Application of weighted sample entropy in detection of abnormal signal

Hong Zhang¹*, Shengze Lv² and Shasha He³

¹xi’an university of science and technology; *zanghong@xust.edu.cn
²xi’an university of science and technology; 18351926906@163.com
³xi’an university of science and technology; 2505169163@qq.com

Abstract. An improved sample entropy (SE) algorithm named weighted sample entropy (WSE) is proposed in this paper to optimize the problems existing in SE based on weighting. The principle of WSE algorithm is given, and the performance of it for dynamical change detection is analyzed using synthetic signal. The results show that WSE can accurately capture amplitude information and amplify the detection results of dynamical changes compared with SE and PE in the case of high complexity, but it requires a higher computation cost than the others. From all the analyses in this paper, we find that WSE has a better performance for dynamical change detection compared with the other three algorithms in abnormal amplitude.

1. Introduction

In 1948, Shannon put forward the theory of information entropy for the first time, and regarded it as a standard to measure the uncertainty of the system. Since then, a variety of new entropy based on information entropy have been proposed. With the development of information entropy, Kolmogrov entropy appeared, and then approximate entropy appeared. Approximate entropy can be applied to the analysis of short-term physiological sequence[1]. Before the appearance of sample entropy, it has been widely used in clinical cardiovascular research[2-10]. However, self-matching is introduced in the calculation of approximate entropy, and illegal values are easy to appear[11]. So Richman and Moorman proposed the sample entropy algorithm. Sample entropy is a single scale entropy, which calculates the exact value of the negative average natural logarithm of the conditional probability on a single scale, does not include the comparison of its own data segments, and the sample entropy has little dependence on the data length.

However, because the sample entropy only considers the complexity of the one-dimensional time series signal, the amplitude information of the time series signal is lost, which leads to the sample entropy can’t detect the sudden change signal well. To solve this problem, the concept of weighted sample entropy is proposed in this paper. The amplitude information is extracted and added to the calculation of entropy.

The paper is organized as follows. In section 1, the research background and the problems to be solved by weighted sample entropy, are introduced. The algorithm of weighted sample entropy is introduced in section 2. In section 3, we design artificial signal for entropy algorithm and the weighted sample entropy is used to analyze the abnormal signal detecting. The concluding remarks are recorded in Section 4.
2. Methodologies

2.1 Weighted Sample Entropy

In the calculation of sample entropy, the parameter of sample entropy is the embedding dimension \( m \) and the similarity tolerance \( r \). In general, when the value of \( m \) is 1 or 2 and \( r \) is between \((0.1 \sim 0.2) \times \text{Std}\), the calculated sample entropy has better statistical characteristics. In this paper, \( m \) is taken as 2 and \( r \) as 0.2.

The general sample entropy algorithm only retains the distance relationship between different reconstruction vectors of time series, but ignores the amplitude difference between time series with the same distance relationship. Therefore, this paper proposes a weighted sample entropy, which takes the amplitude information of time series into account on the basis of general sample entropy. The general steps of calculating weighted sample entropy are as follows:

**Step 1**: A set of \( m \)-dimensional vectors is formed by taking \( m \) points in sequence from a time series \( \{x_i\}_{i=1}^N \) of length \( N \)

\[
X_m(i) = [x(i), x(i+1), \cdots, x(i+m-1)]
\]  
(1)

Where \( 1 \leq i \leq N-m+1 \), and \( m \) denote the embedding dimension.

**Step 2**: \( d[X_m(i), X_m(j)] \) is defined as maximum value of distance difference between \( X_m(i) \) and \( X_m(j) \):

\[
d[X_m(i), X_m(j)] = \max|x(i+k) - x(j+k)|
\]  
(2)

\[
D_{i,j} = \{d[X_m(i), X_m(j)]\}_{(N-m+1) \times (N-m+1)}
\]  
(3)

Where \( 0 \leq k \leq m-1, \ 1 \leq i, \ j \leq N-m+1, \ i \neq j \).

**Step 3**: For the given similarity tolerance \( r > 0 \) , excluding elements which \( d[X_m(i), X_m(j)] \geq r \) in the matrix \( D_{i,j} \), the difference matrix \( S_{i,j} \) is denoted as:

\[
S_{i,j} = \{d[X_m(i), X_m(j)]\}_{(N-m+1) \times (N-m+1)}, \quad d[X_m(i), X_m(j)] < r
\]  
(4)

Where \( 0 \leq k \leq m-1, \ 1 \leq i, \ j \leq N-m+1, \ i \neq j \).

**Step 4**: In this paper, weight \( \omega_{i,j} \) is defined as:

\[
\omega_{i,j} = \frac{\sum_{s_{i,j}: d[X_m(i), X_m(j)] < r} \text{std}(d[X_m(i), X_m(j)])}{\sum_{j=1}^{N-m+1} \text{std}(d[X_m(i), X_m(j)])}
\]  
(5)

**Step 5**: For each row, calculating the ratio of the sum of each row’s weights, \( \sum \omega_{i,j} \), to the total number of vectors \( N-m+1 \) , the ratio is denoted as \( B_m^\omega(r) \):

\[
B_m^\omega(r) = \frac{1}{N-m+1} \sum \omega_{i,j}
\]  
(6)

**Step 6**: \( B_m^\omega(r) \) can be computed as:

\[
B_m^\omega(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln B_m^\omega(r)
\]  
(7)

**Step 7**: Adding \( m \) to \( m+1 \), and repeating the steps (1)–(6), \( B_m^{m+1}(r) \) is calculated.

In theory, the weighted sample entropy should be:

\[
\text{WSampEn}(m, r) = \lim_{N \to \infty} \{-\ln[B_m^{m+1}(r) / B_m^\omega(r)]\}
\]  
(8)
When the N value is limited, the weighted sample entropy can be expressed as:

$$W\text{SampEn}(m,r,N) = -\ln\left[\frac{B_{r+1}^m(r)}{B_r^m(r)}\right]$$  \hspace{1cm} (9)$$

3. Experimental results and analysis

3.1 Synthetic signal

A synthetic signal $y(t)$ is produced using the following equation:

$$y(t) = \begin{cases} 
\sin(2\pi \cdot t) + \sin(2\pi \cdot 10 \cdot t) & 0 < t \leq 10 \\
\sin(2\pi \cdot t) + \sin(2\pi \cdot 10 \cdot t) + \sin(2\pi \cdot 20 \cdot t) & 10 < t \leq 20 \\
2 \cdot \sin(2\pi \cdot t) + 1.5 \cdot \sin(2\pi \cdot 10 \cdot t) + 1.25 \cdot \sin(2\pi \cdot 20 \cdot t) & 20 < t \leq 30 
\end{cases}$$  \hspace{1cm} (10)$$

where the sampling frequency is 100 Hz and the sampling time is 30 s. The curve of synthetic signal $y(t)$ is shown in figure.1.

As shown in the figure.2, WSE, SE, PE and WPE pairs are used to process the signal. In calculation, the sliding window with a length of 200 is used to intercept the signal, and the increment of each intercepting signal by the sliding window is 1.

It can be seen from figure 2(a) that around 10s, WPE and PE have increased significantly, while WSE and SE have changed, but they are not obvious. Around 20s, WSE and SE have a relatively large increase, but the change between WPE and PE is not obvious. This shows that WPE and PE can well detect changes in complexity caused by changes in signal frequency, but they are not sensitive to the
detection of sudden changes in signal amplitude. WSE and SE can detect the change of signal amplitude very well and are sensitive to the change of signal amplitude.

It can be seen from figure 2(b) that at about 10 s, the difference between WPE and PE has a big increase, and around 20 s, the difference between WSE and SE has a big increase. This proves that the weighting strategy proposed in this article can better improve the sensitivity of the entropy algorithm to changes in complexity. In addition, the entropy of WSE, SE, PE and WPE of the synthesized signal $y(t)$ at the trip points are shown in Table 1. It is obvious that the entropy of WSE at the trip points are the largest among the four algorithms. These two parts indicate that WSE algorithm can detect not only the change of complex system, but also the mutation of amplitude. The moment when the entropy changes suddenly is defined as the trip point. The points are shown in figure 3(a), Trip1 is equal to $Mv2$ minus $Mv1$ and Trip2 is equal to $Mv2$ minus $Mv3$, where $Mv1$, $Mv2$ and $Mv3$ represent the mean of entropy value when $t=2s~10s$, $t=12s~20s$ and $t=22s~30s$ respectively.

| synthetic signal | WSE   | SE    | WPE   | PE    |
|------------------|-------|-------|-------|-------|
| trip1            | 0.0166 | 0.0981 | 0.4477 | 0.1445 |
| trip2            | 0.4131 | 0.2938 | -0.0828 | -0.0341 |

3.2 Pulse signal detecting

On the basis of synthetic signal, pulse signal is added to test the ability of WSE algorithm to detect abnormal signal.

It can be seen from the figure 3 that WSE has a good ability to detect the mutation signal at the position where the pulse signal is added, and with the increasing of signal complexity, the ability of
WSE to detect the mutation signal is gradually enhanced, which shows that in the case of high complexity, the ability of WSE to detect the anomaly is better than other entropy.

4. Conclusion
In this paper, an improved algorithm called weighted sample entropy is proposed, which supplements the amplitude information lost in the calculation of sample entropy by weighting, which makes the sample entropy more sensitive to the sudden change of signal amplitude while measuring the overall complexity. The simulation experiment of synthetic signal shows the limitation of SE, PE and WPE, and also proves the advantage of WSE.

Acknowledgments
This research is supported by Natural Science Basic Research Plan in Shaanxi Province of China (Program No. 2017JM5065) and Graduate Education Reform Program of Xi ’an University of Science and Technology (Program No. YJG2020014)

Reference
[1] Dawes GS, Moulden M, Sheil O, Redman CW. Approximate entropy, a statistic of regularity, applied to fetal heart rate data before and during labor. Obstet Gynecol, 1992, 80(5): 763-768
[2] Fleisher LA, Dipietro JA, Johnson TR, and Pincus S. Complementary and noncoincident increases in heart rate variability and irregularity during fetal development. Clin Sci, 1997, 92(4): 345-349
[3] Ho KK, Moody GB, Peng CK, Mietius JE, Larson MG, Levy D, Goldberger AL. Predicting survival in heart failure case and control subject by use of fully automated methods for deriving nonlinear and conventional indices of heart rate dynamics. Circulation, 1997, 96(3): 842-848
[4] Hougue CJ, Domitrovich PP, Stein PK, Despotis GD, ReL, Schuessler RB, Kleiger RE, and Rottman JN. RR interval dynamics before atrial fibrillation in patients after coronary artery bypass graft surgery. Circulation, 1998, 98(5): 429-434
[5] Korpelainen JT, Sotaniemi KA, Makikallio A, Huikuri HV, and Myllyla VV. Dynamic behavior of heart rate in ischemic stroke. Stroke, 1999, 30(5): 1008-1013
[6] Makikallio TH, Ristimae T, Airaksinen KE, Peng CK, Goldberger AL, Huikuri HV. Heart rate dynamics in patients with stable angina pectoris and utility of fractal and complexity measure. Am J Cardiol, 1998, 81(1): 27-31
[7] Nelson JC, Rizwan-uddin, Griffin MP, Moorman JR. Probing the order within neonatal heart rate variability. Pediatr Res, 1998, 43(6): 823-831
[8] Palazzolo JA, Estafano FG, and Murray PA. Entropy measure of heart rate variation in conscious dogs. Am J Physiol Heart Circ Physiol, 274(H1099-H1115), 1998
[9] Schuckers SA. Use of approximate entropy measurements to classify ventricular tachycardia and fibrillation. J Electrocardiol, 1998, 31: 101-105
[10] Tulppo MP, Makikallio TH, Takala TE, Seppanen T, Huikuri HV. Quantitative beat-to-beat analysis of heart rate dynamic during exercise. Am J Physiol Heart Circ Physiol, 1996, 271(1 Pt 2): H244-H252
[11] Alcatra R, Rietta J. A novel application of sample entropy to the electrocardiogram of atrial fibrillation. Nonlinear Analysis: Real World Applications, 2010(2): 1.26-1055
[12] Richman J S, Randall M J. Physiological time-series analysis, using approximate entropy and sample entropy[J]. American Journal of Physiology-Heart and Circulatory Physiology, 2000, 278(6): H2039-H2049.