International Journal of Advanced Information Science and Technology (IJAIST) ISSN: 2319:2682
Vol.4,No.7, July 2015 DOI:10.15693/ijaist/2015.v4i7.147-151

Quantifying the Impact of Technical Artifacts on Entropy based HRV

Manjit Singh
Department of ECE
Guru Nanak Dev University, Regional Campus,
Jalandhar, Punjab, India

Butta Singh
Department of ECE
Guru Nanak Dev University, Regional Campus,
Jalandhar, Punjab, India

***Abstract*— In the measurement of electrocardiogram (ECG) associated with the heart rate, technical artifacts in the ECG recordings deteriorate the data, yielding missing RR interval tachogram. Heart rate variability (HRV) refers the instantaneous changes in beat-to-beat heart rate calculated from RR interval series extracted from the ECG. The linear parameters of HRV are very sensitive to these missing RR intervals. In this study, the effect missing RR intervals on approximate entropy (ApEn) and sample entropy (SampEn) based entropy measures of HRV is investigated, using simulated missing data in real RR interval tachograms. For the simulation, randomly selected data (0–100 RR intervals) were removed from real RR data obtained from the MIT-BIH normal sinus rhythm database. In all, 703 tachograms of 1000 RR interval data length were used for this analysis. The relative error in entropy measures increase more significantly in function of missing RR proportion. The results of the simulation revealed that entropy parameters are more robust measures than linear parameters of HRV in presence of missing RR interval data.**

Index terms - HRV, Approximate entropy, Sample entropy, Missing RR intervals

I. INTRODUCTION

Heart rate variability (HRV), the variation over time of the period between consecutive heartbeats, is predominantly dependent on the extrinsic regulation of the heart rate (HR). It reflects many physiological factors modulating the normal rhythm of heart and provides a powerful means of observing the interplay between the sympathetic and parasympathetic nervous systems [1, 2]. It may be analyzed in either the time, frequency and by using nonlinear approach. One of the traditional, simple and widely used methods to quantify the HRV is the time-domain method [1]. Time-domain HRV parameters can be computed directly by numerical approach based on tachograms of the RR-interval [1]-[3]. Mäkikallio et al. [4] suggests that the time domain measures cannot detect subtle changes in heart rate dynamics and are suitable only for long-term recordings. Since calculations are performed directly in the beat domain, re-sampling is not required for the time-domain parameters. However, HRV spectral parameters, total frequency (TF), very low frequency (VLF), low frequency (LF) and high frequency (HF), are obtained by the sum of the power in the relevant frequency range in the spectrum; this is estimated from data that are regularly re-sampled in the time domain [5]. HR is controlled by several central nervous system oscillators and different control loops. Interactions among these units may induce highly nonlinear, non-stationary, random and complex time courses in the processes, but the underlying sub-processes include well-determined behavior. Therefore, it is presumed that these irregular and complex time courses can be characterized more adequately by dynamic nonlinear analyses rather than by linear time series analyses [6]-[10]. There are strong evidences to consider the complex behavior of HRV as a nonlinear dynamic controlled by the autonomic nervous system (ANS) [2-4]. It has been shown that complexity analysis of HRV by nonlinear dynamics can significantly improve the identification of an increase in sudden cardiac death, in comparison with the conventional linear analysis in the time or frequency domain. The rationale in the emergence of nonlinear measures of HRV is that the heart is not a periodic oscillator under normal physiological conditions [8, 9] and standard linear measures of HRV are not able to detect subtle but important changes in the heart rate time series. Since the linear parameters of HRV do not provide adequate information on the complexity that lies inside beat-to-beat variability, the application of nonlinear techniques is appropriate. Approximate entropy (ApEn) and sample entropy (SampEn) are the recently developed nonlinear techniques to quantify the complexity of time series data like heart rate intervals [11]-[14].

Electrocardiogram (ECG) recordings, during patient movement, exercise stress test and surgical operations are exposed to artifacts. The reason of artifacts may be technical problems, biological events, low amplitude or abnormalities of R waves and errors in the automatic detection. As a result extra beats and missed beats can occur. These artifacts complicate the detection of QRS feature points and yield incomplete RR interval tachogram [15, 16]. Kim et al. evaluated the effect of missing RR intervals on linear (time and frequency) domain HRV parameters [15, 16]. They found that missing RR interval affect significantly both time and frequency domain HRV parameters. The relative error in time domain and frequency domain HRV parameters was found to be: mean <3%; SDSD and SDNN >10%; pNN50 >50%; VLF 0.04% for FFT; 0.12% for mFFT; 0.11% for Welch; 100.4% for Yule-Walker; 12.3 X 103% for Burg; 99.9% for Lomb, LF 7.2% for FFT; 28.8% for mFFT; 58.8% for Welch; 2.2 X 103% for Yule-Walker; 120.8% for Burg; 2.9% for Lomb, HF

36.3% for FFT; 41.0% for mFFT; 28.8% for Welch; 458.7% for Yule-Walker; 31.9% for Burg; 6.3% for Lomb. Similar study was performed by Berntson et al., to investigate the effects of the missed beats of ECGs on spectral parameters [17]. They found that even a single missed or extra beat affects the whole heart rate spectrum. In this study, the impact of consecutive missing RR interval data on entropy based HRV analysis are investigated by simulating missing data in RR interval tachograms recorded from healthy subjects.

II. DATA

In this study, 18 long-term RR tachograms belonging to the MIT-BIH normal sinus rhythm database (http://www.physionet.org/physiobank/database/nsrdb) were used. The RR tachograms were extracted from annotations in the database in which the sampling rate for recording was 128 Hz. These data were recorded over 24 h from subjects who did not exhibit any significant arrhythmias; these subjects included 5 men (age: 26–45 years) and 13 women (age: 20–50 years). In all, 703 RR-interval data sets of data length N=1000 were collected for HRV analysis; they included only normal beats. In each data set, consecutive RR interval data were randomly selected for removal, and the data length removed was increased from 0 to 100 RR intervals in an increment of 5. Therefore, the number of data sets used in these simulations was 14 060 (=703 data sets × 20 missing data). Two random selections of RR intervals were made to analyze the average effects of the missing data. In each case, the ApEn and SampEn based HRV parameters were calculated' A total of 28.12 × 103 calculations were performed using MATLAB for each entropy parameter.

III. NONLINEAR ANALYSIS OF HRV

The development of the nonlinear dynamical system analysis has led to the introduction of a large amount of signal analysis techniques aimed at the extraction of nonlinear parameters from experimental time series. The original objective was the evaluation of the generating system characteristics in order to better understand its nature. In many cases however the generation system is unknown and the output signal is the only information we can have about the system itself. This is precisely the case of the human life support systems among which the heart plays a dominant role. It has been shown that HRV signal changes can be related to the activity of several physiological control mechanisms of different nature. Their interaction produces changes in the beat rate assuring the control activity reacts efficiently to various incoming stimuli, results in the nonlinear deterministic structure in HRV time series signal.

A. Approximate entropy (ApEn)

ApEn is a statistical index to quantify the complexity of a signal. It has been widely adopted by many researchers especially in the field of heart rate variability. The popularity of approximate entropy stems from its capability to provide quantitative information about the complexity of the experimental data that are short in data length [11, 12]. ApEn measures the (logarithmic) likelihood that runs of patterns that are close for m observations remain close on next incremental comparison. Greater likelihood of remaining close, i.e., high regularity, produces smaller ApEn values [12]. While implementing ApEn, calculation requires a priori specification of two unknown parameters, the embedding dimension and *r*, a threshold, which is in effect a noise filter. Pincus, who developed the ApEn method, suggested that *r* should be 0.1 to 0.25 times, the standard deviation of the data, and that *m* be 1 or 2 for data lengths N ranging from 100 to 5,000 data points [11]. Given a signa $u(\mathbf{1}), u(\mathbf{2}), ...u(N)$, where, *N* is the total number of data points. Fix *m*, a positive integer and *r*, a positive real number. For our study we have choose *r* equal to 20% of standard deviation and $m = \mathbf{2}$. ApEn algorithm can be summarized as follows.

1. Form m vectors $X(\mathbf{1})$ to $X(N-m+\mathbf{1})$ defined by

$$X(i) = [u(i), u(i+1), ...u((i+m-1)] \quad 1 \le i \le N-m+1$$

2. Define the distance $d[X(i), X(j)]$ between the vectors $X(i)$ and $X(j)$ as the maximum absolute difference between their respective scalar components:

$$d[X(i), X(j)] = \max_{k=1,2,...m} [|u(i+k-1) - u(j+k-1)|]$$

3. Define for each *i*, for $i = \mathbf{1}, \mathbf{2}, ...N-m+\mathbf{1}$

$$C_i^m(r) = \frac{v^m(i)}{N-m+1}$$

Where $v^m(i)$ =number of $d[X(i), X(j)] \le r$

4. Take the natural logarithm of each $c_i^m(r)$ and average it over *i*

$$\phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} In(C_i^m(r))$$

5. Increase the dimension to $m+\mathbf{1}$ and repeat steps 1 to 4

6. Calculate ApEn value for a finite data length of *N*:

$$ApEn(m, r, N) = \phi^m(r) - \phi^{m+1}(r)$$

A high degree of regularity means that sequences, which are similar for m points, are likely to be similar for the next m+1 points, while this is unlikely to occur for irregular time series. Thus low values of ApEn reflect high regularity.

B. Sample entropy (SampEn)

The SampEn is a modification of ApEn. The differences with respect to ApEn are: (i) self-matches are not counted (ii) only the first $N-m$ vectors of length *m* are considered [13, 14]. SampEn algorithm can be summarized as follows.

1. Form *m* vectors $X(\mathbf{1})$ to $X(N-m+\mathbf{1})$ defined by

$$X(i) = [u(i), u(i+1), ...u((i+m-1)] \quad 1 \le i \le N-m+1$$

2. Define the distance $d[X(i), X(j)]$ between the vectors $X(i)$ and $X(j)$ as the maximum absolute difference between their respective scalar components

$$d[X(i),X(j)] = \max_{k=1,2,...m}[|u(i+k-1)-u(j+k-1)|]$$

3. Define for each *i*, for $i = \mathbf{1},\mathbf{2},...N-m+\mathbf{1}$

$$B_i^m(r) = \frac{v^m(i)}{N-m+1}$$

Where $v^m(i)$ =number of $d[X(i),X(j)] \le r \;\; i = j$

4. Define for each *i*, for $i = \mathbf{1},\mathbf{2},...N-m+\mathbf{1}$

$$A_i^m(r) = \frac{v^{m+1}(i)}{N-m+1}$$

Where $v^{m+\mathbf{1}}(i)$ =number of $d[X(i),X(j)] \le r \;\; i = j$

5. Define

$$B^m(r) = \frac{\sum_{i=1}^{N-m} B_i^m(r)}{N-m}$$

$$A^m(r) = \frac{\sum_{i=1}^{N-m} A^m{}_i(r)}{N-m}$$

6. SampEn for a finite data length of *N* can be estimated as

$$SampEn(m,r,N) = -In(\frac{A^m(r)}{B^m(r)})$$

Similarly to ApEn, we estimated SampEn with *r* equal to 20% of standard deviation and $m = \mathbf{2}$.

IV. STATISTICAL METHODS

The entropy measures of complete RR interval series were compared to the HRV parameters of RR interval series with missing data, In addition independent samples t test was used to analyze the percentage differences in nonlinear parameters with missing RR intervals. A value of P<0.05 was considered statistically significant.

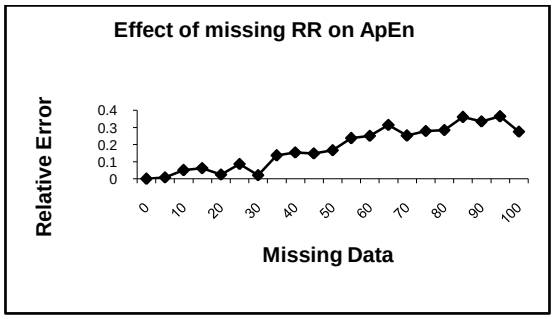
V. RESULTS AND DISCUSSION

Endogenous biomedical signals from physiological systems are acquired for a number of reasons including diagnosis, post surgical intensive care monitoring, neonatal monitoring, guide therapy and research. The problems caused by artifacts in biomedical signals are vast in scope and variety. Their potential for degrading the performance of the most sophisticated signal processing algorithms is high. Table 1 show the impact of missing RR intervals on ApEn and SampEn based HRV of 703 normal complete RR interval time series. Average SampEn was found to be always greater than average ApEn, which is agreement with wide consensus: ApEn method is inherently biased because self-matches are incorrectly counted to avoid the occurrence of natural logarithm of zero in the calculation. The effects of the missing RR interval data on the entropy based HRV measures were evaluated based on the relative errors, compared with the parameters calculated from the original, complete RR interval data. When X1, X2, . . . Xn (n = 2 in this study) is obtained for a HRV parameter of the data set with a missing duration, and Xorigin is the corresponding parameter value of that without any missing data, the relative errors REk are computed as |Xorigin − Xk|/Xorigin × 100 (%), where k = 703. For each HRV parameter and missing duration, 1406 error values were derived and used for the statistical calculations. Figure 1 shows the statistical results for the mean relative errors in each entropy parameter of HRV with an increase in the missing data duration. The entropy parameters are found to be robust to data having missing RR intervals. The mean relative errors are less than 0.364 for ApEn and less than 0.328 for SampEn even in the presence of 100 missing RR intervals. Pattern of relative error of ApEn and SampEn are almost similar. Table 2 shows the significance level (P value) to reject the alternative hypothesis that the entropy parameters of 703 subjects in presence of missing RR intervals is less than mean HRV parameters without missing RR intervals. The significance level found to be increase with increase in missing RR interval data. Table 2 also reveals that the effect of missing RR interval is not significant for entropy based HRV parameters, whereas linear parameters are very sensitive to missing RR intervals [15, 16].

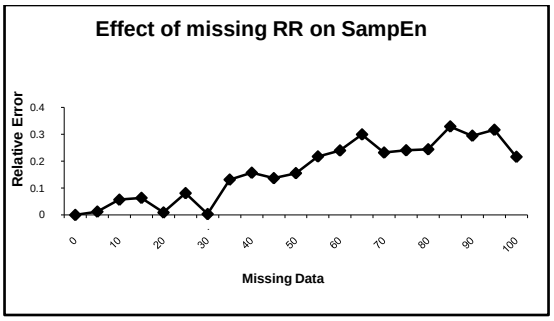
Missing RR	ApEn	SampEn
0	1.5138 ±0.086	1.5534±0.100
10	1.5130±0.085	1.5525±0.099
20	1.5134±0.086	1.5533±0.099
30	1.5135±0.084	1.5534±0.098
40	1.5114±0.084	1.5510±0.097
50	1.5113±0.083	1.5510±0.096
60	1.5100±0.084	1.5497±0.098
70	1.5100±0.084	1.5498±0.097
80	1.5095±0.083	1.5496±0.097
90	1.5087±0.084	1.5488±0.097
100	1.5096±0.083	1.5500±0.096

Table 1. Approximate entropy and sample entropy (mean ± standard deviation) of RR intervals of 703 healthy subjects with missing intervals varies from 0 to 100

Therefore in case of missing RR interval data, entropy based HRV measures are more robust and appropriate parameters than linear HRV measures. Further the correlation coefficients between increase in relative error in entropy measures and increase in amount of missing RR intervals are 0.949 for ApEn and 0.914 for SampEn. Thus, relative errors in the values of entropy measures increase significantly in function to the missing RR proportion.



a



b

Figure 1. Effect of missing RR intervals of healthy subjects on HRV parameters approximate entropy (a), sample entropy (b)

Missing RR	ApEn	SampEn
10	0.594	0.5931
20	0.5477	0.5149
30	0.5404	0.4942
40	0.77	0.7463
50	0.7899	0.7465
60	0.8834	0.843
70	0.8869	0.8376
80	0.915	0.8505
90	0.9461	0.8938
100	0.9084	0.8217

Table 2. Significance level (P value) to reject the alternative hypothesis that the nonlinear HRV parameters of 703 subjects in presence of missing RR intervals is less than mean HRV indices without missing RR intervals.

V. CONCLUSION

By maintaining steady state and isolation conditions for the prescribed duration, artifacts like missing RR intervals are unlikely to occur. But it is virtually impossible to achieve these conditions, therefore artifacts poses serious problems. If missing RR intervals occur frequently in the ECG signals, then one is not able to evaluate neuro-cardiac control through conventional linear (time and frequency-domain) HRV parameters. Present study reveals that entropy based nonlinear techniques are more suitable and reliable parameter for such type of signals. Relative errors in entropy based HRV measures due to missing RR intervals were found to be negligible. Further, error increases in function with amount of missing RR intervals.

The finding of the present study can be partly used as a reference for the acceptable amount of missing RR intervals for the ApEn and SampEn based HRV assessment. The entropy based HRV quantification is not a reliable indicator for variability signal with missing data. However this method has certainly edge over commonly used linear parameters and can be preferred for tachograms having missing RR intervals.

REFERENCES

[1]. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, “Heart Rate Variability—Standards of Measurement, Physiological Interpretation and Clinical Use,” European Heart Journal, 1996, vol. 17, pp. 354-38.

[2]. U. R. Acharya, K.P. Joseph, N. kannathal, C.M. lim, J.S. Suri, “Heart rate variability: A review,” Med Bio Eng Comp, 2006, vol. 44, pp. 1031-1051.

[3]. G. G. Berntson, J. T. Bigger, D. L. Eckberg, P. Grossman, P. G. Kaufmann, M. Malik, H. N. Nagaraja, S. W. Porges, J. P. Saul, P. H. Stone and M. W. van der Molen, “Heart rate variability: Origins, methods, and interpretive caveats,” Psychophysiology, 1997, vol. 34, pp. 623-648.

[4]. T. H. Makikallio, J. M. Tapanainen, M. P. Tulppo, H. V. Huikuri, “Clinical applicability of heart rate variability analysis by methods based on nonlinear dynamics,” Cardiac Electrophysiology Review, 2002, vol. 6, 250–255

[5]. D. Singh, K. Vinod and S.C. Saxena, “Sampling frequency of RR-interval time-series for spectral analysis of heart rate variability,” Journal of Medical Engineering and Technology, 2004, vol. 28, no. 6.

[6]. B. Singh and D. Singh, “Effect of threshold value r on multiscale entropy based heart rate variability,” Cardiovascular Engineering and Technology, 2012,vol. 3, pp. 211-216.

[7]. B. Singh and N. Bharti, “Software tools for heart rate variability analysis,” International Journal of Recent Scientific Research, 2015, vol. 6, no. 4, pp. 3501-3506.

[8]. B. Singh and D. Singh, “Ectopic beats and editing methods for Poincaré-plot-based HRV,” Int. J. Biomedical Engineering and Technology, 2011, vol. 7, no. 4, pp.353–364.

[9]. J.K. Kanters, M.V. Hojgaard, E. Agncr & N-H Holstein-Rathlou, “Short- and long-term variations in non-linear dynamics of heart rate variability,” Cardiovascular Research,, 1996, vol. 31, pp. 400-409.

[10]. M. Singh, B. Singh and V.K. Banga, “Effect of ECG sampling frequency on approximate entropy based HRV,” International. Journal of Bio-Science and Bio-Technology, 2014, vol. 6, no. 4, pp. 179-86.

[11]. S. M. Pincus, “Approximate entropy as a measure of system complexity,” Proc. Natl. Acad Sci. USA, 1991, vol. 88, pp. 2297-2301.

[12]. S.M. Pincus & A. L. Goldberger, “Physiological time-series analysis: what does regularity quantify?,” American Journal of physiology, 1994, vol. 266, pp. H1643-H1656.

[13]. J.S. Richman & J.R. Moorman, “Physiological Time Series Analysis using Approximate Entropy and Sample Entropy,” Am J Physiol Heart Circ Physiol, 2000, vol. 278, pp. 2039-2049.

[14]. D.E. Lake, J.S. Richman, M.P. Griffin & J.R. Moorman, “Sample entropy analysis of neonatal heart rate variability,” Am J Physiol Regl Integr, 2002, vol. 283, pp. R789-R797.

[15]. B. Singh and D. Singh, “Modified multiscale entropy in HRV for automatic selection of threshold value r,” International

Journal of Medical Engineering and Informatics, 2012, vol. 4, pp. 55–65.

[16]. K.K. Kim, Y.G. Lim, J.S. Kim and K.S. Park, “The effect of missing RR-interval data on heart rate variability analysis in the frequency domain,” Physiological measurement, 2009, vol 30, pp. 1039-1050.

[17]. G. G. Berntson and J. R. Stowell, "ECG artifacts and heart period variability: Don't miss a beat," Psychophysiology, 1998, vol. 35, pp. 127-132.

Authors Profile

Manjit Singh received his bachelor degree in Electronics Engineering from Baba Banda Singh College of Engineering Fatehgarh Sahib, Punjab, India in 1998, and Masters Degree in Electronics and communication engineering from National Institute of Technical Teacher Training and Research, Chandigarh, India in 2010. Currently he is pursuing his Ph.D. degree from Punjab Technical University Jalandhar, Punjab, India. He is working as Assistant Professor at Guru Nanak Dev University Regional Campus Jalandhar, Punjab, India. His research interest is in the area of biomedical signal processing.

Butta Singh received his Bachelor’s degree in Electronics and Communication Engineering from Guru Nanak Dev Engineering College, Ludhiana, Punjab, India in 2002, Master’s degree in Instrumentation and Control Engineering from Sant Longowal Institute of Engineering and Technology, Longowal, Sangrur, Punjab, India in 2005 and Ph.D. degree in Engineering from National Institute of Technology, Jalandhar, and Punjab, India. He is serving as Assistant Professor in the Department of Electronics and Communication Engineering, Guru Nanak Dev University, Regional Campus, Jalandhar, Punjab, India. His professional research interests are in signal processing**, in** particular, applied to biomedical applications. He has published over 30 research articles in internationally reputed journals and conference proceedings