizontal visibility graphs are simple. However, the patterns in the time series are not easily recognizable in the corresponding complex networks derived using visibility and horizontal visibility graphs. This will be illustrated in the Section 4.

In this paper, we propose a new method called threshold grouping method to derive complex networks from time series data. The method is easier compared to many of the existing methods. The complex network derived using the proposed method exhibits patterns corresponding to the dynamics of the time series. The data points in the time series are treated as nodes and the connections between the nodes are formed when the difference in the values of nodes is less than a threshold ε. Different groups of nodes are formed by this method and the nodes in each group have approximately the same value (i.e. within ε). The maximum difference between any two nodes in a group is ε.

The proposed idea is simple and the structure of the time series is reflected in the derived complex network. We show that regular networks are formed from periodic time series and random networks are formed from random time series. Moreover, the scale freeness of a complex network derived from a chaotic time series can be captured by this method. The scale free network derived from the chaotic time series of Henon map (Al-Shameri 2012) is discussed in the following sections. The method is then utilized to visualize the distribution of an experimental time series measured from a combustor involving turbulent flow that exhibits intermittency. The observed intermittent time series is composed of bursts of large amplitude periodic fluctuations amidst regions of low amplitude chaotic fluctuations (Nair & Sujith 2013). Different bursts in the time series are converted into different clusters in the complex network.

2. Description of the threshold grouping algorithm

The construction of a complex network using threshold grouping method is illustrated using the Henon map at chaotic state after removing the transients (Al-Shameri 2012). A new time series called $p(t)$ is constructed from the peaks (crest of each cycle) in $x(t)$. In the threshold grouping method, each data bar or data value in the time series is treated as a node. If the difference between the values of two data points is less than a threshold ε, they are connected. Any two nodes with a difference more than ε are not connected. Nodes that are not connected to their neighbors are treated as not connected to other nodes also. For *i* and *j* to be connected, the values of *i*, *j* and *k* ($i<k<j$, *k* denotes the points between *i* and *j*) should have approximately same value for all values of *k*.

The condition is given as,

$$\text{if } |p_i - p_j|, |p_i - p_k|, |p_k - p_j| < \varepsilon, \qquad A_{i,j} = 1;$$
$$\text{else}, \qquad A_{i,j} = 0;$$

The method is illustrated in Figure 1. The basic idea is that nodes of approximately the same value are connected together forming a group. Within any group, the nodes are different only by a value less than **ε.**

Edges in the obtained network do not have direction and weight. However, the method can be extended by assigning weights to edges based on the value of nodes connected by those edges. The edges connecting the nodes with higher values can be assigned higher weights and edges connecting the nodes with lower values can be assigned lower weights. The complex network with weights to edges is not constructed in the present work.

3. Application of the proposed method on model time series

3.1 Chaotic time series of Henon map

The proposed algorithm is applied to 10,000 observations of the classical Henon map ($x_{n+1} = 1 - ax_n{}^2 + bx_n$) (Al-Shameri 2012) at chaotic state (a=1.4, b=0.3). The first 1000 observations are discarded to remove the transients in the iterations. The number of edges that are connected with every node is specified as the degree of that node k. The probability of nodes having k number of connections is represented using $P(k)$. The degree distribution of the obtained network follows a power law behavior ($P(k) \sim k^{-\gamma}$), thereby indicating that the constructed network is “scale free”.

In addition to the scale free behavior, small world effects are captured by the present method. Small world behavior implies that the network has a high clustering coefficient and small characteristic path length (Watts & Strogatz 1998; Watts 1999). The complex network obtained from the first 50 observations of the chaotic time series from the Henon map is illustrated in Fig. 1(c). In this figure, node 4 is connected to a large number of nodes in the networks. Meanwhile, there are other nodes with fewer connections and nodes with even no connection. The nodes that have the largest number of connections are important and are called hubs. The presence of such hubs (nodes that are highly connected) is captured with the present method.

3.2. Periodic time series

The method is then applied on a periodic time series ($x(t) = sin(t)$). There are 20 peaks in the time series. All the nodes (peaks) have the same value. All the nodes are connected to all other nodes. All the nodes have the same number of connections (number of connections is 19). This implies that the obtained network is ‘regular’ (Quenell 1994). The periodic time series and the degree distribution of the corresponding complex network are shown in fig. 3 (a) & (b) respectively.

3.3. Random time series

A random noisy series is generated using the Matlab command called *randn*. The degree distribution of the obtained complex network follows a Poisson distribution, indicating that the corresponding network is a random network (Zhang & Small 2006). Random time series and degree distribution of the random network are shown in fig. 4 (a) & (b) respectively.

The present method is able to convert a periodic time series into a regular network, a random time series into a random network and a chaotic time series into a scale-free network.

In addition, the present method is successful in highlighting specific patterns in the time series in the corresponding complex network. We illustrate this by converting a time series obtained from an experiment (fluctuating pressure time series from a combustor involving turbulent flow) into complex network (Nair & Sujith 2014).

4. Application of proposed algorithm to a time series that exhibits intermittency

The present method is applied to the times series data of fluctuating pressure measured from a combustor with turbulent flow that exhibits intermittency. The obtained network is visualized using the software called Gephi (Bastian *et al.* 2009).

The complex networks derived from three different methods namely, visibility graph, horizontal visibility graph and threshold grouping method are shown (fig. 6 (a), (b) & (c) respectively) for comparison. The nodes are appearing as small circles filled with colors in the obtained networks. The edges are shown as the lines connecting the nodes. The nodes are colored based on the degree of that node. For example, green color in the visibility and horizontal visibility graph implies that the corresponding node has a degree of 2.

The intermittent time series shows periodic bursts. In the visibility and horizontal visibility graph, the nodes in the periodic bursts are connected only with nodes which are next to it and have a degree of 2. The number of nodes with green color is more compared to nodes with other colors in the visibility and horizontal visibility graphs. This is due to the presence of periodic bursts in the time series.

The threshold grouping method converts the intermittent bursts in the experimental time series into different clusters in the complex network. The clusters of various sizes can be seen in the network corresponding to the bursts in the time series. The proposed algorithm is thus able to capture the structure (pattern) that corresponds to the intermittent bursts present in the time series. Further, the properties of the complex network (for example clustering coefficient, short path length etc.) can be calculated. The variation of these network properties for different dynamical states (for example chaos, limit cycle oscillations etc.) is useful in studying the dynamical transitions. This will be presented in the future publications.

5. Conclusion

The threshold grouping algorithm converts the time series into complex networks. The structure of the dynamical features in the time series is preserved in the obtained complex network. The method is able to capture the network specific properties such as scale freeness, small world effects and the presence of hubs. The threshold grouping method is useful to visualize the intermittent bursts in the experimental time series data in the constructed complex network.

Acknowledgements

The authors would be thankful to Dr. V. Vineeth Nair for providing the data used in the present work. The authors would like to thank Dr. V. Vineeth Nair and Mr. E. A. Gopalakrishnan for their valuable suggestions.

References

Albert, R., Jeong, H., & Barabasi, A. L. (1999), "Internet: Diameter of the world-wide web", *Nature* **401**(6749), 130-131.

Albert, R., Jeong, H., & Barabasi, A. L. (2000), "Error and attack tolerance of complex networks", *Nature* **406**(6794), 378-382.

Alm, E., & Arkin, A. P. (2003), "Biological networks", *Current opinion in structural biology* **13**(2), 193-202.

Al-Shameri, W. F. H. (2012), "Dynamical properties of the Hénon mapping", *International Journal of Mathematical Analysis* **6**, 49.

Arianos, S., Bompard, E., Carbone, A., & Xue, F. (2009), "Power grid vulnerability: A complex network approach", *Chaos: An Interdisciplinary Journal of Nonlinear Science* **19**(1), 013119.

Bastian, M., Heymann, S., & Jacomy, M. (2009), "Gephi: an open source software for exploring and manipulating networks", *ICWSM* **8**, 361-362.

Barabasi, A. L., & Oltvai, Z. N. (2004), "Network biology: understanding the cell's functional organization", *Nature Reviews Genetics* **5**(2), 101-113.

Chen, G., Dong, Z. Y., Hill, D. J., Zhang, G. H., & Hua, K. Q. (2010), "Attack structural vulnerability of power grids: A hybrid approach based on complex networks", *Physica A: Statistical Mechanics and its Applications* **389**(3), 595-603.

Donner, R. V., Zou, Y., Donges, J. F., Marwan, N., & Kurths, J. (2010), "Recurrence networks - A novel paradigm for nonlinear time series analysis", *New Journal of Physics* **12**(3), 033025.

Eser, J., Zheng, P., & Triesch, J. (2014), "Nonlinear Dynamics Analysis of a Self-Organizing Recurrent Neural Network: Chaos Waning", *PloS one* **9**(1), e86962.

Kuramoto, Y. (2003), *Chemical oscillations, waves, and turbulence*, Courier Dover Publications.

Lacasa, L., Luque, B., Ballesteros, F., Luque, J., & Nuño, J. C. (2008), "From time series to complex networks: The visibility graph", *Proceedings of the National Academy of Sciences* **105**(13), 4972-4975.

Luque, B., Lacasa, L., Ballesteros, F., & Luque, J. (2009) "Horizontal visibility graphs: Exact results for random time series", *Physical Review E* **80**(4), 046103.

Nair, V., & Sujith, R. I. (2014), "Multifractality in combustion noise: predicting an impending combustion instability", *Journal of Fluid Mechanics* **747**, 635-655.

Nair, V., & Sujith, R. I. (2013), "Identifying homoclinic orbits in the dynamics of intermittent signals through recurrence quantification", *Chaos: An Interdisciplinary Journal of Nonlinear Science* **23**(3), 033136.

Pagani, G. A., & Aiello, M. (2013), "The power grid as a complex network: a survey", *Physica A: Statistical Mechanics and its Applications* **392**(11), 2688-2700, Elsevier.

Quenell, G. (1994), "Spectral Diameter Estimates for *k* Regular Graphs", *Advances in mathematics* **106**(1), 122-148.

Strogatz, S. H., & Stewart, I. (1993), "Coupled oscillators and biological synchronization", *Scientific American* **269**(6), 102-109.

Tan, C. O., Cohen, M. A., Eckberg, D. L., & Taylor, J. A. (2009), "Fractal properties of human heart period variability: physiological and methodological implications", *The Journal of physiology* **587**(15), 3929-3941.

Watts, D. J., & Strogatz, S. H. (1998), "Collective dynamics of 'small-world' networks", *Nature* **393**(6684), 440-442.

Watts, D. J. (1999), *Small worlds: the dynamics of networks between order and randomness*, Princeton university press.

Winfree, A. T. (2001), *The geometry of biological time*, Vol. **12**, Springer.

Zhang, J., & Small, M. (2006) "Complex network from pseudoperiodic time series: Topology versus dynamics", *Physical Review Letters* **96**(23), 238701.

Zhang, J., Sun, J., Luo, X., Zhang, K., Nakamura, T., & Small, M. (2008), "Characterizing pseudoperiodic time series through the complex network approach", *Physica D: Nonlinear Phenomena* **237**(22), 2856-2865.

Zorick, T., & Mandelkern, M. A. (2013), "Multifractal Detrended Fluctuation Analysis of Human EEG: Preliminary Investigation and Comparison with the Wavelet Transform Modulus Maxima Technique", *PloS one* **8**(7), e68360.

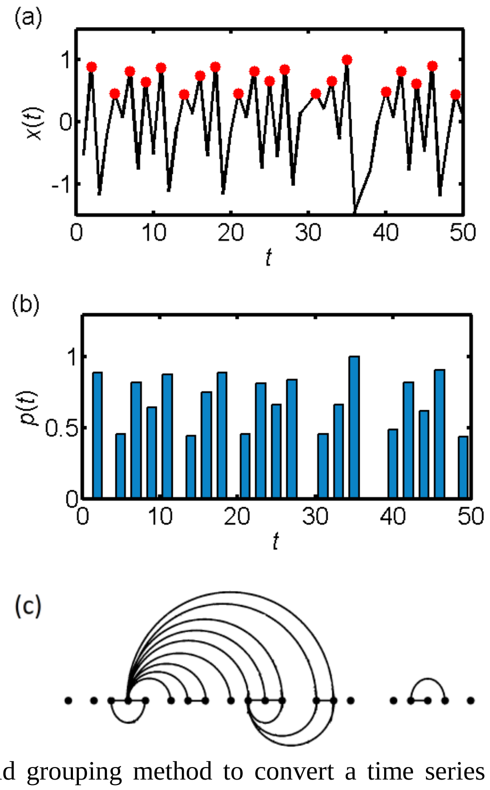

Figure 1. Illustration of threshold grouping method to convert a time series into a complex network. (a) $x(t)$ represents the time series to be converted into a complex network (For illustration, we show the first 50 observations of chaotic time series from the Henon map ($x_{n+1} = 1 - ax_n{}^2 + bx_n$) (Al-Shameri 2012) at chaotic state (a=1.4, b=0.3)), (b) The peaks in $x(t)$ are shown as $p(t)$, (c) Each peak is considered as a node. Two nodes (i and j) are connected if p_i, p_j and p_k ($i<k<j$, k denotes the points between i and j) have approximately the same value for all values of k. Here, $\varepsilon = 0.25$.

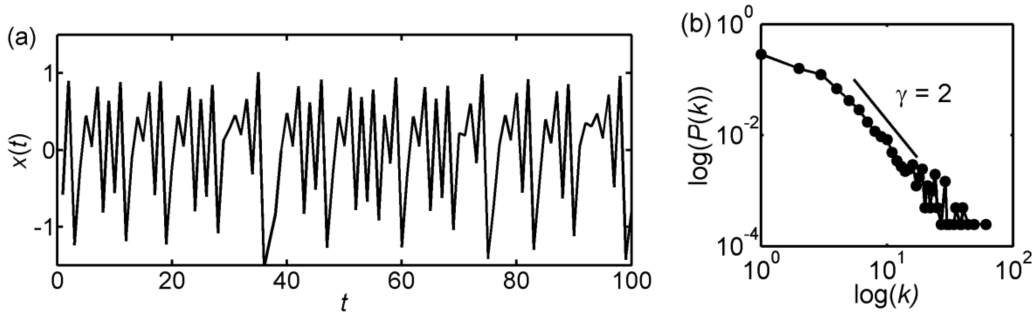

Figure 2. Chaotic time series obtained from the Henon map is converted into a scale-free network using the threshold grouping method. (a) The first 100 observations of the Henon map (Al-Shameri 2012) ($x_{n+1} = 1 - a{x_n}^2 + bx_n$) at chaotic state ($a = 1.4, b = 0.3$) after removing the transients, (b) degree distribution of the complex network obtained from the Henon map with $\varepsilon = 0.25$.

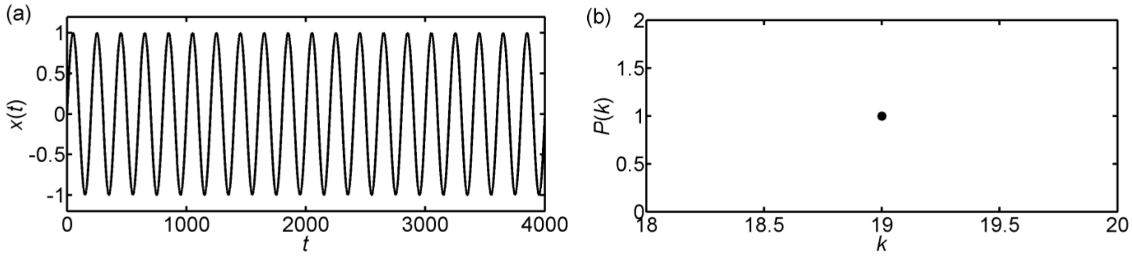

Figure 3. A regular network is obtained from a periodic time series using the threshold grouping method (a) Periodic time series ($x(t) = sin(t)$) with 4000 data points, there are 20 peaks in the time series and are converted into nodes, (b) degree distribution of the corresponding complex network. All the nodes have equal number of connections (number of connections is 19).

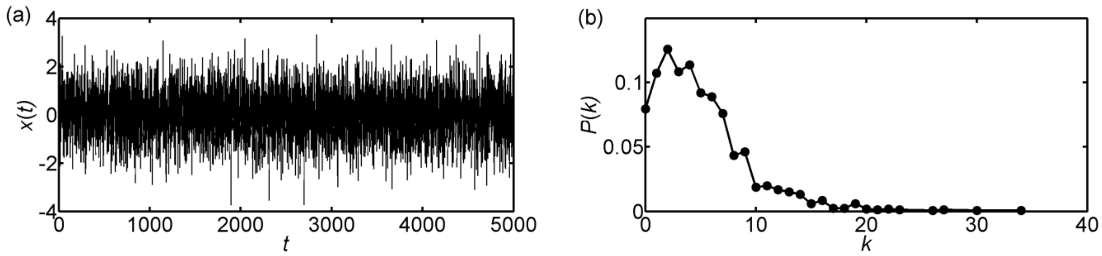

Figure 4. Random time series is converted into a random network (a) Random time series generated using the Matlab command called randn with 5000 data points, (b) degree distribution of the corresponding complex network.

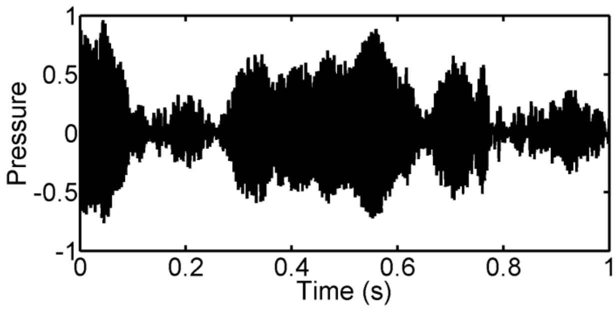

Figure 5. Time series data of the fluctuating pressure is measured from a combustor involving turbulent flow. The combustor is 700 *mm* long with square cross section (90 *mm* × 90 *mm*). Fuel (LPG) and oxidizer (air) are flowing at high Reynolds number (Re = 25812). The flame is stabilized using a flame holding device called a circular bluff body. Further details of the experiments can be found in (Nair & Sujith 2014). Intermittent periodic

bursts are observed in the time series data in a near random manner.

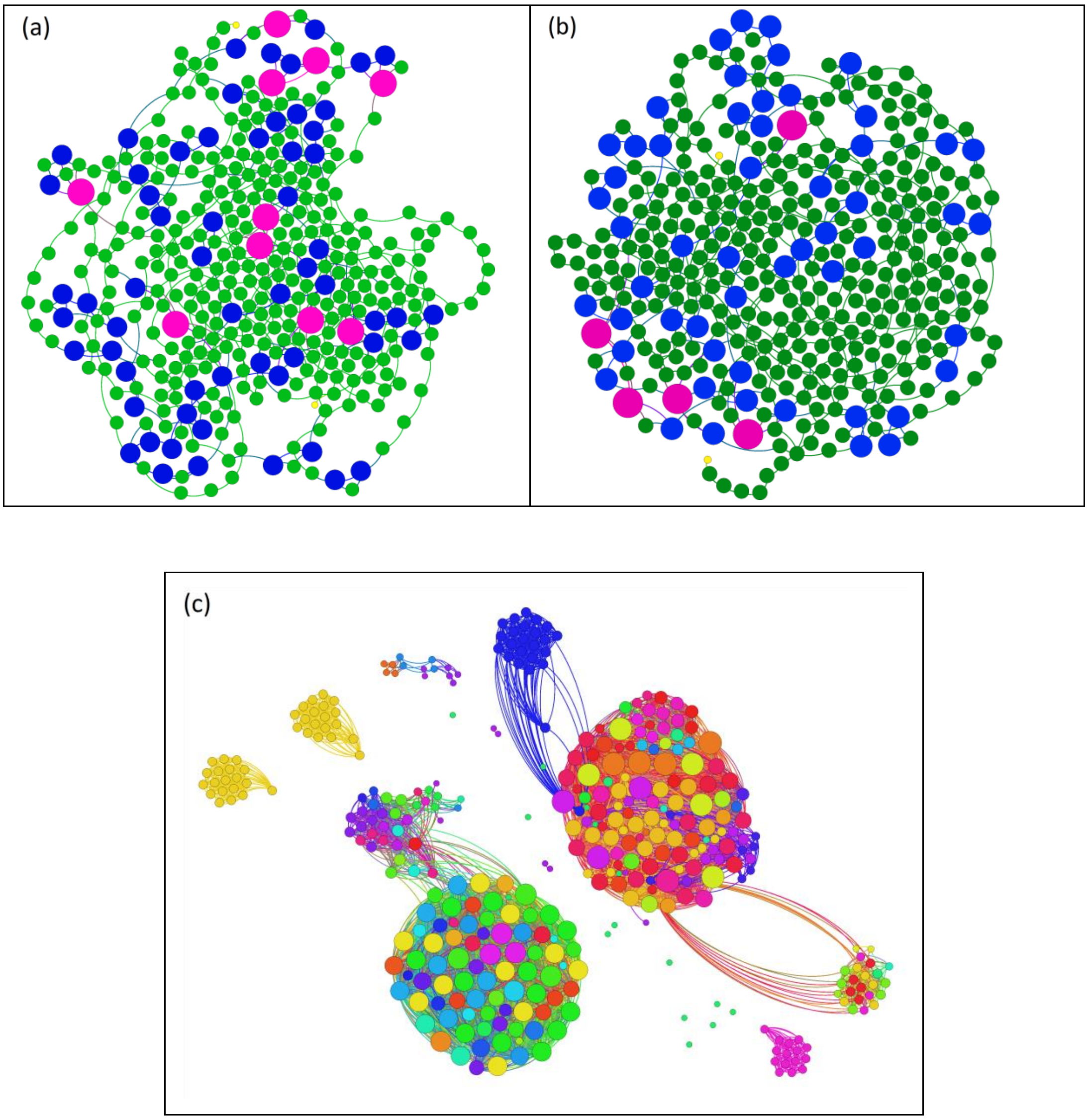

Figure 6. The complex networks are derived from a time series shown in Fig. 5 using (a) visibility condition, (b) horizontal visibility condition & (c) Threshold grouping method. The visualization of the complex networks are accomplished using the software called as Gephi (Basian *et al.* 2009).