journal homepage: www.intl.elsevierhealth.com/journals/cmpb

Influence of QRS complex detection errors on entropy algorithms. Application to heart rate variability discrimination

Antonio Molina–Picó[a], *David Cuesta–Frau*[a], *Pau Miró–Martínez*[b], *Sandra Oltra–Crespo*[a], *Mateo Aboy*[c]

[a] *Technological Institute of Informatics, Polytechnic University of Valencia, Alcoi Campus, Plaza Ferrandiz y Carbonell 2, Alcoi, Spain*
[b] *Department of Statistics, Polytechnic University of Valencia, Alcoi Campus, Alcoi, Spain*
[c] *Department of Electrical Engineering at Oregon Institute of Technology, Portland, OR, USA*

ARTICLE INFO

Article history:
Received 3 April 2012
Received in revised form
11 September 2012
Accepted 27 October 2012

Keywords:
Heart rate variability
QRS detection
Approximate entropy
Sample entropy

ABSTRACT

Signal entropy measures such as Approximate Entropy (ApEn) and Sample Entropy (SampEn) are widely used in Heart Rate Variability (HRV) analysis and biomedical research. In this article, we analyze the influence of QRS detection errors on HRV results based on signal entropy measures. Specifically, we study the influence that QRS detection errors have on the discrimination power of ApEn and SampEn using the Cardiac Arrhythmia Suppression Trial (CAST) database. The experiments assessed the discrimination capability of ApEn and SampEn under different levels of QRS detection errors. The results demonstrate that these measures are sensitive to the presence of ectopic peaks: from a successful classification rate of 100%, down to a 75% when spikes are present. The discriminating capability of the metrics degraded as the number of misdetections increased. For an error rate of 2% the segmentation failed in a 12.5% of the experiments, whereas for a 5% rate, it failed in a 25%.

1. Introduction

The accurate detection of QRS complexes by finding the associated R wave provides the basis for a number of electrocardiogram (ECG) analysis tasks such as heart rate estimation [1], arrhythmia detection [2], beat/wave segmentation [3], signal compression [4], and many other diagnosis aids [5].

This paper focuses on the application of QRS detection methods to extract beat–to–beat time intervals (RR series) and the influence that detection errors may have on the automatic analysis of these series. Particularly, our study assesses the robustness of two widely used signal entropy measures, Approximate entropy (ApEn) and Sample entropy (SampEn), against the presence of outliers due to detection errors. These non–linear signal processing techniques, and other metrics such as Lempel–Ziv Complexity (LZC) and Multiscale entropy (MSE), have been used successfully in the analysis of RR records [6–9]. Entropy estimation methods can potentially distinguish among a wide variety of RR records by revealing small changes in the dynamics of the heart rate. The underlying cardiac pathologies and results of the entropy estimation are frequently correlated.

RR series often include outliers, which usually manifest in the form of signal ectopic spikes (Fig. 1) Karlsson et al. (2012). These spikes (impulses) are caused by QRS detection errors [11] due to signal noise, and usually lead to abnormal and abrupt long (missing heartbeats) or short (false positives) RR intervals.

The influence of such spikes on the entropy calculations remains to be studied. Although it has been claimed that ApEn and SampEn are robust to outliers [12,13], we demonstrated in

E-mail address: dcuesta@disca.upv.es (D. Cuesta–Frau).
0169-2607/$ – see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cmpb.2012.10.014

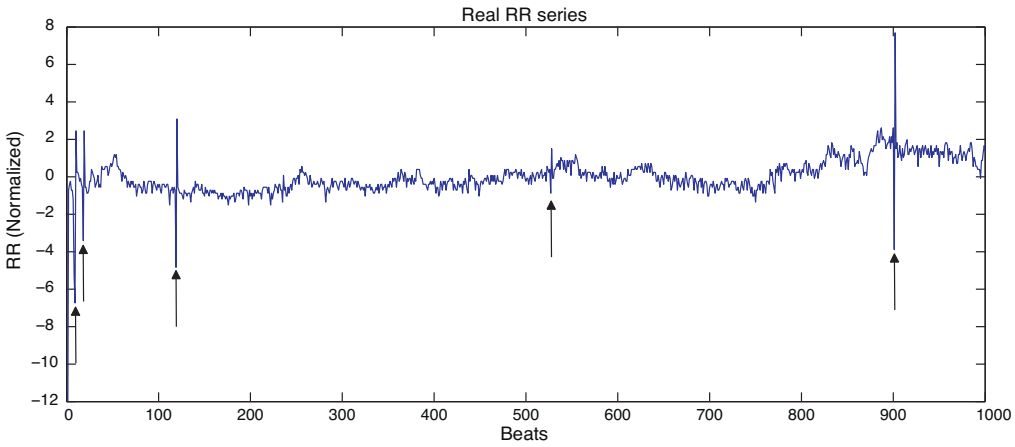

Fig. 1 – Example of a real RR record with spurious peaks due to R wave detection errors. An abnormal above–trend peak usually corresponds to one or more undetected beats. This false negative detection results in an abruptly longer beat, Below–trend peaks account for incorrectly detected heartbears. This false positive detection results in an abruptly shorter beat. It is quite common to find a false positive followed by a longer heartbeat afterwards, as shown in beats 10, 19, 120, 528, and 902 (indicated by arrows).

a previous work [14] that, in general, a small fraction of extraneous spikes may elicit a great entropy variation. Only a few other studies have addressed the characterization of entropy measures in the context of RR time series [15,16], but using a different approach.

We assess the performance of ApEn and SampEn outliners in the framework of Heart Rate Variability (HRV) [17]. HRV is a widely used non-invasive marker for the Autonomic Nervous System (ANS) [18]. It is applied to spot early signs of cardiac disease [19] and to obtain feedback on patient outcomes [20]. In addition to a pathological condition, HRV also depends on other factors [21–24]. Extrinsic factors that influence HRV are activity, mental or physical stress, sleep apnea, smoking, posture, or pharmacological condition. Intrinsic factors are age, gender, respiratory sinus arrhythmia, baroreceptor reflex regulation, thermoregulation, neuroendocrine secretion and circadian rhythms.

There are three basic approaches to quantify HRV: time domain [25] and geometric [26] analyses, frequency domain [27] analysis, and non-linear analysis [16]. All of them are based on the segmentation of heartbeats from an ECG by means of a QRS detection algorithm [11]. The analysis of changes in the associated RR series constitutes the basis for HRV studies.

Our hypothesis, based on previous findings [14], is that entropy changes induced by the presence of spikes in RR records might blur the discrimination capability of ApEn and SampEn measures. We used a non–linear HRV analysis for problem contextualization. The experimental dataset was collected from the CAST study database [28–31] publicly available at Physionet [32].

2. Methods

This paper is aimed at unveiling the influence that some outliers, specifically abnormal signal spikes, may have on the results of entropy metrics. It is obvious that any change in a time series will produce an effect on such calculations but, what remains to be investigated, is if these changes also degrade the signal class segmentation capability of entropy measures, probably the most used capability of such metrics.

The method proposed to achieve this goal includes the computation of ApEn and SampEn as the representatives of the measures to characterize, over a publicly available RR records experimental set. These records were prefiltered before being added synthetic spikes, and the entropy results were statistically assessed in terms of signal class separation accuracy. All these steps are described next.

2.1. Signal entropy measures

2.1.1. ApEn

ApEn [33] is an estimator of the regularity of a data series. It has been widely applied to a number of biomedical signal analysis problems. In the context of HRV, ApEn has been used in numerous studies (over 155 works in PubMed), such as the ones described in [34–36]. ApEn is currently one of the most popular non-linear signal processing methods.

The ApEn of a time series $\langle x(n)\rangle$ of embedded dimension m, distance threshold r, and length N, ApEn(m, r, N) is computed as follows:

(a) Take $N-m+1$ vectors of length m, $X_m(1)$, $X_m(2)$, …, $X_m(N-m+1)$, defined as $X_m(i)=x(i)$, $x(i+1)$, …, $x(i+m-1)$, for $1\le i\le N-m+1$. These vectors are m consecutive values of x, commencing at the ith sample.

(b) Define a distance between vectors $X_m(i)$ and $X_m(j)$, $d[X_m(i), X_m(j)]$ as:

$$d[X_m(i), X_m(j)] = \max_{0\le k\le m-1}(|x(i+k)-x(j+k)|)$$

(c) For a given $X_m(i)$, count the number of $j(1\le j\le N-m)$, such that $d[X_m(i), X_m(j)]\le r$. To this end, define:

$$C_{r,m}(i) = \frac{k_{i,m}(r)}{N-m+1}$$