Detecting the phenomenon of strange non-chaotic attractors

Alan Oxley[1]

(Received 16 December 2009; revised 10 August 2010)

Abstract

I comment on how chaos might be defined. A sample of dynamical systems that have quasi-periodic forcing functions is then considered. The normal approach found in the literature is to start with an ordinary differential equation, change to a difference equation, and then plot a graph. The question of how to detect a strange non-chaotic attractor without the underlying ordinary differential equation is posed and some pointers are given as to a possible method of solution using statistical analysis.

Contents

1 Introduction

This article concerns dynamical systems. In a chaotic system, a small change in the initial conditions has a significant effect on the future states of the system. This property is termed "sensitive dependence on initial conditions" which is defined as follows.

- $\mathbf{X}$ is a set, $x \in \mathbf{X}$.

- $\mathbf{N}$ is the neighbourhood of x, whose points are within a distance of δ from x; that is, if $y \in \mathbf{N}$ then $|x - y| < \delta$.

- $f : \mathbf{X} \to \mathbf{X}$ has sensitive dependence on initial conditions if there exists $\epsilon > 0$ such that for all $x \in \mathbf{X}$ and all $\delta > 0$ there exists a point $y \in \mathbf{N}$ and a value $n \geqslant 0$, such that

$$|f^n(x) - f^n(y)| > \epsilon \,,$$

where $f^n(x)$ and $f^n(y)$ are the nth iterates (an nth iterate is the nth state that the system is in after n equal incremental time periods.)

Some dynamical systems have an attractor of the type referred to as a 'strange non-chaotic attractor' (SNA). Let us focus on a range of dynamical systems that often have SNAs. The range we choose is one where the forcing function

is quasi-periodic. An example of such a function is

$$\sin(2\pi t) + 3\cos(2\pi\omega t).$$

With such a system, the usual approach is to change from considering an ODE to considering a difference equation. We plot a graph in order to gain an understanding of the shape of the attractor. Let us assume that our original ODE has y as the dependent variable and t as the independent variable. We do not plot y versus t (on the horizontal axis), instead we plot y versus (ωt mod 1). This is referred to as the stroboscopic map.[1] Before constructing the plot we need to convert our original ODE to a difference equation. We are interested in how y varies over the stroboscopic map. (ωt mod 1) varies from 0 to 1 even though t ranges from 0 onwards. In view of this it is preferable to define the variable

$$x = \omega t \mod 1.$$

We are interested in discrete values of $t = 1, 2, \ldots$ and so we re-write the above equation as

$$x_0 = 0, \quad x_{n+1} = x_n + \omega \mod 1.$$

The y values can be written in the form

$$y_{n+1} = g(x_n, y_n).$$

The description above considered a time series where the value of a dependent variable has been recorded (or generated) at fixed time intervals. The work described here assumes that we recorded the values by observing a dynamical system. Further, it is assumed that it is difficult, or impractical, to formulate a mathematical model for the dynamical system under study. Our work

---

[1]Our choice of ωt is based on the above example where the frequencies of the components differ by a factor of ω. Also, the amplitude of the cos term is three times that of the sin term and so is the major contributor to forcing.

describes a preliminary study into the use of statistical methods to detect the SNA 'pattern' in raw data.

The detection of patterns in large volumes of data gathered from dynamical systems has practical applications. Let us take a computer grid, for example. A computer grid is a network of computers in which the resources are pooled. 'Jobs' enter the grid and a software manager decides which computer the job is to be allocated to. Jobs may even be split up and the parts allocated to different computers. The problem of how best to schedule jobs on a grid is an open research question. If we develop a new scheduling algorithm, then we would wish to evaluate it to determine whether or not it is superior to existing algorithms. We can use example data gathered from live grids for this testing purpose. Such data are termed 'workload traces'. However, unless we are able to statistically analyse a trace to determine what patterns lie therein, then it is possible that the traces that we try are all statistically similar, in some sense, and not representative of the variety of traces that the algorithm should be able to handle. If we perfect our statistical analysis of traces, then we can begin to generate synthetic traces having any property or pattern that we wish.

2   Related work

There has been some difficulty in defining chaos. As an example Banks et al. [1] criticises an earlier definition. Let us try to establish when the dynamics is chaotic. We looked at a property of chaos—the system has sensitive dependence on initial conditions. Often researchers assumed that this property is a sufficient condition for chaos. Let us consider another property of chaos. In a chaotic system, it is only possible to predict the states of the system in the very near future, if at all. However, it is possible for a dynamical system to have sensitive dependence on initial conditions but for n to be very large. It is therefore possible to have a definition of chaos in which sensitive dependence on initial conditions is a necessary but not

sufficient condition. It could be argued that for a system to be chaotic there must be sensitive dependence on initial conditions and that changing the initial conditions manifests itself in a rapid, in some sense, spatial difference in the resultant states. To check whether the change is rapid enough we could calculate the Liapounov exponents. Having positive exponents indicates a rapid change.

Examples of SNAs have been given by Keller [7] and Glendinning [6] for example. Let us consider the example

$$y_{n+1} = 2\beta \cos(2\pi x_n) \tanh y_n, \quad \beta > 0.$$

We see that $|y_{n+1}| < 2\beta$. Glendinning [4] investigated the attractors associated with this difference equation.

Cases giving simple attractors There are two cases

1. $y_1 = 0$ and $y_2, y_3, \ldots = 0$.
2. $y_1 \neq 0$, $\beta < 1$. The attractor is 0 and y is stable about 0.

Case giving complicated attractors This occurs when $y_1 \neq 0$, $\beta > 1$. The attractor is difficult to describe because of its complexity. Figure 1 shows an example for

$$\beta = 2, \quad \omega = \frac{(\sqrt{5} - 1)}{2}.$$

We can clearly see the shape to which the y-values are attracted. Let us recap. We stated in the Introduction that we begin with an ODE in which y is the dependent variable and t is the independent variable. We also stated that we are only interested in discrete values of $t = 1, 2, \ldots$. If we were to plot a graph of y versus t, or a part of this graph, then the eye would be unable to detect a pattern. By using the modulo arithmetic on the t-values we are effectively overlapping the graph on itself. Doing this reveals the attractor.
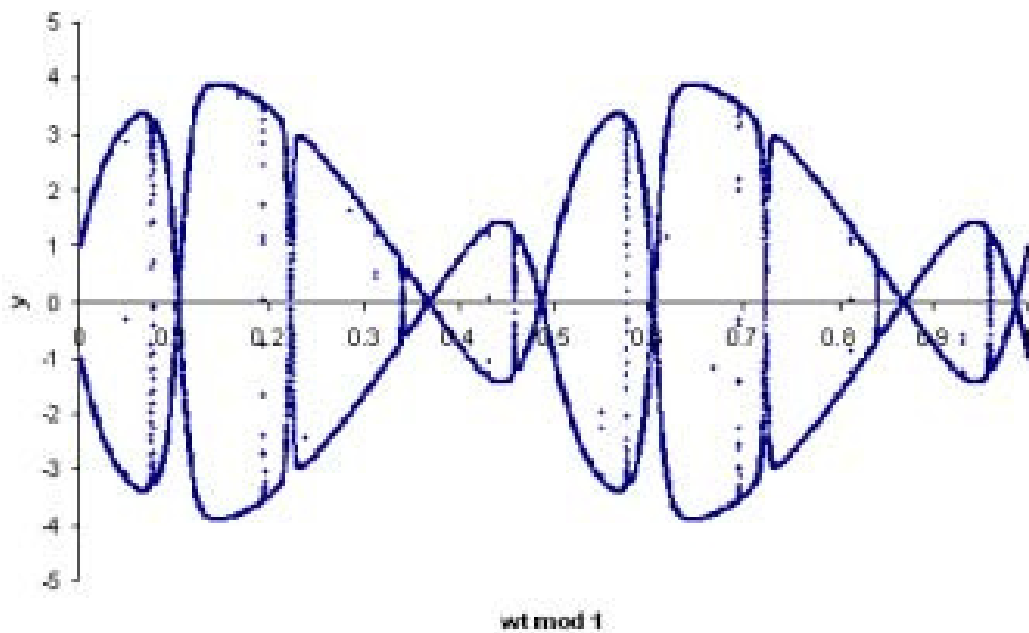
FIGURE 1:   A   plot   of $2\beta \cos(2\pi x_n) \tanh y_n$ for   points (x_n, y_n), $n =$ $1, 2, \ldots, 20,000$

Effect on changing initial conditions   Glendinning et al. [5] shows that this case has 'sensitive dependence on initial conditions'.

2.1   SNAs and statistical analysis

The statistical analysis of Time Series is a well established field [2]. Much work has been done on the analysis of some types of computer networks and parallel computers [3, 9, 10]. Relatively little work has been done on the analysis of workload traces for grids. That which has been done [8] has focused on identifying some patterns that occur in the data. It analyses a trace called LCG1 taken from the Large Hadron Collider Grid. The patterns include pseudo-periodicity, long range dependence, and multifractals. There is an absence of detail in the article which makes it difficult should one wish to duplicate the work. I have been unable to find other articles by this author, or others, which give these details. Little, or no, work seems to have been done on identifying the presence of attractors.

3   Methodology

3.1   An actual workload trace

As part of our work I felt it necessary to gain proficiency in the practical ways of using statistical methods to analyse traffic in a grid environment. To give an indication of what this involves, in this section I explain processing of trace LCG1. We are interested in analysing the job arrival times for both the trace as a whole and for some virtual organisations (VOs). We performed five analyses. In Analysis I we wished to plot two graphs, one showing the distribution of jobs amongst the VOs, and another showing the distribution of jobs amongst the users. In the remainder of the analyses we were interested in trying to identify patterns in job arrival times. Initially, in Analyses II

to IV, we did not consider the whole trace. Instead we considered one VO, whose ID is 'g1' and whose name is 'lhcb'. Later, in Analysis V, we considered the whole trace.

In Analysis II we looked at the time between each job arriving and the arrival of the succeeding job, referred to as the 'inter-arrival time'. We plotted a histogram showing the distribution of the inter-arrival times. Next we found the autocorrelation of the job inter-arrival times, at different lags. Subsequently we performed a Fourier analysis on the autocorrelation output.

In Analyses III and IV, rather than considering the job arrival times as our raw data, we used the number of jobs arriving in each 64 second period as the raw data. We refer to this new data as 'job counts'. In Anaysis III we plotted the distribution of job counts. Next we performed an autocorrelation of the job counts. In our next task we performed a Fourier analysis of the job counts. We found the power spectral estimates. In Analysis IV we found the entropy of the power spectral estimates.

In Analysis V we considered the trace as a whole. We plotted the distribution of job counts for the whole trace. Next we found the autocorrelation of the job counts, at different lags. Subsequently we performed a Fourier analysis on only 800 values of the autocorrelation output.

3.2   A synthetic trace possessing an SNA

In the example above we considered a time series where the value of a dependent variable has been recorded (or generated) at fixed time intervals. The 20,000 values used to draw Figure 1 were generated from a difference equation. In what follows, let us assume that we recorded these 20,000 values by observing a dynamical system. Further, let us assume that it is difficult, or impractical, to formulate a mathematical model for the dynamical system under study. The following work describes a preliminary study into the use of statistical methods to detect the SNA 'pattern' in raw data. We use some

statistical methods that have been used in the analysis of patterns occurring in the workload traces of grids.

Let us consider the $20{,}000$ y-values used to draw Figure 1. To say that these are the values of a dependent variable taken at $20{,}000$ points in time is only one way of interpreting these values as a time series. Let us consider an alternative meaning. The y-values could indicate the time between job arrivals, called the 'interarrival time'. The first five y-values from the $20{,}000$ are 0.3, 1.16525, -2.42666, -0.34429 and -0.8063. As they stand the values do not make sense in the context of job arrivals because an interarrival time cannot be negative. The number -4 is a lower bound to the $20{,}000$ values so if we add 4 to each of the values then all are positive. The first five values become 4.3, 5.16525, 1.573338, 3.655712 and 3.193696. The meaning of these numbers is now described. We start a clock at time 0. After 4.3 seconds a job arrives on the grid. $5.16525\,\mathrm{s}$ later another job arrives. $1.573338\,\mathrm{s}$ later a third job arrives, and so on. Even though the data has a different meaning and the values have been transformed, the SNA pattern, if present, will be preserved.

The problem is now one of trying to detect the presence or otherwise of an SNA pattern using statistical analyses. The work which follows is an initial attempt to do this using a few of the many statistical methods at our disposal.

4   Results

If we try to plot the transformed data as a time series with the x-axis representing the arrival time of a job and the y-axis representing the number of jobs then all the y-values will be one and the x-values will be unequally spaced. We change this into a time series having equal time increments by considering a fixed time quantum, say $64\,\mathrm{s}$, and counting the number of jobs that arrive in each $64\,\mathrm{s}$ period. This histogram is shown on the top left of Figure 2. One statistical method is to perform a correlation of data with
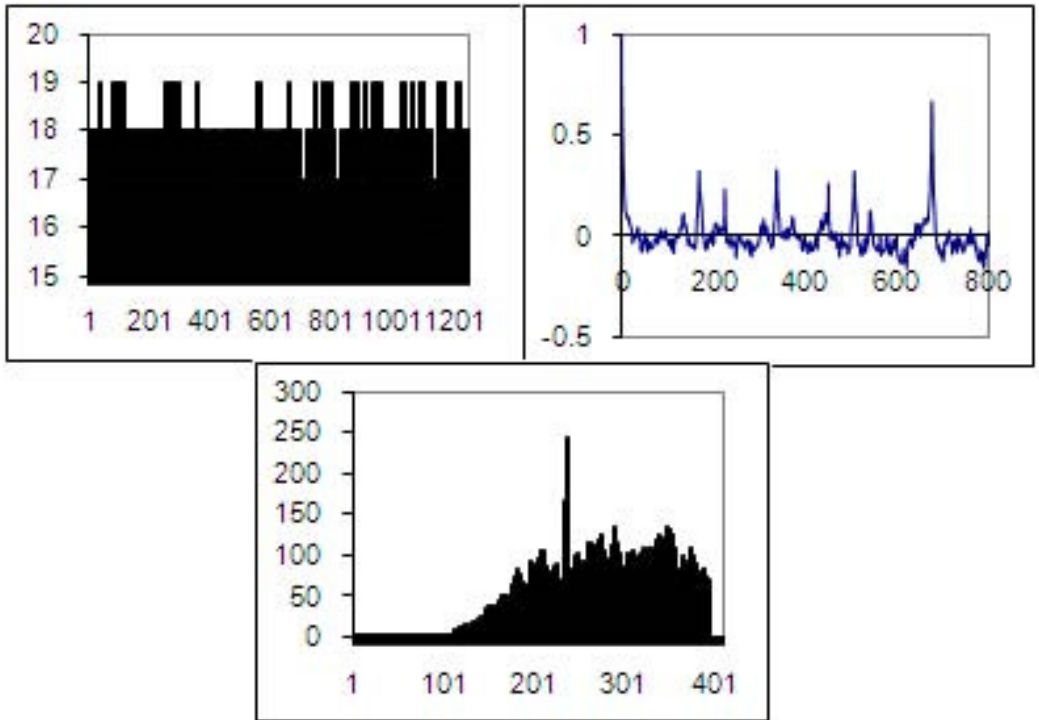
FIGURE 2: The job arrival times, autocorrelation (ACF), and discrete Fourier transform (DFT) for the trial data

itself—an autocorrelation. With a Time Series we correlate the original data with a time-shifted version of itself (called a 'lag'). A lag of one of the data in our study compares the original data with the data shifted by $64\,$s. I repeat this process for a number of different lags. This produces the top right graph of Figure 2.

Another statistical method is to perform Fourier analysis on the ACF. The algorithm used to evaluate the magnitude of the sine waves at the various frequencies is the Fast Fourier Transform (FFT). This algorithm accepts as input a set of data points that number a power of two. For example, the algorithm will accept 512 data values or $1,024$ data values. In the top right graph of Figure 2 we have 800 values. I resort to zero-padding, that is inserting 224 zeros. Normally this is done by having half of them on the left and half on the right. Unfortunately, this zero-padding is likely to be too drastic an approach and we have to perform what is known as windowing. This means that we give weights to the y-values in the top right graph of Figure 2. The y-values at the ends are almost zero whilst those in the middle are almost one. The analysis is complicated by the fact that many windowing methods exist—rectangular, Hanning, Hamming, Blackman, and so on. Furthermore, we can apply windowing to all 800 values or fewer. We could apply windowing to the 15%, say, of the values at the left hand end and the 15% at the right hand end. An example power spectrum is shown in the bottom graph of Figure 2.

5   Conclusion

The topic of chaos has been described as well as the phenomenon known as a Strange Non-chaotic Attractor. An example of an SNA has been given. These descriptions of chaos and SNAs add to the body of knowledge on these topics.

The significance of our work is threefold. Firstly, we have found that the detection of attractors in network traffic is a topic that has been neglected.

Secondly, we have shown how to transform an analytical description of an SNA into a workload trace. Thirdly, we have drawn attention to the fact that whilst we can observe properties such as self similarity and long range dependence by looking for traits in specific statistical graphs, there may be other properties, or patterns, inherent in the data for which we do not, as yet, know how these manifest themselves in statistical graphs. In particular, constructing a synthetic workload that possesses an attractor enables us to study how this pattern manifests itself. In the future, when we know the answer to this, we will be able to analyze actual workload traces to determine whether or not an attractor exists.

The problem of how one goes about detecting the presence of an SNA in raw data is posed. A small number of statistical methods which might be suitable for further investigation are described. However, even if the methods are the ones to use there are still many options when using them. In grouping data we must choose an appropriate size for the interval. In the example described, 64 s was chosen fairly arbitrarily. This is an analogous problem to finding the correct width of a column in a histogram. If the column width is too small or too large then the pattern is hidden.

When using the ACF a window size of 800 was chosen, once again fairly arbitrarily. Work needs to be undertaken on finding out the best size. Finally, when we calculate the Power Spectrum we need to find out which windowing technique is superior and to what percentage of the original data it should be applied.

References

[1] Banks J., Brooks J., Cairns G., Davis G. and Stacey P., On Devaney's definition of chaos, *American Mathematics Monthly*. **99** (1992), 332–334. C615

[2] Box G. E. P, Jenkins G. M. and Reinsel G. C., *Time series analysis* John Wiley & Sons, New Jersy, 2008. C618

[3] Feitelson D. G., Workload modeling for performance evaluation, In: *Lecture notes in computer science*, vol. 2459, Springer, Berlin, 2002, 114–141. C618

[4] Glendinning P., The nonsmooth pitchfork bifurcation, *Discrete and Continuous Dynamical Systems B.* **4** (2004), 457–464. C616

[5] Glendinning P., T. H. Jager and G. Keller, How chaotic are strange non-chaotic attractors?, *Nonlinearity.* **19** (2006), 2005–2022. C618

[6] Glendinning P., View from the Pennines: strange nonchaotic attractors, *Mathematics Today.* **44** (2008), 119–120. C616

[7] Keller G., A note on strange nonchaotic attractors, *Fundamenta Mathematicae.* **151** (1996), 139–148. C616

[8] Li H., Workload dynamics on clusters and grids, *Journal of Supercomputing.* **47)** (2009), 1–20. C618

[9] Lublin U., Feitelson D. G., The workload of parallel supercomputers: modeling the characteristics of rigid jobs, *J. Parallel Distrib Comput.* **63(11)** (2003), 1105–1122. C618

[10] Squillante M. S., Yao D. D and Zhang L. The impact of job arrival patterns on parallel scheduling, *ACM SIGMETRICS Perform Eval Rev.* **26(4)** (1999), 52–59. C618

Author address

1. **Alan Oxley**, Department of Computer and Information Sciences, Universiti Teknologi Petronas, Bandar Seri Iskandar, 31750 Tronoh, Perak Darul Ridzuan, Malaysia.
   mailto:alanoxley@petronas.com.my