

Detección de señales cardíacas arrítmicas a partir de registros de ECG utilizando el plano Entropía-Complejidad

Detection of Arrhythmic cardiac signals from ECG recordings using the Entropy-Complexity plane

Presentación: 6 y 7 octubre 2020

Doctorando:

Pablo Martínez Coq

Centro de Procesamiento de Señales e Imágenes - CPSI, Facultad Regional Buenos Aires. Universidad Tecnológica Nacional, Medrano 951, C1179AAQ, CABA, Argentina; pabломartinezcoq@gmail.com

Director/a:

Walter Legnani

Co-director/a:

Ricardo Armentano

Resumen

El objetivo de este trabajo fue analizar en el plano Entropía-Complejidad (**HxC**) un conjunto de series de tiempo provenientes de **ECG** de seres humanos con el objetivo de discriminar registros de dos grupos diferentes de pacientes: los de ritmo sinusal normal, y los de arritmias cardíacas.

Las series de tiempo utilizadas en este artículo se obtuvieron de la base de registros de ECG de PhysioNet. El set estuvo compuesto por 47 señales de pacientes con arritmias cardíacas diagnosticadas y 18 señales de pacientes con ritmo sinusal normal.

El plano HxC se construyó con la entropía de Shannon como eje **X**, y con la complejidad estadística como eje **Y**. Dado que el contenido de información de las series de tiempo se transmite en forma de una función de distribución de probabilidad, **FDP**, ésta se obtuvo por medio de la metodología propuesta por Bandt & Pompe (2002).

Los valores medios de complejidad estadística y entropía de Shannon normalizada se calcularon y analizaron en el plano HxC para cada serie de tiempo, donde los valores medios de complejidad de las series de tiempo de ECG de pacientes con arritmias diagnosticadas fueron mayores que el grupo de ritmo sinusal normal. Por otro lado, los valores medios de la entropía de Shannon para los pacientes con arritmias fueron más bajos que los del grupo de ritmo sinusal normal. Estas notorias características hicieron posible la discriminación de los diferentes espacios para ambos grupos de señales en el plano HxC. Los resultados se analizaron mediante una hipótesis de prueba estadística multivariada.

La metodología propuesta tiene una notable simplicidad conceptual, velocidad computacional, y robustez al ruido, mostrando además una promisoría eficacia en la detección de patologías cardiovasculares.

Palabras clave: Plano Entropía-Complejidad, ECG, Señales cardíacas arrítmicas, Entropía de Shannon, Entropía de permutación; Bandt & Pompe.

Abstract

The aim of this paper was to analyze in the Entropy-Complexity plane (**HxC**) a set of time series coming from **ECG** of humans, with the objective to discriminate recordings from two different groups of patients: normal sinus rhythm and cardiac arrhythmias.

The time series dataset used in this paper were ECG recordings obtained from PhysioNet. These were 47 long-term signals of patients with diagnosed cardiac arrhythmias and 18 long-term signals from normal sinus rhythm patients.

The HxC plane was built with the Shannon's Entropy on **x**-axis and the statistical complexity on **y**-axis. As the information content about the time series is conveyed in the form of a probability distribution function (**PDF**), the methodology proposed by Bandt & Pompe (2002) was applied to compute it.

Average values of statistical complexity and normalized Shannon entropy were calculated and analyzed in the HxC plane for each time series, where the average values of complexity from ECG time series of patients with diagnosed arrhythmias were bigger than normal sinus rhythm group. On the other hand, the Shannon entropy average values for arrhythmias patients were lower than the normal sinus rhythm group. These notorious characteristics made possible to discriminate different spaces of both groups of signals in the HxC plane. The results were analyzed through a multivariate statistical test hypothesis.

The methodology proposed has a remarkable conceptual simplicity, computational speed, and robustness to noise. It shows a promissory efficiency in the detection of cardiovascular pathologies.

Keywords: Entropy-Complexity plane, ECG, Arrhythmic cardiac signals, Shannon's entropy, Permutation entropy; Bandt & Pompe.

Introducción

As its published by World Health Organization (WHO) Of the 56.9 million deaths worldwide in 2016, light more than a half (54%) were due to the top 10 causes. Ischemic heart disease and stroke are the world's biggest killers, accounting for a combined 15.2 million deaths in 2016. These diseases have remained the leading causes of death globally in the last 15 years (WHO, 2019).

The American Heart Association explain that some arrhythmias are so brief (for example, a temporary pause or premature beat) that the overall heart rate or rhythm isn't greatly affected. However, if arrhythmias last longer, they may cause the heart rate to be too slow or too fast or the heart rhythm to be erratic – so the heart pumps less effectively (AHA, 2016).

The most important clinical significance of arrhythmias is related to an association with sudden cardiac death (Goldstein et al.; Recommended General Bibliography p. xvii; 1994) (Bayés de Luna & Baranchuk, 2017). It is also important to remember that frequently arrhythmias (especially atrial fibrillation) may lead to embolism, including cerebral emboli, often with severe consequences. In addition, it must be remembered that sometimes, fast arrhythmias may trigger or worsen heart failure.

For the analysis of certain time series coming from biomedical signals, such as an electrocardiogram (ECG) (Khan, 2008), (Bayés de Luna, 2007), (NASPE, 1996), there is a need to characterize the degree of organization or complexity of it. For this reason, the theory of information and the analysis of dynamics of the systems under study provide a way to obtain relevant results.

In order to analyze these time series, the first step is the computation of the PDF. For this, Bandt and Pompe introduced a simple and robust symbolic method that takes into account the time causality connected with the dynamics of the system. Without any kind of assumption of the dynamics originating the signal, the permutation's Bandt & Pompe probability distribution function allows then, to compute the entropy in such a way of with the complexity, computed from the Jensen-

Shannon information divergence, constitutes the system's behavior localization in an entropy-complexity plane (Kolmogorov, 1959), (Sinai, 1959), (Olivares, Plastino & Rosso 2012), (Rosso, Larrondo, Martin, Plastino & Fuentes, 2007), (Rosso, Olivares & Plastino 2015), (Rosso et al., 2013), (Amigó, 2010).

The concept of entropy has many interpretations arising from a wide diversity of scientific and technological fields (Olivares, Souza, Legnani & Rosso, 2019). Among them is associated with disorder, with the volume of state space, and with a lack of information too. There are various definitions according to ways of computing this important magnitude to study the dynamics of the systems, and one of the most frequent that could be considered of foundational definition is the denominated Shannon entropy (Shannon & Weaver, 1948), which can be interpreted as a measure of uncertainty. The Shannon entropy can be considered as one of the most representative examples of information quantifiers.

Let be a continuous Probability Distribution Function associated, (**PDF**), be noted by $\rho(x)$ with $x \in \Omega \subset \mathbb{R}$ and $\int_{\Omega} \rho(x) dx = 1$; its associated *Shannon Entropy* $S[\rho]$ is defined as (Brissaud, 2005):

$$S[\rho] = \int_{\Omega} \rho(x) \ln(\rho(x)) dx. \quad (1)$$

This concept means a global measure of the information contained in the time series; it has a low degree of sensitivity to strong changes in the distribution originating from a small-sized region of the set Ω .

For a time series $X(t) \equiv \{x_i; t=1, \dots, M\}$, a set of M measures of the observable X and the associated **PDF**, given by $P = \{p_i; i=1, \dots, N\}$, with $\sum_{i=1}^N p_i = 1$ and N as the number of possible states of the system under study, the *Shannon entropy* is defined in (2).

$$S[\rho] = -\sum_{i=1}^N p_i \ln(p_i). \quad (2)$$

Equation (2) constitutes a function of the probability $P = \{p_i; i=1, \dots, N\}$, which is equal to zero when the outcomes of a certain experiment denoted by the index k associated with probabilities $p_k \approx 1$ will occur. Therefore, the known dynamics developed by the dynamical system under study is complete. If the knowledge of the system dynamics is minimal, all the states of the system can occur with equal probability; thus, this probability can be modeled by a uniform distribution $P_e = \{p_i=1/N; \forall i=1, \dots, N\}$. It is useful to define the so-called *normalized Shannon entropy*, denoted as $H[P]$ which its expression is (3).

$$H[P] = \frac{S[P]}{S_{max}}. \quad (3)$$

Another information quantifier applied in this work is the Statistical Complexity Measure (SCM) which is a global informational quantifier. All the computations made in the present work were done with the definitions introduced in (López-Ruiz, Mancini & Calbet, 1995) and improved by (Lamberti, Martin, Plastino & Rosso, 2004). For a discrete probability distribution function $P = \{p_i; i=1, \dots, N\}$, associated with a time series, this functional $C[P]$ is given by (4).

$$C[P] = Q_J[P, P_e] \cdot H[P] \quad (4)$$

where H denotes the amount of "disorder" given by the normalized Shannon entropy (3) and Q_J is called "disequilibrium", defined in terms of the Jensen-Shannon divergence, given by (5).

$$Q_J[P, P_e] = Q_0 J[P, P_e] = Q_0 \left\{ S \left[\frac{P+P_e}{2} \right] - \frac{S[P]}{2} - \frac{S[P_e]}{2} \right\} \quad (5)$$

and, Q_0 denotes the normalization condition for the disequilibrium which corresponds to the inverse of the maximum possible value of Jensen-Shannon divergence, that is $Q_0 = J[P_0, P_e]$.

The $C[P]$ quantifies the existence of correlational structures giving a measure of the complexity of a time series. In the case of perfect order or total randomness of a signal coming of a dynamical system, the value of the $C[P]$ is identically null that means the signal possesses no structure. Between these two extreme instances, a large range of possible stages of physical structure may be realized by a dynamical system. These stages should be reflected in the features of the obtained PDF and quantified by a no-null $C[P]$.

The global character of the SCM arising in that its value does not change with different orderings of the PDF. So, the $C[P]$ quantifies the disorder but also the degree of correlational structures. Given that the statistical complexity not only quantifies randomness but also the degree of correlation between structures and consequently is not a trivial function of entropy, in the sense that, for a given value of H , there is a range of possible values of C between a minimum value C_{min} and a maximum value C_{max} (Martin, Plastino & Rosso, 2006). It is shown in the Figure 1.

Once evaluated the Shannon's entropy and the statistical complexity, the results can be displayed in the HxC plane (global quantifier) with the objective to discriminate an ECG coming from an arrhythmic patient from a normal sinus rhythm one.

Desarrollo

The groups of time series analyzed in this work were obtained from the PhysioNet platform (managed by members of the Computational Physiology Laboratory of the M.I.T. Institute), which are available at <https://physionet.org/>

The registry of patients with normal sinus rhythm is made up of eighteen ECG recordings registered at the Beth Israel Deaconess Medical Center. On the other hand, the registry of patients with cardiac arrhythmias is made up of forty-seven ECG recordings.

Several computational codes were developed to evaluate the values of statistical complexity C and the normalized Shannon entropy H for each one of the time series. Mean values of everyone pair of that calculus was represented in the HxC plane; where the statistical complexity was represented on Y-axis, and the normalized Shannon entropy was represented on X-axis.

Table 1 shows mean values, standard deviation, and the mean error of statistical complexity and the normalized Shannon entropy of the records under analysis. Results arising from Table 1 are represented in the **HxC** plane as it shown in Figure 1.

	Normal Sinus Rhythm		Cardiac Arrhythmias	
	Entropy - H	Complexity - C	Entropy - H	Complexity - C
Mean	0,89511	0,19060	0,81984	0,28982
S.D.	0,03002	0,04562	0,04924	0,05739
Mean Error	0,00728	0,01106	0,00726	0,00846

Table 1. Mean HxC Values for the whole database under analysis.

Conclusiones

As is shown detailed in Figure 1, the average values of Statistical Complexity obtained for the group of patients with arrhythmias were higher than those of the normal heart rate group. On the other hand, the average values of Shannon Entropy for the group of patients with arrhythmias were smaller than those of the normal heart rate group. The combination of the above characteristics allowed to discriminate the mean values of both groups of patients in the **HxC** plane. The difference between two groups of interest was analyzed through multivariate statistical tests (Test of Royston and Hotelling for samples of different sizes and with binormal distribution (Mardia, Kent & Bibby, 1995)). See Tables 2 and 3.

From values obtained from previous multivariate statistical tests, there was accepted the hypothesis than mean values of both groups of interest in the HxC plane are different.

	Normal	Cardiac Arrhythmias
Royston's statistic	1,82783	0,15397
Equivalent deg. of freedom	0,527507	0,92392
P-value	0,082330	0,662806

Table 2. Royston Test for both groups of patients with a given significance of 0.050.

According to values shown in Table 2, both groups of interest have a normal distribution.

	T^2	$F(2,60)$	p
Normal vs Arrhythmias	40,141	19,741	0,0000002583

Table 3. Hotelling Test for both groups of interest.

This is showing that the signals from patients with normal ECG have less statistical complexity in their waveform (measured from Jensen-Shannon's divergence of information and higher entropy values. This result confirms (as studied in the literature (Sprott, 2003) that a normal ECG would be closer to the regions of the entropy complexity plane that usually brings together the dynamic systems called dissipative (by decreasing their n-dimensional volume in the phase space) and that they comprise the systems that develop deterministic chaos (characteristic of the control system of a healthy heart), while an ECG from a patient with arrhythmia is located in a region of the HxC plane that characterizes the so-called k-noise and fractional Brownian motion. This last result is indicating that cardiac arrhythmia could occur due to the loss of non-linear control (with the development of deterministic chaos behavior) of the heart giving a way to a more disorderly (random) dynamics typical of the pathology.

With the results obtained, an objective interpretation of a cardiac pathology can be made from the informational measures (complexity and entropy) when they are used together, forming what is agreed to be called the HxC plane, explaining in part of the disease process in itself.

Given that the representation in the complexity-entropy plane is a very useful method for the analysis of signals from the ECG of patients with different pathologies, providing an efficient way to make a statistical and deterministic analysis; the proposed methodology has a remarkable conceptual simplicity and shows promising efficacy in the detection of cardiovascular pathologies.

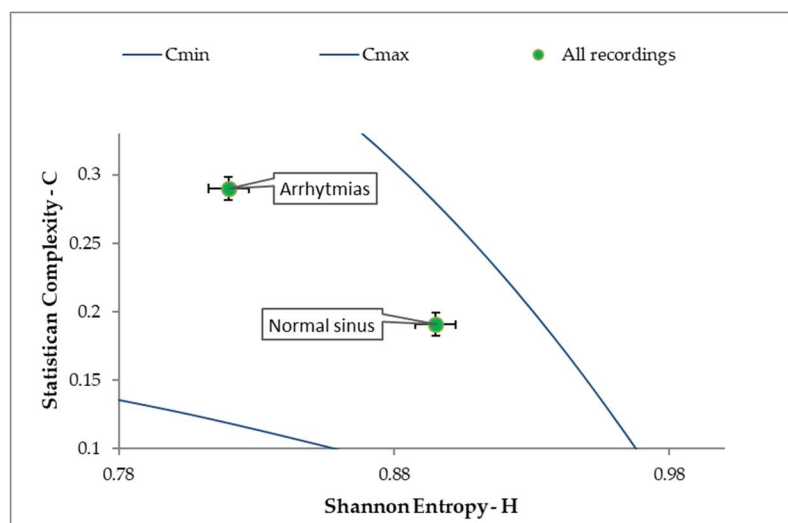


Figure 1. Detailed HxC for discriminating the mean values of both groups of interest.

Referencias

- Bandt, C.; Pompe, B. (2002). Permutation entropy: a natural complexity measure for time series. *Phys. Rev. Lett.*, 88, 174102.
- WHO. The top 10 causes of death. Last visited, June 2019. Available online at: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>
- American Heart Association. About Arrhythmia. 2016. Last visited, May 2019. Available online: <https://www.heart.org/en>
- Bayés de Luna, A; Baranchuk, A. (2017). *Clinical Arrhythmology*. John Wiley & Sons Ltd.
- Khan, G. (2008). *Rapid ECG interpretation*. Humana Press Inc.
- Bayés de Luna, A. (2007). *Basic Electrocardiography: Normal and Abnormal ECG patterns*. Blackwell Publishing.
- Heart rate variability - Task Force of The European Society of Cardiology and The North American Society of Pacing and electrophysiology. (1996). *European Heart Journal* 17, 354-381
- Kolmogorov, A. (1959). A new metric invariant for transitive dynamical systems and automorphisms in Lebesgue spaces. *Doklady Akademii Nauk SSSR.*, 119:861-864
- Sinai Y. (1959). On the concept of entropy for a dynamical system. *Doklady Akademii Nauk SSSR.*, 124:768-771
- Olivares, F; Plastino, A; Rosso, O. (2012). Ambiguities in the Bandt and Pompe's methodology for local entropic quantifiers. *Physica A: Statistical Mechanics and Its Application.*; 391:2518-2526.
- Rosso, O.A; Larrondo, H.A; Martin, M.T; Plastino, A; Fuentes, M.A. (2007). Distinguishing noise from chaos. *Physical review letters*, 99(15):154102.
- Rosso, O; Olivares, F; Plastino, A. (2015). Noise versus chaos in a causal Fisher- Shannon plane. *Papers in Physics*; 7:070006.
- Rosso, O.A; Olivares, F; Zunino, L; De Micco, L; Aquino, A.L.L; Plastino, A; Larrondo, H.A. (2013). Characterization of chaotic maps using the permutation Bandt–Pompe probability-distribution. *The European Physics Journal, B*, 86(4):116–128.
- Amigó, J. (2010). *Permutation Complexity in Dynamical Systems*. Springer-Verlag.
- Olivares F; Souza L; Legnani W; and Rosso O. (2019). Informational Time Causal Planes: A Tool for Chaotic Map Dynamic Visualization. *Nonlinear Systems - Volume 2*. DOI: 10.5772/intechopen.88107.
- Shannon, C.E.; Weaver, W. (1948). The mathematical theory of communication. *Bell Syst. Tech. J*, 27, 379–423.
- Brissaud, JB. (2005). The meaning of entropy. *Entropy*, 7:68-96
- López-Ruiz R; Mancini H. L, Calbet X. (1995). A statistical measure of complexity. *Phys. Lett, A*, 209:321–326.
- Lamberti, P; Martín, M; Plastino, A; Rosso, O. (2004). Intensive entropic non-triviality measure. *Physica A: Statistical Mechanics and Its Applications*; 334:119-131
- Martin, M.T.; Plastino, A.; Rosso, O.A. (2006). Generalized statistical complexity measures: Geometrical and analytical properties. *Phys. A Stat. Mech. its Appl*, 369, 439–462.
- Mardia, K; Kent, J; Bibby, J. (1995). *Multivariate Analysis*. Academic Press.
- Clinton Sprott J. (2003). *Chaos and Time-Series analysis*. Oxford University Pres.