From time series to complex networks: The visibility graph

Lucas Lacasa*[†], Bartolo Luque*, Fernando Ballesteros[‡], Jordi Luque[§], and Juan Carlos Nuño[¶]

*Departmento Matemática Aplicada y Estadística, ETSI Aeronáuticos, Universidad Politécnica de Madrid, 28040 Madrid, Spain; [‡]Observatorio Astronómico, Universidad de Valencia, 46010 Valencia, Spain; [§]Departmento de Teoria del Senyal i Comunicacions, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain; and [¶]Departmento Matemática Aplicada a los Recursos Naturales, ETSI Montes, Universidad Politécnica de Madrid, 28040 Madrid, Spain

In this work we present a simple and fast computational method, the *visibility algorithm*, that converts a time series into a graph. The constructed graph inherits several properties of the series in its structure. Thereby, periodic series convert into regular graphs, and random series do so into random graphs. Moreover, fractal series convert into scale-free networks, enhancing the fact that power law degree distributions are related to fractality, something highly discussed recently. Some remarkable examples and analytical tools are outlined to test the method's reliability. Many different measures, recently developed in the complex network theory, could by means of this new approach characterize time series from a new point of view.

Brownian motion | complex systems | fractals

Fig. 1. Example of a time series (20 data values) and the associated graph derived from the visibility algorithm. In the graph, every node corresponds, in the same order, to series data. The visibility rays between the data define the links connecting nodes in the graph.

I n this article we present a tool in time series analysis: the *visibility graph*. This algorithm maps a time series into a network. The main idea is to study to which extent the techniques and focus of graph theory are useful as a way to characterize time series. As will be shown below, this network inherits several properties of the time series, and its study reveals nontrivial information about the series itself.

For illustrative purposes, in Fig. 1 we present a scheme of the visibility algorithm. In the upper zone we plot the first 20 values of a periodic series by using vertical bars (the data values are displayed above the plot). Considering this as a landscape, we link every bar (every point of the time series) with all those that can be seen from the top of the considered one (gray lines), obtaining the associated graph (shown in the lower part of the figure). In this graph, every node corresponds, in the same order, to series data, and two nodes are connected if visibility exists between the corresponding data, that is to say, if there is a straight line that connects the series data, provided that this "visibility line" does not intersect any intermediate data height.

More formally, we can establish the following visibility criteria: two arbitrary data values (t_a, y_a) and (t_b, y_b) 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:

$$y_c < y_b + (y_a - y_b) \frac{t_b - t_c}{t_b - t_a}. \qquad [1]$$

We can easily check that by means of the present algorithm, the associated graph extracted from a time series is always:

1. Connected: each node sees at least its nearest neighbors (left and right).
2. Undirected: the way the algorithm is built up, there is no direction defined in the links.
3. Invariant under affine transformations of the series data: the visibility criterion is invariant under rescaling of both horizontal and vertical axes, and under horizontal and vertical translations (see Fig. 2).
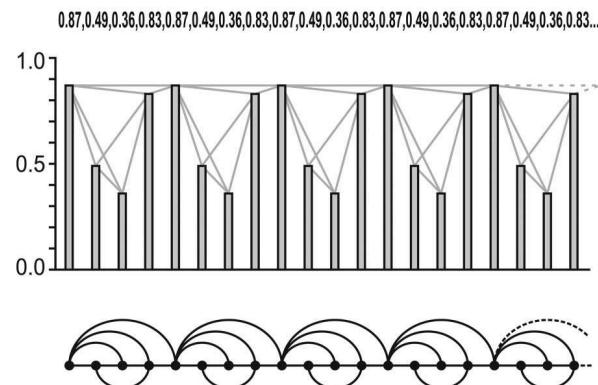
In a recent work (1), Zhang and Small (ZS) introduced another mapping between time series and complex networks. Although the philosophy is similar to this work (to encode the time series in a graph to characterize the series by using graph theory), fundamental differences exist between both methods, mainly in what refers to the range of applicability (ZS only focus on pseudoperiodic time series, associating each series cycle to a node and defining links between nodes by temporal correlation measures, whereas the visibility graph can be applied to every kind of time series) and the graph connectedness (in ZS the giant component is assured only ad hoc; meanwhile, the visibility graph is always connected by definition).

The key question is to know whether the associated graph inherits some structure of the time series, and consequently whether the process that generated the time series may be characterized by using graph theory. In a first step we will consider periodic series. As a matter of fact, the example plotted in Fig. 1 is nothing but a periodic series with period 4. The associated visibility graph is regular, as long as it is constructed by periodic repetition of a pattern. The degree distribution of this graph is formed by a finite number of peaks related to the series period, much in the vein of the Fourier power spectrum of a time series. Generically speaking, all periodic time series are mapped into regular graphs, the discrete degree distribution being the fingerprint of the time series periods. In the case of
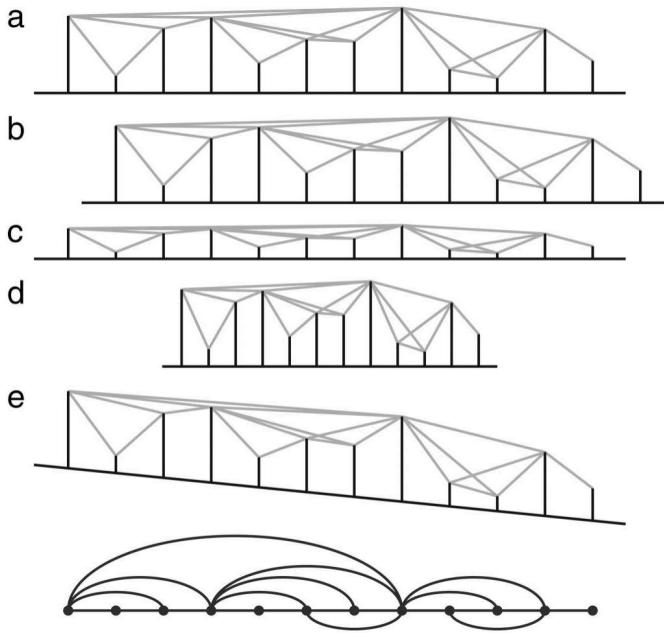
Fig. 2. The visibility graph of a time series remains invariant under several transformation of the time series. (*a*) Original time series with visibility links. (*b*) Translation of the data. (*c*) Vertical rescaling. (*d*) Horizontal rescaling. (*e*) Addition of a linear trend to the data. As can be seen in the bottom diagram, in all these cases the visibility graph remains invariant.

Fig. 3. Random series. (*Left*) First 250 values of $R(t)$, where R is a random series of 10^6 data values extracted from U[0,1]. (*Right*) Degree distribution $P(k)$ of the visibility graph associated with $R(t)$ (plotted in semilog). Although the beginning of the curve approaches the result of a Poisson process, the tail is clearly exponential. This behavior is due to data with large values (rare events), which are the hubs.

periodic time series, its regularity seems therefore to be conserved or inherited structurally in the graph by means of the visibility map.

As an opposite to periodic series, in a second step we will tackle a series $R(t)$ of 10^6 data values extracted from an uniform distribution in [0, 1]. Although one would expect in a first moment a Poisson degree distribution in this case [as for uncorrelated random graphs (2)], a random time series has indeed some correlation, because it is an ordered set. In fact, let k_t be the connectivity of the node associated with the data t. If k_t is large (related to the fact that the data have a large value and that consequently they have large visibility), one would expect that k_{t+1} would be relatively small, because the time series is random and two consecutive data values with a large value are not likely to occur. It is indeed because of these "unlikely" large values (the hubs) that the tail of the degree distribution deviates from the Poisson distribution. Two large values in the series data can be understood as two rare events in the random process. The time distribution of these events is indeed exponential (3), therefore we should expect the tail of the degree distribution in this case to be exponential instead of Poissonian, as long as the form of this tail is related to the hub's distribution.

In the left side of Fig. 3 we depict the first 250 values of $R(t)$. In the right side we plot the degree distribution $P(k)$ of its visibility graph. The tail of this distribution fits an exponential distribution quite well, as expected. Note, at this point, that time series extracted randomly from other distributions than uniform have also been addressed. In every case the algorithm captures the random nature of the series, and the particular shape of the degree distribution of the visibility graph is related to the particular random process.

Hitherto, ordered (periodic) series convert into regular graphs, and random series convert into exponential random graphs; order and disorder structure in the time series seem to be inherited in the topology of the visibility graph. Thus, the question arises: What kind of visibility graph is obtained from a fractal time series? This question is in itself interesting at the present time. Recently, the relationship between self-similar and scale-free networks (4–8) has been intensively discussed (9–12). Although complex networks (5) usually exhibit the small-world property (13) and consequently cannot be size-invariant, it has been recently shown (9) that by applying fitted box-covering methods and renormalization procedures, some real networks actually exhibit self-similarity. So, whereas self-similarity seems to imply scale freeness, the opposite is not true in general.

To explore these issues in more detail, the following two fractal series will be considered: the well known Brownian motion $B(t)$ and the Conway series (14). Whereas the Brownian motion represents a well known case of self-affinity [indeed, the following relation holds: $B(t) = a^{1/2}B(t/a)$], the Conway series $a(n) - n/2$ is the recursively generated fractal series from:

$$a(1) = a(2) = 1$$

$$a(n) = a(a(n-1)) + a(n - a(n-1)); n > 2. \qquad [2]$$

In Fig. 4 we have plotted the behavior of these series, the degree distribution $P(k)$ of their respective visibility graphs and their
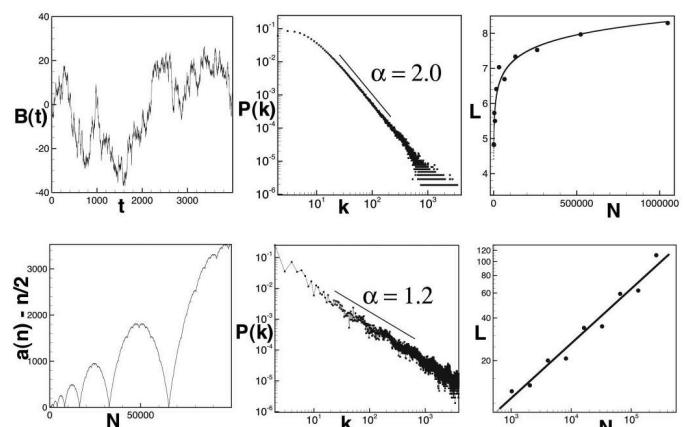
Fig. 4. Fractal series. (*Upper* from *left* to *right*) First 4,000 data values from a Brownian series of 10^6 data values. (*Center*) The degree distribution of the visibility graph associated with the Brownian motion. This one is a power law $P(k) \sim k^{-\alpha}$ with $\alpha = 2.00 \pm 0.01$. (*Right*) Plot of the mean path length of this network as a function of the network size N. The best fitting provides a logarithmic scaling $L(N) = 1.21 + 0.51 \log(N)$. This network shows small-world effect in addition to being scale-free. (*Lower* from *left* to *right*) First 10^5 data values from a Conway series of $4 \cdot 10^6$ data values. (*Center*) The degree distribution of the visibility graph associated with the Conway series. This one is a power law $P(k) \sim k^{-\alpha}$ with $\alpha = 1.2 \pm 0.1$. The mean path length as a function of the size N is depicted in *Right*. The best fitting provides a power law scaling $L(N) = 0.76N^{0.38}$. Then, this network is scale-invariant.
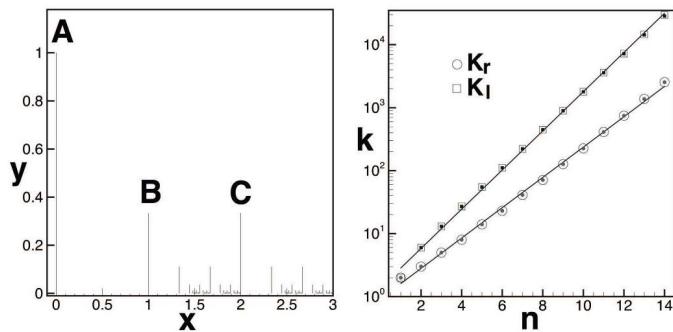
Fig. 5. Simple deterministic fractal. (*Left*) Fractal series obtained by iteration of the original pattern (points A, B, and C) with $p = 10$ steps. (*Right*) Values of K_r (circles) and K_l (squares) as a function of the fractal size, related to Eqs. **3** and **5**. Note that the plot is log-linear; the relation is thus exponential. The straight lines correspond to the approximations deduced in Eqs. **4** and **6**.

mean path length $L(N)$ as a function of the series length. First, both series have visibility graphs with degree distributions that correspond to power laws of the shape $P(k) \sim k^{-\alpha}$, where we get different exponents in each case: this result enhances the fact that, in the context of the visibility algorithm, power law degree distributions [that is, scale-free networks (6–9)] arise naturally from fractal series. Moreover, this relation seems to be robust as long as the preceding examples show different kinds of fractality: whereas $B(t)$ stands for a stochastic self-affine fractal, the Conway series is a deterministic series recursively generated. However, whereas the Brownian visibility graph seems to evidence the small-world effect (Fig. 4 *Upper Right*) as $L(N) \sim \log(N)$, the Conway series shows in turn a self-similar relation (Fig. 4 *Lower Right*) of the shape $L(N) \sim N^\beta$. This fact can be explained in terms of the so-called hub repulsion phenomenon (11): visibility graphs associated with stochastic fractals such as the Brownian motion or generic noise series do not evidence repulsion between hubs (in these series it is straightforward that the data with highest values would stand for the hubs, and these data would have visibility between each other), and consequently will not be fractal networks according to Song *et al.* (11). However, the Conway series actually evidence hub repulsion: this series is concave-shaped and, consequently, the highest data values will not in any case stand for the hubs; the latter ones would be located most likely in the monotonic regions of the series, which are indeed hidden from each other (effective repulsion) across the series local maxima. The Conway visibility graph is thus fractal.

Because a fractal series is characterized by its Hurst exponent, we may argue that the visibility graph can actually distinguish different types of fractality, something that will be explored in detail in further work. Note at this point that some other fractal series have been also studied [Q series (15), Stern series (16), and Thue–Morse series (17), etc.] with similar results. Moreover, observe that if the series under study increases its length, the resulting visibility graph can be interpreted in terms of a model of network growth and may eventually shed light into the fractal network formation problem.

To cast light into the relation between fractal series and power law distributions, in Fig. 5 *Left*, we present a deterministic fractal series generated by iteration of a simple pattern of three points. The series starts (step 0) with three points (A, B, and C) of coordinates (0, 1), (1, 1/3), and (2, 1/3), respectively. In step p, we introduce 2^{p+1} new points with height 3^{-p-1} and distanced 3^{-p}. The series form a self-similar set: applying an isotropic zoom of horizontal scale 3^p and vertical scale 3^p to the pattern of order p, we recover the original pattern.

Note that this time series is not data uniformly spaced as in the previous examples. However, its usefulness is set on the fact that it is simple enough to handle it analytically, that is, to find the degree distribution of its visibility graph. The main idea is to find a recurrence behavior in the way that a given node increases its connectivity when the fractal step (that is, the fractal size) is increased (18). Then we calculate how many nodes (self-similar to it) appear in each step, and from both relations we come to a degree distribution for these kinds of nodes.

First, from a quick visual exploration of Fig. 5 *Left* one comes to the conclusion that nodes A and B have typically the same degree. However, the degree of node C can be decomposed in two terms: the left degree (due to visibility of nodes at the left of C) and the right degree (due to visibility of nodes at the right of C). The degree of A and B is the same as the right degree of C (statistically speaking, A and B increase their connectivity as the fractal size increases much in the way as the right part of C). Thereby, the degree of C provides the whole information of the system. We will quote $K_r(C, n)$ the right degree of node C in a n-step fractal (respectively, $K_l(C, n)$ stands for the left degree). Applying the visibility criterion, one can geometrically find that

$$K_r(C, n) = \sum_{m=1}^{n} \frac{1}{m} \sum_{d|m} \mu(d) \cdot 2^{m/d}, \qquad [3]$$

where μ is the Moëbius function. Note that this summation agrees with the number of irreducible polynomials of degree at most n over the Galois field GF(2) (19), something which deserves an in-depth investigation. This expression can be approximated by

$$K_r(C, n) \sim 2^{4n/5}. \qquad [4]$$

However, there is a recurrence in the left degree that reads

$$K_l(C, n) = 2K_l(C, n-1) + 1, \qquad [5]$$

whose leading term is

$$K_l(C, n) \sim 2^n. \qquad [6]$$

The node C will thus have a degree $K(C) = K_r(C, n) + K_l(C, n)$. In Fig. 5 *Right* we plot the values of K_r (circles) and K_l (squares) as a function of the fractal size (the number of iterations n). Numerical values are plotted as the outer circles and squares, whereas the inner circles and squares come from plotting Eqs. **3** and **5**. Note that both formulas reproduce the numerical data. The straight lines correspond to the approximation Eqs. **4** and **6**. Now, in a generic step p, 2^p nodes that are self-similar to C appear (by construction). Those nodes will have a degree $K(C, n-p) = 2^{4/5(n-p)} + 2^{n-p}$ that, for large values of $n-p$, can be approximated to $K(C, n-p) \simeq 2^{n-p}$. Defining $f(K)$ as the degree distribution, we get that $f(K(C, n-p)) = 2^p$, and with the change of variable $u \equiv 2^{n-p}$, it is easy to come into:

$$f(u) \sim u^{-1}, \qquad [7]$$

that is, the degree distribution related to the C nodes is a power law. Although this simple example does not provide a general explanation of why fractality is traduced into power law distributions, it may stand as a generic way of dealing with deterministic fractal series generated by iteration.

Once the visibility method has been presented, some remarks can be stated: note that typically two series that only differ by an affine transformation will have the same visibility graph; in this sense the algorithm absorbs the affine transformation. Furthermore, it is straightforward to see that some information regarding the time series is inevitably lost in the mapping from the fact

that the network structure is completely determined in the (binary) adjacency matrix. For instance, two periodic series with the same period as $T_1 = \{\ldots, 3, 1, 3, 1, \ldots\}$ and $T_2 = \{\ldots, 3, 2, 3, 2, \ldots\}$ would have the same visibility graph, albeit being quantitatively different. Although the spirit of the visibility graph is to focus on time series structural properties (periodicity, fractality, etc.), the method can be trivially generalized by using weighted networks (where the adjacency matrix is not binary and the weights determine the slope of the visibility line between two data values), if we eventually need to quantitatively distinguish time series like T_1 and T_2, for instance.

Although in this article we have only tackled undirected graphs, note that one could also extract a directed graph (related to the temporal axis direction) in such a way that for a given node one should distinguish two different connectivities: an ingoing degree k_{in}, related to how many nodes see a given node i, and an outgoing degree k_{out}, that is the number nodes that node i sees. In that situation, if the direct visibility graph extracted from a given time series is not invariant under time reversal [that is, if $P(k_{in}) \neq P(k_{out})$], one could assert that the process that generated the series is not conservative. In a first approximation we have studied the undirected version and the directed one will be eventually addressed in further work.

There are some direct applications of the method that can be put forward. The relation between the exponent of the degree distributions and the Hurst exponent of the series will be addressed in further work. In particular, it turns out that the method presented here constitutes a reliable tool to estimate Hurst exponents, as far as a functional relation between the Hurst exponent of a fractal series and the degree distribution of its visibility graph holds (J.C.N., B.L., L.L., and F.B., unpublished work). Note that the estimation of Hurst exponents is an issue of major importance in data analysis that is yet to be completely solved (see, for instance, ref. 20). Fractional Brownian motions, a concept of great interest in a large variety of fields ranging from electronic devices to biology, will also be considered in relation with the preceding point.

Moreover, the ability of the algorithm to detect not only the difference between random and chaotic series but also the spatial location of inverse bifurcations in chaotic dynamical systems is another fundamental issue that will also be at the core of further investigations (unpublished work). Finally, the visibility graph characterizes nontrivial time series and, in that sense, the method may be relevant in specific problems of different garments, such as human behavior time series recently put forward (21).

In summary, a algorithm that converts time series into graphs is presented. The structure of the time series is conserved in the graph topology: periodic series convert into regular graphs, random series into random graphs, and fractal series into scale-free graphs. Such characterization goes beyond the preceding points, since different graph topologies arise from apparently similar fractal series. In fact, the method captures the hub repulsion phenomenon associated with fractal networks (11) and thus distinguishes scale-free visibility graphs evidencing the small-world effect from those showing scale invariance. With the visibility algorithm, a natural bridge between complex networks theory and time series analysis has now been built.

1. Zhang J, Small M (2006) *Phys Rev Lett* 96:238701.
2. Bollobás B (1998) *Modern Graph Theory* (Springer, New York).
3. Feller W (1971) *An Introduction to Probability Theory and Its Applications* (Wiley, New York).
4. Barabási AL, Albert R (1999) *Science* 286:509.
5. Albert R, Barabási AL (2002) *Rev Mod Phys* 74:47–92.
6. Newman MEJ (2003) *SIAM Rev* 45:167–256.
7. Dorogovtsev S, Mendes JFF (2002) *Adv Phys* 51:4.
8. Bocaletti S, Latora V, Moreno Y, Chávez M, Hwang DU (2006) *Phys Rep* 424:175–308.
9. Song C, Havlin S, Makse HA (2005) *Nature* 433:392.
10. Goh KI, Salvi G, Kahng B, Kim D (2006) *Phys Rev Lett* 96:018701.
11. Song C, Havlin S, Makse HA (2006) *Nat Phys* 2:275.
12. Kim JS, Goh KI, Salvi G, Oh E, Kahng B, Kim D (2007) *Phys Rev E* 75:016110.
13. Watts DJ, Strogatz SH (1998) *Nature* 393:440–442.
14. Conway J (1988) *Some Crazy Sequences*, Lecture presented at AT&T Bell Labs, Murray Hill, NJ, 15 July.
15. Hofstadter D (1980) *Gödel, Escher, Bach* (Vintage Books, New York), pp 137–138.
16. Stern MA (1858) *J Reine Angew Math* 55:193–220.
17. Schroeder MR (1991) *Fractals, Chaos, and Power Laws* (Freeman, New York).
18. Barabási A-L, Ravasz E, Vicsek T (2001) *Physica A* 299:559564.
19. Hicks KH, Mullen GL, Sato I (2002) Distribution of irreducible polynomials over F2. *Finite Fields with Applications to Coding Theory, Cryptography and Related Areas* (Oaxaca, 2001) (Springer, Berlin), pp 177–186.
20. Karagiannis T, Molle M, Faloutsos M (2004) *IEEE Internet Comput* 8:5.
21. Vázquez A, Gama Oliveira J, Deszö Z, Goh K, Kondor I, Barabási AL (2006) *Phys Rev E* 73:036127.

APPLIED MATHEMATICS