Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 22 (2018) 621–625

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

11th International Conference Interdisciplinarity in Engineering, INTER-ENG 2017, 5-6 October 2017, Tirgu-Mures, Romania

Measure of regularity in discrete time signals

Zoltán Germán-Salló[a],*

[a] "Petru Maior" University of Tirgu Mures, Romania

Abstract

Discrete time series can be treated through linear or non-linear mathematical procedures in order to find specific properties. Last decades, non-linear analysis methods brought valuable results in discrete time series analysis and prediction. Statistical signal processing methods as entropy measures have become important tools in the analysis of time series data, especially in physiology and medicine. Generally, entropy measures the degree of regularity in systems and usually it should be able to quantify the complexity of any underlying structure in the discrete time signals. This paper proposes approximate entropy and sample entropy calculations on synthesized test signals with specific properties in order to try to find hidden properties. Approximate entropy and sample entropy being mathematical algorithms created to measure the regularity or predictability within a time series, are extremely sensitive to their used parameters as length of the data segment and length of data.

Peer-review under responsibility of the scientific committee of the 11th International Conference Interdisciplinarity in Engineering.

Keywords: discrete time signals; approximate entropy; sample entropy.

1. Background

The entropy as a metric was first defined in the field of thermodynamics and interpreted as the amount of information needed to completely specify the physical state of a system. A regular system with orderly behavior has a low entropy value. Shannon defined entropy as the information content of a communication system [1].

* Corresponding author. Tel.: +40-265-233-112; fax: +40-265-233-212.
E-mail address: zoltan.german-sallo@ing.upm.ro

2351-9789
Peer-review under responsibility of the scientific committee of the 11th International Conference Interdisciplinarity in Engineering.
10.1016/j.promfg.2018.03.090

The concept of entropy remained pure theoretical until Pincus developed approximate entropy as a measure of changing complexity which could be applied to real-world data sets [2]. After that, various other entropy measures have been proposed for the same purpose. Richman and Moorman introduced sample entropy [3], a modified version of approximate entropy, eliminating the self-match bias and to improve on several other statistical properties. Bandt and Pompe proposed permutation entropy [4] as an alternative measure of complexity for time series. Costa et. al. developed multi-scale entropy [5] to account for structural interactions across multiple time scales. Nowadays the concept of entropy is widely used in the field of non-linear dynamic analysis and chaos as a measure of the complexity and a regularity of a system.

1.1. Approximate Entropy

The Approximate Entropy (ApEn) is a recently developed mathematical formula quantifying regularity over time-series data. It has proved to be a useful tool because of its ability to distinguish different system's dynamics when there is only available relatively short-length noisy data. Incorrect parameter selection (embedding dimension m, threshold r and data length N) and the presence of noise in the signal can undermine the ApEn discrimination capacity. This entropic measure was first proposed by Pincus [2], and it exhibits a good performance in the characterization of randomness.

In order to compute the approximate entropy, the embedded dimension m, that is, the length of the vectors to be compared, and a noise filter threshold r are required. Given N data points u(1), u(2), ...u(N) of time discrete signal, a sequence of vectors x(1), x(2), … x(N-m+1) is formed, where:

$$x(i) = \left[u(1), u(2), \ldots u(i+m-1)\right] \tag{1}$$

Using the previous sequence x(i) for each satisfying 1≤i≤N-m+1, on construct

$$C_i^m(r) = \frac{nr \quad j < N-m+1 \quad such \quad that \quad d\left[x(j), x(i)\right] < r}{N-m+1} \tag{2}$$

where d[x(j),x(i)] represents the distance between the vectors x(j) and x(i), given by the maximum difference in their respective scalar components. It's important to observe that j takes on all values, so the match provided when i=j will be counted (the subsequence is matched against itself). This distance is calculated as follows

$$d_{ij} = \max_{k=0,\ldots m-1}\left(\left|x(i+k)\right| - \left|x(j+k)\right|\right) \tag{3}$$

The parameter $C_i^m(r)$ expresses the prevalence of repetitive patterns of length m. The approximate entropy ApEn(m,r,N) as a parameter of embedded dimension m, threshold r and signal length N is defined as follows:

$$ApEn(m,r,N) = \begin{cases} \Phi^m(r) - \Phi^{m+1}(r) \quad for \quad m > 0 \\ -\Phi^{-1}(r) \quad for \quad m = 0 \end{cases} \tag{4}$$

where

$$\Phi^m(r) = \frac{1}{N-m+1}\sum_{i=1}^{N-m+1} \ln C_i^m(r) \tag{5}$$