Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification

Joseph P. Zbilut a,1, Alessandro Giuliani b, Charles L. Webber, Jr. c
a Department of Molecular Biophysics and Physiology, Rush University, 1653 W. Congress, Chicago, IL 60612, USA
b Ist. Superiore di Sanita, TCE Lab., Via Regina Elena 299, Rome 00161, Italy
c Department of Physiology, Loyola University Chicago, Stritch School of Medicine, 2160 S. First Ave., Maywood, IL 60153, USA

Received 20 March 1998; revised manuscript received 27 May 1998; accepted for publication 28 May 1998
Communicated by C.R. Doering

Abstract

The demonstrated ability of recurrence quantification analysis to detect very subtle patterns in time series was exploited to devise a filter able to recognize and extract signals buried in large amounts of noise. The proposed technique, cross-recurrence quantification, demonstrates the ability to extract signals up to a very low signal-to-noise-ratio and to allow an immediate appreciation of their degree of periodicity. The lack of any stationarity dependence of the proposed method opens the way to many possible applications, including encryption. © 1998 Elsevier Science B.V.

PACS: 05.40; 05.45; 07.05.Rm; 07.05.Kf
Keywords: Signal detection; Recurrence quantification analysis; Nonlinear dynamics; Time series; Encryption; Noise

1. Introduction

The detection of signals in the presence of large amounts of contaminating noise is important in many scientific situations. We have previously demonstrated the utility of combining recurrence quantification analysis (RQA) with principal components analysis to determine the probability of the existence of a deterministic signal in an apparently noisy time series [1]. Once it has been established that there is a good probability for the existence of a signal, there still remains the problem of locating the signal itself. In this respect, a modification of RQA may provide a quick solution before more computationally expensive methods are employed. The general context we are interested in, is the discovery of a signal of known structure (probe) in a very noisy situation at very low signal-to-noise-ratio (SNR) 2.

In these situations, the classical ways of filtering such as Fourier analysis or singular value decomposition fail in unequivocally identifying the signal against the noise [2,3]. More recent methods require libraries of ODEs to extract correlations [4], multiple input models [5], or more complex data transforms [6–9]. RQA, originally conceived by Kamphorst, Eckmann and Ruelle [10], on the other hand, does not require preexisting models or transformations; instead

1Corresponding author. E-mail: jzbilut@rush.edu.

2Although this Letter emphasizes the logical continuation of a previous Letter [1], we also point out that this method may be useful on its own; i.e., without knowledge of target signal structure. See below.
it is based upon simple tallies of recurrences, which may be an important consideration, especially in complex biological signals. The modification of RQA, termed cross-recurrence quantification (CRQ), deals with cross-recurrences instead of “auto-recurrences”.

2. Methods

RQA does not depend on any distributional assumptions on the data set, since it is a simple tabulation technique aimed to find the recurrent patterns inside a time series\(^3\), and has demonstrated its utility in areas ranging from molecular dynamics and nerve signals to bifurcation analysis [13–27]. The recurrent patterns are identified on the basis of an embedded distance matrix between subsequent epochs (rows) of the given series, as the number and position of pairs of points whose mutual distance is below a predefined threshold (radius). Essentially this technique is a form of autocorrelation.

\(^3\) We note that recurrences are probably the most basic method of identifying patterns. See, e.g., Refs. [11,12].

For CRQ as a signal detector, the algorithm is changed in order to transform the procedure from the computation of the auto-recurrence structure of a given series to the computation of cross-recurrences between a known signal (probe) and the noisy series of interest. The two series (the probe (\(P\)) and the test (\(T\) series) are mutually aligned starting from the first point to the last point (\(n\)) of the same length, and a distance matrix between the two series is computed as \(D_{ij} = \sqrt{(P_i - T_j)^2}\) for all the \(i, j\) pairs with \(i\) and \(j\) going from 1 to \(n\).

The distance matrix is then analyzed by scoring as recurrent all the \(i, j\) pairs corresponding to an observed distance between the \(i, j\) elements below a predefined threshold, and then coloring the corresponding pixel on the computer screen. The pattern of recurrent points on the screen is directly linked to the cross-recurrence structure of the \(P\) and \(T\) series.

3. Results

The probe signal was a simple square wave of 1000 points made by alternating patches of 100 points oscillating between 10 and 20. The intent here was to
use a signal which can be unambiguously corrupted, and difficult to recover, based upon our previous analysis (see Ref. [1]). (We have used other signals such as a sine wave, with similar results.) The test series was twelve noise corrupted versions of the original square wave, where a Gaussian noise was added to the original series going from a SNR = 10 to a SNR = 0.2 (SNR = variance of signal/variance of noise) (Fig. 1).

The application of CRQ in this case allowed a clear recognition of the signal embedded in the noisy series, even in the most extreme example (Fig. 2). The length of the recurrent horizontal lines in the screen corresponds exactly to the periodicity of the embed-
Correlation of signal-to-noise-ratio of signals with percent recurrence.

CRQ of two random signals (left), and a random signal with original square wave probe (right). Note the absence of any recurrence in the right panel.

The choice of the particular radius value was dictated by the necessity to make only the strong superpositions between the two series emerge (0.1, matrix normalized on the unit interval). It is worth noting that at an SNR of 0.2, Fourier analysis does not highlight any remarkable peak corresponding to the signal frequency (Fig. 3). The coherence function between the two signals was found to be 1 at all frequencies, thus again failing to locate any signal of interest.

The variable REC (percentage of recurrence) quantifies the general amount of superposition between the probe and test series and was found to scale very well with SNR (Fig. 4).

The pattern of cross-recurrences between two independent random series (uniform distributions between 0 and 1) does not show any regular pattern, despite an average recurrence value similar to the square wave case. On the other hand, the original square wave, when compared to the uniform distribution using the same radius adopted for the noise corrupted series, did not highlight any recurrence (Fig. 5).

4. Conclusions

The reason for the high performance of CRQ in detecting very weak signals in noisy environments resides in the extreme redundancy of the distance matrix formalization with respect to the original series. Comparing Figs. 2 and 5, it appears clear that the scoring of a significant recurrence between the two series is
amplified by the presence of a regular signal, so that, given the redundancy of the distance formalization, the same point is recognized as recurrent whenever it encounters the corresponding phase of the probe signal (Fig. 2). This in turn allows us to identify in the noisy series the exact phase of each point in terms of the embedded signal. In the case of pure noise (Fig. 5) the scored recurrences, even if in the same proportion of Fig. 2 (with respect to the amount of recurrence (0.23%)), are sporadically counted all along the sequence and do not show any ordered pattern.

We also note that the probe signal need not be the same to detect the possibility of a signal. The signal of Fig. 1 (bottom) was probed with a sine wave. The resultant CRQ reveals periodic banding, indicative of a hidden periodic signal; although, clearly the probe signal is not exact (Fig. 6).

Another demonstration that signal probes need not be exact, involves a high amplitude (10 a.u.), low frequency square wave probe (1 Hz), and a low amplitude (2 a.u.), low frequency sine wave test signal (0.75 Hz), contaminated with Gaussian noise (SNR = 0.22). To demonstrate this, it is useful to note CRQ behavior in the noiseless case (Fig. 7). An important feature of the plot is the number of recurrent bands in the first vertical stripe, which is 3. This coincides with the number of times the probe co- recurs with ascending and descending limbs of the sine wave. In the case of the noisy signal, a similar stripe occurs, with some blurring due to the noise in the top most band. Analysis reveals similar recurrence “times” in these bands (Fig. 8).

An additional possible use for this technique is the encryption of data. A noisy (chaotic) signal can be added to a signal which needs to be encrypted to act as a “key” [28]. On the receiving end, CRQ can be employed to detect the chaotic signal’s specific CRQ signature, and then subtract the signature signal to retrieve the original message (Figs. 9, 10).

The advantage of CRQ is that it does not require extensive a priori models, equations, or transformations, since it depends only on the empirical signal’s recurrence structure. We do emphasize that CRQ is an exploratory data technique. Where hidden signals exhibit
a well-defined structure more sophisticated techniques should be employed. Nonetheless, in experimental circumstances, where clear or unambiguous models are not available, CRQ can provide information not easily available otherwise.

References


