Time Series Analysis based on Visibility Graph Theory

Shijie Yan

School of Automation and Electrical Engineering
University of Science and Technology Beijing,
Beijing, China
yanshijie0303@sina.com

Danling Wang*

School of Mathematics and Physics
University of Science and Technology Beijing,
Beijing, China
wang_dan_ling@sina.com

Abstract—**The dynamics of a complex system is usually recorded in the form of time series. In recent years, the visibility graph algorithm and the horizontal visibility graph algorithm have been recently introduced as the mapping between time series and complex networks. Transforming time series into the graphs, the algorithms allows applying the methods of graph theoretical tools for characterizing time series, opening the possibility of building fruitful connections between time series analysis, nonlinear dynamics, and graph theory. This paper analyzes mainly the topological properties of visibility graphs and horizontal visibility graphs associated to random series and fractional Brownian motions series, with a special emphasis on degree distribution of the associated graphs. As an example, this paper study the visibility graph and horizontal visibility graph constructed from the total daily turnover of stock market, and unveil that the degree distribution of visibility graph has power-law tails, and the degree distribution of horizontal visibility graph has exponential tail.**

Keywords-Time series analysis; Visibility graph; Horizontal Visibility Graph; Degree distribution; Daily turnover of stock.

I. INTRODUCTION

Random graph theory is considered as basis of research on the topological structure of complex networks, however, most of real complex networks are not random graph. The structure of complex networks is irregularly, complex and always changing with time, thus when the research objects are changed from simple and small networks to complex networks with thousands of vertices and edges, researchers begin to focus on the integrally dynamical property in complex networks. In 1998, Watts and Strogatz [1] proposed the famous model of small world networks. In 1999, Barabási and Albert [2] build evolutionary model of scale free networks. These two great works bring the research of complex networks into a new age. Since the end of 20th century, the researches and applications of complex networks have already spread to many fields, such as mathematics, biology, information sciences. The understanding on the features of complex networks in quantities and qualities has become a hot and challenging issue at present [3-5].

In the research of complex networks, one of the most critical aims is designing an intuitional, simple and effective method to build the model of complex networks for certain research objects. Since 2000, utilizing the analysis of topological structure to understand the dynamical behavior of time series via mapping them to complex networks has attracted researcher's great attention [6-10]. Zhang and Small [11,12] proposed an approach that map a time series to a cycle network. In this method, each vertex in network is a fragment in the time series. Edges between different vertices are determined by the similarity values or space distance of fragments. In 2008, Lacasa et al. [20] proposed the model of Visibility Graph (VG). In visibility graph, each vertex is each value in time series, edges between different vertices would exist if the two values can be "seen" with each other. The visibility graph is suitable for all kind of time series and the networks retain the structure information of time series. Particularly, periodic sequence is mapped to regular graph, random sequence is mapped to random graph and fractal sequence is mapped to scale free network. Meanwhile, in lacasa's work there is a linear relationship between the degree distribution exponent of visibility graph vertices and Hurst exponent of fractional Brownian motion which generate the visibility graph [21]. For some stochastic process, the fractal property has tight relationship [22-23]. The visibility graph theory has been applied in many fields, such as the analysis of Chinese macro economy time series in seasons [24], the test of wide speed in central Argentina [25] and the dynamics of human's heart ect [26]. Although the visibility graph theory is based on the single variable time series and in this sense the research on the degree distribution cannot provide us more information, however, it is still useful for numerical analysis of time series.

As the research on visibility graph going deeper and deeper, Luque et al. [27] proposed Horizontal Visibility Graph (HVG) model. HVG is an improvement of VG, for the same time series, its HVG is the subgraph of its VG. Although the degree distributions of HVGs generated from chaos time series and correlated time series are all exponential distribution. However, the distribution values are different. When the value is less than $\ln(1.5)$, the HVG describe the chaos series, while when the value is larger than $\ln(1.5)$, the HVG describe the correlated random series, specially, when the distribution exponent is equal to $\ln(1.5)$, the HVG describe random series with independent identically distributed variable. All the topological properties of HVG generated from fractional Brownian motion are depended on Hurst exponent: As the increasing of Hurst exponent, the absolute value of the degree distribution exponent of HVG is increasing, meanwhile, the clustering

IEEE
computer
society

coefficient is decreasing; the average length of path is increasing exponentially.

VG and HVG algorithms are efficient tools to construct complex networks and they are bridges between complex networks theory and analysis of time series. In this paper we applied the two algorithms on financial time series data and make a comparison on the degree distribution of VG and HVG which are generated from the same series data.

II. METHOD

A. Visibility Graph

In 2008, Lacasa et. al. proposed a method to construct visibility graph. In VG, the networks inherit some properties of time series. It is a brand new view on analysis of time series, and VG is a bridge between time series and complex networks and enriches the qualified analysis method form time series.

In general, the VG generate from a time series obey the following "rules of visibility": For any two value in time series (t_a, y_a) and (t_b, y_b) who will have visibility, and consequently will become two connected nodes of the associated graph, if any other data (t_c, y_c) placed between them fulfills:

$$y_c < y_b + (y_a - y_b)\frac{t_b - t_c}{t_b - t_a}$$

The VG algorithm map a time series $\{x_i\}_{i=1,...,N}$ to a complex network which can express by a adjacent matrix $A = (a_{ij})_{N \times N}$ if x_i and x_j in time series are satisfied the rule above, a_{ij}=1, otherwise a_{ij}=0. In Fig.1 the part above is a bar plot of values of a periodical time series and the part below is its VG. In this VG, every node corresponds, in the same order of the original time series and two nodes are connected if visibility existed between the two values in time series which means if there is a straight line that connects the value, and it is not intersect any intermediate data value


0.87,0.49,0.36,0.83,0.87,0.49,0.36,0.83,0.87,0.49,0.36,0.83,0.87,0.49,0.36,0.83,0.87,0.49,0.36,0.83...
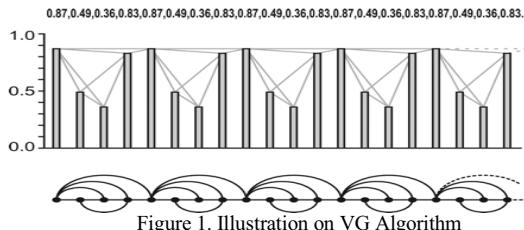
Figure 1. Illustration on VG Algorithm

Thus, following the algorithm above, a VG that is generated from a time series have the properties below:

(1) Connected: the VG is a full connected graph because there is one link between each node and its nearest neighbors
(2) Undirected: there is no direction defined in the links thus the VG is a undirected graph.
(3) Invariant: The VG structure is invariant under affine transformations of the series data.

The VG that is generated from factional Brownian motion is not only a scale free network, but also a small world network. Its degree distribution follows the power law: $P(k) \sim k^{-\gamma}$. If the series has a peak value $B_H(t) = h$ at time t, then the probability of the peak value has degree k can be denoted as:

$$P(k) \sim P_{fr}(k)r(k) \qquad (2)$$

where $P_{fr}(k)$ is the probability that series has the same peak value after k times step; $r(k)$ is the percentage between number of the vertices that time t can see and the total number of vertices in VG from time t to time $t + k$ in k times step. From factional Brownian motion series, $P_{fr}(k) \sim k^{H-2}$ and $r(k) \sim k^{H-1}$, then formula (2) can be write as

$$P(k) \sim k^{H-2}k^{H-1} = k^{2H-3} \qquad (3)$$

Then, there is a linear relationship between the exponent of the degree distribution in VG and the Hurst exponent of the accordingly fractional Brownian motion series.

$$\gamma(H) = 3 - 2H \qquad (4)$$

The tail of the degree distribution of VG generated from random time series follows a power law: $P(k) \sim e^{-\lambda k}$. At the beginning of the curve, the distribution is like Poisson distribution while the tail obeys exponential distribution. Actually, let k_t denote as the connection between value t and other vertices, if the value of k_t was large, then k_t+1 would be very small, that's because the time series is random. Thus at the tail of degree distribution, it become to obey exponential distribution.

B. Horizontal Visibility Graph

As the research on the algorithm for VG construction going deeper, an algorithm for horizontal visibility graph construction is proposed by Luque et al. in 2009 which is a simpler and more effective method. HVG is much simpler and more precise on analysis of the topological properties of networks. Meanwhile, HVG can describe the statistical properties of random series more effectively.

Let $\{x_i\}$,i=1,...,N be a time series o f N data. The HVG algorithm assigns each value of the series to a node. The two nodes i and j in the graph are connected if one can draw a horizontal line in the time series joining x_i and x_j that does not intersect any intermediated data height. Thus, the two connected nodes i and j should satisfied with the following rule:

$$x_i, x_j > x_n, \forall n \,|\, i < n < j \qquad (5)$$

Accordingly, $\{x_i\}$,i=1,...,N is transformed into HVG G, and $A = (a_{ij})_{N \times N}$ is the adjacency matrix of G, where a_{ij} =1 if the value x_i and x_j can be connected according the horizontal visible rule above, and a_{ij} =0 if not.

The HVG algorithm is an improvement and simplification of the VG algorithm. Actually, the HVG is always a subgraph of the VG which are generated from the same time

312

series. Accordingly, the HVG associated with a time series always has the following features:

(1) Connected: the HVG is a full connected graph because there is one link between each node and its nearest neighbors;

(2) Invariant: The VG structure is invariant under affine transformations of the series data.

(3) Undirected and directed characters of the mapping: the algorithm generates undirected graphs, however we could also get a directed graph in such a way that for a given node which have the ingoing degree and outgoing degree;

(4) Reversible and irreversible characters of the mapping: some information of the time series is inevitably lost when the network is constructed. For example, the two periodic series with the same period as T1=…,2,3,2,3,… and T2=…,4,5,4,5,…would have the same visibility graph and the same adjacency matrix. However T1 and T2 still could be distinguished from each other quantitatively via constructing weighted HVGs.
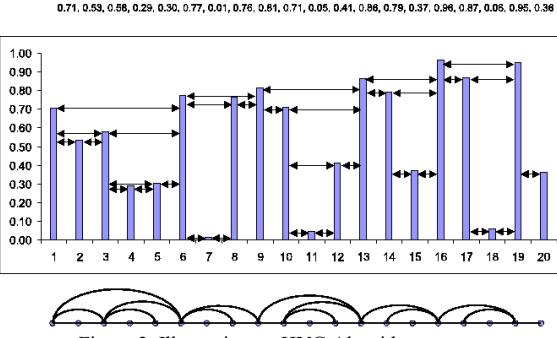

Figure 2. Illustration on HVG Algorithm.
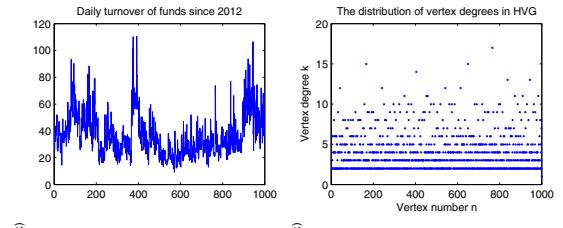
As the increasing of the Hurst exponent H, the absolute value of the slope λ of HVG generated from the fractional brown motion series increases and the clustering coefficient decreases, the average path distance L increase exponentially. The HVG generated from any random series is a small world networks. And its degree distribution obeys the unified exponential form:

$$P(k) = (\frac{1}{3})(\frac{2}{3})^{k-2} \qquad (6)$$

which means the degree distribution have no relationship with probability density function.

III. RESULTS AND DISCUSSION
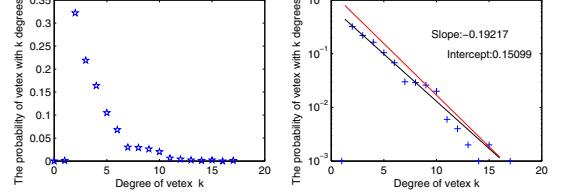
In this section we construct VG based on the turnovers per day time series of the Chinese stock market since 2012 to 2014 and describe their degree distribution curve in figure 3, figure 4. For the same data, we also construct HVG by matlab and give the semi-log plot and log-log plot of its degree distribution in figure 5 and figure 6. For both of the two degree distribution semi-log plot and log-log plot, the linear least square fitting method is used to obtain the slope .


Figure 3. Degree Distribution of VG
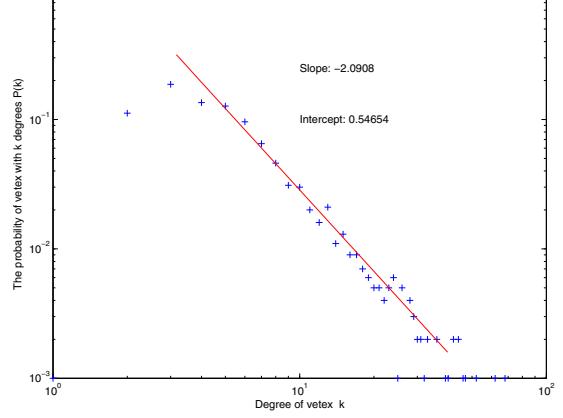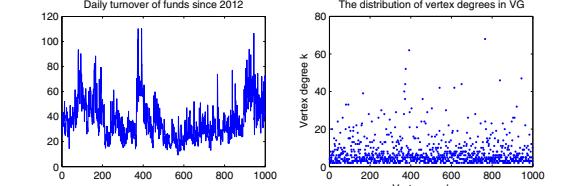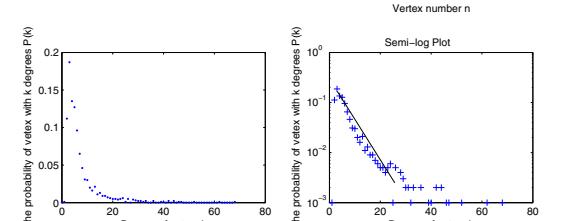
Figure 4. Linear Fitting on Degree Distribution of VG

Figure 5. Degree Distribution of HVG

In figure 3, the degree distribution of VG is center on $P(k) = 0$ around. The nodes with degree probability between 0.05 and 0.2 are too few which leads most of the nodes in the figure are located in the left bottom. Thus the local degree distribution cannot describe the statistical of the series. However in figure 4, the degree distribution of HVG is more uniform, the points in the figure are proximal in one straight line. $\log(P(k)) = -0.19217k + 0.15099$

313
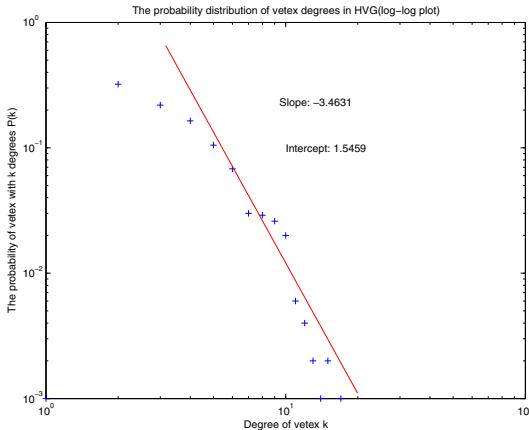
Figure 6. Linear Fitting on Degree Distribution of HVG

Because of $\left|\dfrac{-0.17609-(-0.19217)}{-0.17609}\right|<10\%$, thus the degree distribution of VG of the stock turnover time series obey the power low and its probability distribution function is satisfied with (6).

In figure 5, there are seldom available points and in the tail of the fitting straight line of the degree distribution, the points are not located uniformly, thus it is not clear that whether the degree distribution tail of the HVG follows the power law. In figure 6, there are more available points and they are uniformly located around the fitting straight line. Thus the degree distribution of HVG generated from the stock turnover since 2012 obey the power law $P(k)\sim k^{-\gamma}$ and the HVG is a scale free network.

IV. CONCLUSION

The algorithms of VG and HVG are applied for the analysis of time series. The statistical properties can be obtained from the topological properties of the networks. In this paper we applied the two algorithms on the analysis of Chinese stock turnover time series since 2012. The degree distribution of VG followed the power law $P(k)\sim k^{-\gamma}$, its VG is a scale free network, while the degree distribution of HVG of the time series data is satisfied with the exponential distribution, the probability distribution function is approximately equal to $P(k)=(1/3)(2/3)^{k-2}$.

REFERENCES.

[1] Watts D J,Strogatz S H. Collective dynamics of 'small-world' networks[J]. Nature, 1998, 393(6684):440-442.

[2] Barabási A L,Albert R. Emergence of scaling in random networks[J]. Science,1999,286(5439):509-512.

[3] Lv J H. Mathematical model of complex dynamic network and synchronization criterion. System Engineering Theory and Practice, 2004,4(6): 17-22, in Chinese

[4] Wang X.F., Li X., Chen G.Y. Theory and application on complex networks, Tsinghua University Publisher, 2006, Beijing, in Chinese

[5] Fang J.Q., Wang X.F, Zheng Z.G., A brand new interdiscipline: Networks Science. Advance in Physics, 27(3),2007, 239-343, in Chinese.

[6] Fang J.Q., Wang X.F. and Liu Z. R. A brief introduction on complicated problems and study on non-linear complex networks systems, Science and Technology Review, 2,2004, 9-12. in Chinese.

[7] Wang B. H., Zhou T., Wang W. X. The current research content on complex system. Complex Systems and Complexity Science, 2008,5(4),21-28,in Chinese

[8] Caraiani P. Characterizing emerging European stock markets through complex networks: From local properties to self-similar characteristics. Physica A,2012,391(13): 3629-3637.

[9] Ahmadlou M,Adeli H,Lichtenstein A G,et al. Improved visibility graph fractality with application for diagnosis of Autism Spectrum Disorder. Physica A,2012,391(20):4720-4726.

[10] Zhang J,Small M. Complex network from pseudoperiodic time series:Topology versus dynamics[J]. Phys. Rev. Lett.,2006,96(23):238701.

[11] Xu X,Zhang J,Small M. Superfamily phenomena and motifs of networks induced from time series. Proc. Natl. Acad. Sci. USA,2008,105(50):19601-19605.

[12] Donner R V,Zou Y,Donges J F,et al. Recurrence networks: A novel paradigm for nonlinear time series analysis. New J. Phys.,2010,12(3):033025.

[13] Gao Z K,Jin N D. A directed weighted complex network for characterizing chaotic dynamics from time series. Nonlinear Analysis,2012,13(2):947-952.

[14] Dong Y,Huang W,Liu Z,et al. Network analysis of time series under the constraint of fixed nearest neighbors. Physica A,2012,392(4):967-973.

[15] Marwan N,Wessel N,Meyerfeldt U,et al. Recurrence plot based measures of complexity and its application to heart rate variability data. Phys. Rev. E,2002,66(2):026702.

[16] Donner R V,Small M,Donges J F,et al. Recurrence-based time series analysis by means of complex network methods. International Journal of Bifurcation and Chaos,2011,21(4):1019-1046.

[17] Yang Y,Yang H. Complex network-based time series analysis. Physica A,2008,387(5):1381-1386.

[18] Thiel M. etc. Spurious structures in recurrence plots induced by embedding. Nonlinear Dynamics,2006,44(1-4):299-305.

[19] Lacasa L,Luque B,Ballesteros F,et al. From time series to complex networks: The visibility graph. Proc. Natl. Acad. Sci.,2008,105(13):4972-4975.

[20] Lacasa L,Luque B,Luque J,et al. The visibility graph: a new method for estimating the Hurst exponent of fractional Brownian motion. EPL(Europhys. Lett.),2009,86(3):30001.

[21] Ni X H,Jiang Z Q,Zhou W X. Degree distributions of the visibility graphs mapped from fractional Brownian motions and multifractal random walks. Phys. Lett. A, 2009, 373(42):3822-3826.

[22] Newman M E J. The structure and function of complex networks. SIMA review,2003,45(2):167-256.

[23] Boccaletti S,Latora V,Moreno Y,et al. Complex networks: Structure and dynamics. Physics reports,2006,424(4):175-308.

[24] Wang N,Li D,Wang Q. Visibility graph analysis on quarterly macroeconomic series of China based on complex network theory. Physica A,2012,391(24):6543-6555.

[25] Pierini J. O. etc. Visibility graph analysis of wind speed records measured in central Argentina. Physica A, 2012, 391(20): 5041-5048.

[26] Shao Z G. Network analysis of human heartbeat dynamics. Appl. Phys. Lett., 2010, 96(7): 073703.

[27] Luque B,Lacasa L,Ballesteros F,et al. Horizontal visibility graphs: Exact results for random time series. Phys. Rev. E,2009,80(4): 046103.