Computers in Biology and Medicine 56 (2015) 167–174

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Permutation entropy analysis of vital signs data for outcome prediction of patients with severe traumatic brain injury ☆

Konstantinos Kalpakis [a,*], Shiming Yang [b], Peter F. Hu [b], Colin F. Mackenzie [b], Lynn G. Stansbury [b], Deborah M. Stein [b], Thomas M. Scalea [b]

[a] Department of Computer Science and Electric Engineering, University of Maryland, Baltimore County, MD 21250, United States
[b] University of Maryland School of Medicine, Baltimore, MD 21201, United States

ARTICLE INFO

Article history:
Received 21 May 2014
Accepted 7 November 2014

Keywords:
Permutation entropy
Ordinal pattern
Traumatic brain injury
Vital signs
Prediction

ABSTRACT

Permutation entropy is computationally efficient, robust to outliers, and effective to measure complexity of time series. We used this technique to quantify the complexity of continuous vital signs recorded from patients with traumatic brain injury (TBI). Using permutation entropy calculated from early vital signs (initial 10–20% of patient hospital stay time), we built classifiers to predict in-hospital mortality and mobility, measured by 3-month Extended Glasgow Outcome Score (GOSE). Sixty patients with severe TBI produced a skewed dataset that we evaluated for accuracy, sensitivity and specificity. The overall prediction accuracy achieved 91.67% for mortality, and 76.67% for 3-month GOSE in testing datasets, using the leave-one-out cross validation. We also applied Receiver Operating Characteristic analysis to compare classifiers built from different learning methods. Those results support the applicability of permutation entropy in analyzing the dynamic behavior of TBI vital signs for early prediction of mortality and long-term patient outcomes.

1. Introduction

Traumatic brain injury (TBI) is the most common cause of admission to emergency care and trauma-related death in the U.S. civilian population and is a major cause of death and disability in combat causalities [1,2]. In most modern intensive care units (ICUs), vital signs (VS), such as heart rate (HR), blood pressure (BP), and oxygen saturation (SpO2), among others, are collected in high-quality, automated, continuous electronic data streams, as sequential assessments of important physiological functions, providing basic evidence of patients' status. Because VS are an early warning system of physiologic perturbation, they are usually recorded hourly in the ICU setting. However, in most modern ICUs, the massive quantities of high-quality data produced create both a challenge to store, analyze, and interpret and an opportunity to explore novel advanced analytic methods for predicting outcomes. Such predictive algorithms can support advanced instrumentation and decision-assist tools that have the potential to significantly improve clinical outcome for these very ill patients.

To discover the intrinsic patterns that characterize continuous, multivariate, clinical time series, a variety of methods can be used, such as entropy, auto-correlation, autoregressive models, and structure models [3]. One strategy is to embed the time series into higher dimensional space and then compute various entropies for the elements of the embedded time series. Conventional entropies such as Shannon entropy, Rényi entropy and Tsallis entropy can be calculated given the distribution of elements of the embedded time series. The Rényi entropy of a time series has been used to detect spatially varying multivariate relationships [4] and to study brain injuries [5] and heart rate variability [6,7]. The Tsallis entropy of the elements of a time series has been used to monitor brain injuries after cardiac arrests [8], and to improve the accuracy of gene regulatory networks inference [9].

Bandt and Pompe [10] introduced permutation entropy as a new measure of complexity of non-linear time series. Zanin et al. [11] provide an extensive review of various biomedical applications of permutation entropy. Permutation entropy has been used to predict the onset of epileptic episodes from EEG data by considering changes in the permutation entropy of the EEG time series over time [12,13]. Veisi et al. [14] find that permutation entropy can be used to effectively classify EEG signals into normal vs. epileptic with an accuracy of 85% even for highly noisy EEG. Physiologically, epileptic episodes/symptoms are manifested with deterministic behavior of the EEG signals, while healthy states are characterized by higher non-chaotic state variability [7]. Permutation entropy has

☆An earlier version of this work has appeared at the 8th International Conference, Machine Learning and Data Mining (MLDM), Berlin, Germany, Jul. 13–20, 2012.

* Corresponding author. Tel.: +1 410 455 3143.
E-mail address: kalpakis@umbc.edu (K. Kalpakis).

http://dx.doi.org/10.1016/j.compbiomed.2014.11.007
0010-4825/

also been used to study sleep using EEGs [15,16], to identify motifs in the EEG signals of patients given fast acting anesthetic drugs [17,18], and to identify temporal gene expression profiles [19]. Bian et al. [20] use permutation entropy to identify heart rate variability under different physiological conditions. Berg et al. [21] use ordinal patterns of beat-to-beat heart rate variability from an EKG signals of 40 patients who suffered from myocardial infraction, and try to classify them based on whether they survived for more than two years or not. They achieve a classification accuracy of 85%. Permutation entropy can also highlight forbidden patterns: state-space patterns/permutations appear very infrequently or not at all [22]. It can also be used to quantify non-linear interactions among time series by considering the relative entropy of the joint Takens embeddings of such time series versus the product of independent Takens embeddings [23].

In this research effort, we use permutation entropy to derive features from continuous, multivariate, time series for outcome prediction of patients with severe TBI. The remainder of this paper is organized as follows. In Section 2, we briefly introduce the permutation entropy and the entropy map that we used for quantifying the characteristics of the dynamic system. We use different independence tests to assist in variable selection. In Section 3, we describe the dataset and experiment design. We apply the permutation entropy to predict mortality and 3-month Extended Glasgow Outcomes Scale (GOSE), and present experiment results, evaluated by accuracy and the area under the receiver operating characteristic (ROC) curve. We conduct preliminary interpretation of ordinal patterns derived from the VS. Finally, in Section 4, we discuss and summarize the results.

2. Method

2.1. Ordinal patterns and permutation entropy

We assume that the physiological status of living things is dynamic but has identifiable and repeated patterns. Likewise, we assume that these patterns will be different in the healthy, injured, and/or ill individuals and that the patterns will be discernibly different from each other. For instance, if the patient is also losing blood, blood pressure (BP) will fall. Heart rate (HR) increases to compensate for the decreased BP to ensure adequate circulation and oxygenation of the brain, and the increase in HR usually increases the BP, at least temporarily. If blood loss continues, BP falls, and clinicians will usually give fluid, including blood, to raise the BP and ensure adequate oxygenation. These changing patterns of HR and BP are accompanied by changes in intracranial pressure (ICP), cerebral perfusion pressure (CPP), and so on.

Bandt and Pompe [10,24] suggested an approach to time series analysis in which they embedded a continuous time series as a symbolic sequence into another space, a process which they called "permutation entropy". One major ingredient of permutation entropy is the ordinal pattern. The ordinal pattern of a sequence of elements $x_1, \ldots, x_n$ is the permutation (re-arrangement) $\pi = (i_1, i_2, \ldots, i_n)$ that sorts the amplitude values in ascending order so that $x_{i_1} \le x_{i_2} \le, \ldots, \le x_{i_n}$.

The order L permutation entropy of a time series $x_{1,\ldots,N}$ is calculated as follows. Let π_t be the ordinal pattern (i.e. the sorting permutation) for the segment of the time series under the sliding window of length L that ends at x_t, i.e. the subsequence $x_{t-L+1}, \ldots, x_t$. Let $S_L = \{\pi_t\}$ be the set of all those unique (alphabet) ordinal patterns π_t. The time series $x_{1,\ldots,N}$ corresponds to the sequence $\langle \pi_t : t = L, \ldots, N \rangle$ of $N-L+1$ ordinal patterns from the alphabet S_L. The entropy of this sequence of ordinal patterns is the permutation entropy of the time series $x_{1,\ldots,N}$. For example, the Shannon permutation entropy is defined in (1),

$$H_L = -\sum_{\pi_k \in S_L} P(\pi_k)\log(P(\pi_k)), \tag{1}$$

where $P(\pi_k)$ is the frequency of π_k in the sequence $\langle \pi_t \rangle$. In the work presented here, we use the Rényi entropy with parameter α of the sequence $\langle \pi_t \rangle$ defined as

$$R_L^{\alpha} = \frac{1}{1-\alpha}\log\left(\sum_{\pi_k \in S_L} P(\pi_k)^{\alpha}\right). \tag{2}$$

The parameter α in the Rényi entropy acts as a selector of probabilities. It assigns almost equal weight to each possible probability when α is sufficiently close to zero. When α is larger, it puts more weight on higher probabilities. We can use this parameter to assign different weights on events of different probabilities.

2.1.1. Theoretical foundations of permutation entropy

The idea of permutation entropy, introduced by Bandt and Pompe [10,24], relies on a large body of previous work on using information theory to study the phase space (state-space) of dynamical systems [23]. For example, the Kolmogorov–Sinai (KS) entropy is used extensively to characterize the probability distributions (random processes) induced by finite partitions of the state-space of dynamical systems [23,25].

The underlying distribution of the states is an invariant measure of a dynamical system (invariant under smooth transformations of the state space), while entropy functions provide us with a way to compare such distributions. Due to the intractability in deriving explicit analytic expressions of the state distributions, researchers have resorted to numerical estimates from the data. To this end, of particular importance is the Takens–Whitney delay embedding and reconstruction theorem [23] that relates the dimension d of the system's attractor and the dimension $(2d+1)$ of the embedding space that is sufficient to reconstruct those properties of the system's attractor that are invariant under smooth transformations. Characterizing the attractor of a dynamical system enables us to predict the system's long-term behavior (since the attractor contains all states that are mapped by the system back into a state in the attractor). The Takens–Whitney theorem provides an effective way to estimate the dimension of the attractor by estimating the Kolmogorov–Sinai entropy of an embedding of a system's state-space.

One particular partition of the space of a Takens delay embedding is obtained via permutations as follows. Consider for simplicity a univariate discrete-time time series x_t, and its Takens delay embedding of order m with delay lag τ: $X_t = (x_{t-(m+1)\tau}, \ldots, x_{t-\tau_t}, x_t) \in R^m$. Partition the space R^m into $m!$ subsets, each labeled by a unique permutation π of $[1, \ldots, m]$, with each subset containing all points in R^m that can be sorted by the subset's labeling permutation.

Permutation-based partitions are more robust to noise and other non-linear distortions and artifacts than value-based fixed-size partitions of the state space, since they depend on the relative order rather than the exact values of the time series. Furthermore, in order to obtain reliable entropy estimates with fixed-size partitions, one needs long time series (in the order of 2^m in order to cover all blocks of such fixed-size partitions); permutation-based entropy estimates do not require long time series. The robustness of permutation entropy makes it particularly attractive for mining vital signs collected in real clinical settings, without expensive pre-processing and cleaning of such signals.

Recall that the uniform distribution has maximum entropy among discrete distributions of bounded support. Large values of permutation entropies correspond to dynamical systems with substantial uncertainty/randomness (divergence in time of initially