Commun Nonlinear Sci Numer Simulat 37 (2016) 115–124

Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

journal homepage: www.elsevier.com/locate/cnsns

Dissimilarity measure based on ordinal pattern for physiological signals

Jing Wang[a], Pengjian Shang[b], Wenbin Shi[b], Xingran Cui[c,*]

[a] Beijing Key Lab of Traffic Data Analysis and Mining, School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, PR China
[b] Department of Mathematics, School of Science, Beijing Jiaotong University, Beijing 100044, PR China
[c] Division of Interdisciplinary Medicine and Biotechnology, Department of Medicine, Beth Israel Deaconess Medical Center/ Harvard Medical School, Boston, MA, US

ARTICLE INFO

Article history:
Received 10 September 2015
Revised 3 December 2015
Accepted 11 January 2016
Available online 20 January 2016

Keywords:
Ordinal pattern
Dissimilarity
Heart rate
Physiologic state

ABSTRACT

Complex physiologic signals may carry information of their underlying mechanisms. In this paper, we introduce a dissimilarity measure to capture the features of underlying dynamics from various types of physiologic signals based on rank order statistics of ordinal patterns. Simulated $1/f$ noise and white noise are used to evaluate the effect of data length, embedding dimension and time delay on this measure. We then apply this measure to different physiologic signals. The method can successfully characterize the unique underlying patterns of subjects at similar physiologic states. It can also serve as a good discriminative tool for the healthy young, healthy elderly, congestive heart failure, atrial fibrilation and white noise groups. Furthermore, when investigated into the details of underlying ordinal patterns for each group, it is found that the distributions of ordinal patterns varies significantly for healthy and pathologic states, as well as aging.

1. Introduction

Physiological systems generate complex fluctuations in their output signals that reflect the underlying dynamics [1]. For example, the variations of beat-to-beat intervals (RR-intervals) extracted from ECG signals may reflect the complex interactions of control loops of the cardiovascular system and its nonlinear response to perturbations and external stimuli. Cardiac diseases often manifest themselves in characteristic changes in the heart rate variability and in the corresponding patterns of beat-to-beat intervals (RR intervals) [2]. Consequently, exploring and analyzing the hidden dynamical structures within these physiological signals are of both basic and clinical interests.

Recently, various methods and techniques have been applied to analyze the temporal evolution of heart rate from ECG recordings. They range from traditional linear methods, such as conventional heart rate variability (HRV) analysis, including both time and frequency domain analysis (e.g., SD, SDNN, pNNx, LF, HF, etc.) [3–13], to nonlinear methods derived from the theory of nonlinear dynamics and chaos theory, such as methods based on scale invariant (fractal) behavior (power law analysis [14–16] and detrended fluctuation analysis [17–21]) and regularity (approximate entropy [22–25], sample entropy [26,27], multiscale entropy [28–33], etc.). To some extent, the nonlinear approaches are capable of extracting informative features from heart rate signals, and moreover, they are more superior to the traditional linear methods. Nonetheless, such

* Corresponding author. Tel.: +1 8572772830.
E-mail address: xingrancui@gmail.com (X. Cui).

http://dx.doi.org/10.1016/j.cnsns.2016.01.011
1007-5704/

series of techniques assume that the signals are stationary or originate from a low dimensional nonlinear system [34]. A real ECG signal or its corresponding RR-interval series is nonstationary, nonlinear and usually carrying noises. Therefore, we should be cautious when applying such approaches to heart rate signals. It is thus important to develop new approaches to characterize the variations of heart rate signals in different physiological and pathological states.

Recently, Bandt and Pompe proposed an ordinal time series analysis, called permutation entropy [35], which measures the irregularity of time series. The basic premise of this method is consideration of the order relations between the values of a time series but not the values themselves. The advantages of this method, are its simplicity, extremely fast calculation and its robustness even in the presence of observational and dynamical noise [35]. This method exhibits a fundamental distinction between deterministic chaos and noisy systems [36]. Such advantages facilitate the use of techniques based on the Bandt–Pompe algorithm to investigate the intrinsic ordinal structures in complex time series from physical systems and physiological systems [2,34,37]. Besides, a rank order statistics based on the concept of permutation entropy has also been developed to study directionality in the coupling between different time series [38].

The output signals as information carriers are commonly observed in nature. Some are unique human creations, such as paintings, music and languages; some are generated by natural process, such as DNA sequences and heart rate signals. One core problem is that how to effectively characterize the information carried by different signals and then categorize the information based on their origins. As we know, different composers usually display distinct and differentiable “rhythms” in their musical compositions which can be recognized by experienced listeners; different writers may have their personal preference for “words” used in their literature. Then some interesting questions come to us: For heart rate signals or even general physiologic signals, are there unique underlying repetitive patterns equivalent to “rhythms” in musical compositions or “words” in literatures? Can these repetitive patterns be characterized by ordinal patterns? Furthermore, can it be quantitatively compared between one style of repetitive ordinal patterns and another?

In this paper, we will discuss above questions based on heart rate signals. We also develop a new dissimilarity measure, which characterizes the distinction between two dynamical systems or states. This measure can be exploited to detect and quantify the temporal structures in physiologic time series based on ordinal pattern statistics. The structure of the paper is as follows. In Section 2, we briefly introduce ordinal pattern and dissimilarity measure. In Section 3, some simulated signals are generated to validate the effectiveness of dissimilarity measure. Section 4 is devoted to analysis of similarity measure on physiological signals. Finally, a conclusion is presented in Section 5.

2. Methods

In this section, we first briefly review the construction of ordinal patterns, and then introduce the calculation of dissimilarity measure.

2.1. Ordinal patterns

Ordinal patterns consider the order relations between the values of a time series instead of the values themselves [39]. Given a one-dimensional time series, $\{x_t\}_{t=1}^N$ of length N, an embedding procedure is applied to generate $N-(m-1)\tau$ vectors $V_1, V_2, \ldots, V_{N-(m-1)\tau}$ defined as:

$$V_t = \{x_t, x_{t+\tau}, \ldots, x_{t+(m-1)\tau}\} \tag{1}$$

where m is the embedding dimension, τ is the time lag. An ordinal pattern of order m at time t (also called permutation) $\pi_m^\tau(t) = (r_1, r_2, \ldots, r_m)$ of (1,2, …,m) satisfies

$$x_{t+r_1\tau} \le x_{t+r_2\tau} \le \cdots \le x_{t+r_{m-1}\tau} \le x_{t+r_m\tau} \tag{2}$$

which is defined to rearrange the elements in vector V_t in an increasing order. In order to obtain a unique result, we set $r_{l-1} < r_l$ in the case of $x_{t+r_{l-1}\tau} = x_{t+r_l\tau}$. There will be $m!$ possible ordinal patterns for order m. As shown in Fig. 1a, for $m=3$, there are six possible ordinal patterns. Fig. 1b illustrates the definition of ordinal patterns for the white noise time series. Given a white noise time series {0.27,0.37,0.99,0.85,0.46,0.85,0.50, …}, in order obtain permutation $\pi_3^2(3)$, one has to compare the values $x_{t+r_1\tau} = x_{3+0\tau} = x_3 = 0.99$, $x_{t+r_2\tau} = x_{3+1\tau} = x_5 = 0.46$, $x_{t+r_3\tau} = x_{3+2\tau} = x_7 = 0.50$. Clearly, $x_{t+r_2\tau} \le x_{t+r_3\tau} \le x_{t+r_1\tau}$, implying $\pi_3^2(3) = (231)$. Similarly, $\pi_3^4(9) = (321)$.

Applying above procedures, we map original time series $\{x_t\}_{t=1}^N$ to a new series $\{\pi_m^\tau(t)\}$ constituted of ordinal patterns π_k $(k = 1, \ldots, m!)$, which represents a unique pattern of fluctuations for a given time series. We then count the occurrences of each ordinal patterns π_k, denoted as $n(\pi_k)$, and calculate the relative frequency $p(\pi_k) = n(\pi_k)/(N-(m-1)\tau)$. They are then re-sorted according to descending frequency, to gain a rank-frequency distribution. Fig. 1c and d plots $p(\pi_k)$ and corresponding rank-frequency distribution for white noise and sine wave noise. For instance, the first ranked pattern π_k corresponds to one type of ordinal patterns that is most frequently found in the original time series, such as (123) and (321) in sine function series. In contrast, the last ranked pattern π_k defines the rare type of ordinal pattern in original series.