A Robust Martingale Approach for Detecting Abnormalities in Human Heartbeat Rhythm

Jonathan Etumusei[a], Jorge Martinez Carracedo[a] and Sally McClean[a]

[a]*Ulster University, Jordanstown, United Kingdom*

Abstract

The analysis of electrocardiogram data is vital to the healthcare system to improve and monitor health conditions. Existing algorithms are effective in discovering abnormalities in electrocardiogram data streams but most of these approaches do not focus on the intensity and duration of these anomalies. In this paper, we propose a new method called alignment of the martingale sequence (AMS) that improves previous approaches using dynamic time warping and particle swarm optimisation to obtain the optimal parameter that maximises F1. Our proposed method can also estimate the severity and extent of an abnormal heartbeat rate. Experimental results show that the proposed approach makes some improvements over the traditional method.

Keywords

Heart rate, dynamic time warping, ECG sequence, martingales

1. Introduction

Persistent heart failure (PHF) is a dynamic, crippling condition that can lead to cardiac disorders and hospitalisation. Common symptoms of PHF include fatigue, shortness of breath and peripheral oedema. These symptoms can cause several effects to human health and disruption in daily life activities [1]. The predicaments caused due to PHF have not only negatively impacted patients and their loved ones but also the health care system and society generally. PHF diagnosis is around 2% of the general population in developed nations [2]. The British heart foundation reckons that heart failure affects around 2% of the UK population [3]. The ageing demographic, in the developed world, are more inclined to be affected by this disease. HF increases from 1% among those within the age group of 45 – 55 years old to over 5% in the

CERC 2021: Collaborative European Research Conference, September 09-10, 2021, Cork, Ireland
etumusei-j@ulster.ac.uk (J. Etumusei); j.martinez-carracedo@ulster.ac.uk (J.M. Carracedo); si.mcclean@ulster.ac.uk (S. McClean)
0000-0002-1337-1471 (J. Etumusei); 0000-0001-8017-2598 (J.M. Carracedo); 0000-0002-6871-3504 (S. McClean)

age group from 80 years old above [4].

It is assumed that heartbeat rate plays a key role in the risk of a heart attack. Heart diseases such as PHF, coronary heart disease, congenital heart disease and congestive heart failure are the main cause of mortality for men and women in many countries [5]. PHF patients will need to constantly monitor their heart rate for a sign of irregular behaviour which might be a sign of potential heart failure. There are several ways by which we can handle or manage the threat of PHF as stipulated in the guidelines from the European Society of Cardiology [6]. These procedures include:

- Monitoring of symptoms associated with PHF.
- Self-management of multiple chronic illnesses.
- Educating patients to observe their health conditions and be able to identify illness seriousness.
- Consistent exercises and physical activities.
- Consistent monitoring of the heartbeat for irregular signs.

The recent evolution of microelectronics and sensor technology has led to the development of many wireless sensor applications which can measure the heartbeat for signs of cardiovascular diseases [6]. The heart's rhythm can be measured using signals that are recorded by specialised devices to identify the normal functioning of the heart through heartbeat or heart rate. Heart rate is the number of periods the heart beats per minute while heartbeat is one complete pulsation of the heart. A normal heart rate for resting adults is within the range of 60 to 100 beats per minute [7]. A lesser heart rate suggests a more dynamic heart functionality.

Arrhythmia is an abnormal heart rate or rhythm that happens when electric impulses that originate from hearth beats do not function properly. Arrhythmia might cause concern and can be life-threatening exhibiting symptoms like shortness of breath, palpitations, fatigue, feeling dizzy, fainting. The most efficient avenue to diagnose an arrhythmia is through an electrical recording of the heart rhythm called an electrocardiogram (ECG)[8]. The use of ECG time series to identify heartbeat makes it possible to intervene in the situation of PHF. Abnormal heart rate intensity (AHI) is the extent or degree to which the heart rate becomes too slow or fast in an uncertain way. Abnormal heart rate duration (AHD) is the time or period when the heart becomes too slow or fast in an uncertain way.

An ECG is a recorded signal that can be used to check your heart's rhythm and electrical activity for the diagnosis process. ECGs are one of the primary diagnostic tests to detect cardiovascular abnormalities. Using ECG, it is also possible to estimate the dimension and position of the heart chambers to discover any form of damage in the heart [9].

An anomaly can be informally defined as anything that deviates from a normal standard.

Mathematically, an anomaly is often perceived as a point that is far away from the mean of a sequence. Anomaly detection is the discovery of irregularities that are different from the rest of the time series such as ECG data. ECG time series consist of real quasi-periodic signals and current approaches will need to learn about the relationship of the sequence before they can discover anomalies [10]. Identifying abnormalities in the ECG time series is very crucial in the medical and health area.

Anomaly detection in ECG can be very challenging as a result of heartbeat variation in patients. Consequently, some heart rate variations might not be life-threatening and this can produce misleading information that can affect the interpretation of ECG readings. This situation motivates us to propose a method that uses machine learning techniques to detect severe heart rates for quick intervention to save lives.

Recently, diverse methods have been created for analysing ECG signals. However, the complexity of these techniques has limited the performance of identifying anomalies in heart rate [11]. Most of these techniques can detect irregularities in heart rate but are also unable to completely isolate noise interference in ECG signal [12]. This situation gives rise to a false alarm rate (FAR). In this work, we are developing an algorithm that will discover abnormalities in ECG signals. Unlike most successful change detection approaches, our proposed approach can detect change intensity and duration in heart rate that occurs in ECG sequence. Our method uses the martingale frame to reduce the impact of noise interference in the ECG data set. The rationale behind our suggested technique is that it allows PHF patients and medical staff to monitor their heart rate intensity and abnormal duration to diagnose early signs of arrhythmia. Also, our method can be useful in measuring the heart rate to evaluate effort between several exercises or workout sessions.

To handle any challenges of similarity measures of ECG signal points produced by our algorithm, we implement dynamic time warping (DTW). DTW [13] is a popular technique that locates the optimal alignment between two sequences under certain conditions. The optimisation and DTW concepts are further discussed in Section 3 respectively.

To obtain the optimal parameters that improve the performance of the algorithm, we use particle swarm optimisation (PSO). PSO [14] is a stochastic optimisation technique that is motivated by the intuitive mutual(swarm) behaviour of animals such as a swarm of bees, a flock of birds and schools of fish. We use the PSO approach to identify the parameter values that maximise $F1$. The method explores or searches simultaneously through a group of individuals or particles to obtain the optimum value in a swarm whose trajectories are modified by stochastic and a deterministic component [15]. For this study, we use PSO rather than genetic algorithm (GA) [16] due to the following reasons:

- PSO can be adjusted to handle complex problems.
- PSO is computationally more efficient concerning speed and memory requirements [17].

PSO will be further explained in Section 3. In this work, we use the PSO to obtain the optimal

fitness function value using F1. We use F1 instead of accuracy as it takes into consideration both false negatives and false positives. F1 is a better metric to evaluate sequences where imbalance classes exist. We benchmark our proposed technique with traditional methods and obtain competitive prediction outcomes.

The paper structure is as follows: In section 2, we review the latest work done on identifying changes in ECG data. In Section 3, we introduce our novel approaches. In section 4, we show our experimental results and compare them with the existing Martingale algorithm. We finish the paper in section 5 discussing the results that we got and the next steps that we will take in the research.

2. Related work

In the last decades, new change detection techniques have been developed to discover transitions in a human heartbeat using ECG data. For instance, Varon et al. [18] proposed a methodology for the instinctive discovery of sleep apnea from an ECG sequence. The approach uses two novel well-known features common in heart variability analysis: standard deviation and serial correlation coefficients of the interval between heartbeats. The first feature utilises the main components of QRS complexes(the spread of impulses through the ventricles of the heart) that represent abnormalities in their structure as a result of increased sympathetic activity during sleep apnea conditions. The second novel feature captures the information distributed between the respiration system and heart rate using orthogonal subspace projections. The respiratory information is obtained using the ECG signal through three robust algorithms. The features use the radial basis function (RBF) kernel implemented as input to the least-square support machine classifier. Two independent ECG data sets which include hypopneas and apnea points were analysed. The algorithm can achieve a comparable result of 100% accuracy rate in classifying sleep apnea and also able to determine the contamination level of each ECG timing.

The rise in electronic medical observation and sensors applications such as electrocardiograms are becoming available as a result of the big data revolution. However, most of these signal recorded remains unlabelled thereby making anomaly detection challenging. This situation motivates Pereira et al. [19] to introduce an unsupervised method that uses a technique to learn about the features of the ECG sequence to discover any abnormality using numerous detection strategies. Experimental result shows that the suggested method can learn demonstrative representations of ECG time series to discover divergence with scores that outperform conventional supervised and unsupervised approaches respectively.

High false alarm rates (FAR) occur in ECG signals as a result of the inability to distinguish between actual ECG signals and ECG artefacts (electromagnetic alterations that are unrelated to cardiac impulse activities) as both signals are similar in terms of structure and frequency. These characteristics lead to a misconception of ECG readings. Sivaraks et al. [20] were motivated by this fact to propose a robust approach that can discover abnormalities while minimising FAR

in ECG data. The method design takes into consideration the cardiologist and motif identification approaches. Every step of the algorithm complies with the review of a cardiologist. The approach can make use of both single-lead and multi-lead ECGs respectively. Experimental results show that the algorithm can achieve 100% on accuracy on detection, specificity, sensitivity and positive predictive with 0% FAR. The outcome depicts that the suggested method performs better compared to conventional anomaly detection techniques.

Conventional change point detection approaches can discover changes in electromagnetic (EM) signals but are often limited by the issue of noise interference. This situation motivates Etumusei et al. [21] to propose two approaches that utilise the martingale framework to discover abnormalities in EM signals. The methods can isolate noise and makes use of cross-validation to optimise its parameters. Experimental result shows the proposed algorithm makes improvements over the previous technique within the martingale framework.

In this paper, we present a novel method based on enhancing the performance of our previous work [21] known as the moving median of the martingale sequence. The proposed method is known as the alignment of the moving median of the martingale sequence (AMS). The method improves the previous moving median of the martingale sequence (MMMS) by applying time warping and optimisation techniques to enhance performance for discovering abnormality in ECG time series. This technique is applied to analyse ECG sequences. Our unsupervised method uses previous methods such as randomised power martingale to distinguish between normal and anomalous data points by learning the ECG sequences. To make this anomalous data point outstanding, our method uses the moving median approach to isolate ECG artefacts (noise). We apply dynamic time warping (DTW) to align any displaced points and PSO to optimise the algorithm parameters. We compare the suggested approach results with the original methods called the randomised power martingale (RPM) and running average of the martingale sequence (RAS) respectively. These methods will be discussed in detail in Section 3. We have summarised the proposed system in Figure 1. The next paragraph explains the heart rate model.

3. Heart rate Model

This Section discusses the pre-processing approach used and the model for discovering anomalies in ECG sequence.

3.1. Data pre-processing

The first step to implement the proposed algorithm involves obtaining accelerometry data from a Shimmer wireless sensor platform (SWSP) [22] attached to healthy participants. The participants perform activities in different scenarios within a home environment (for more details about this data set see [23]). For each scenario, ECG and accelerometer signals were captured

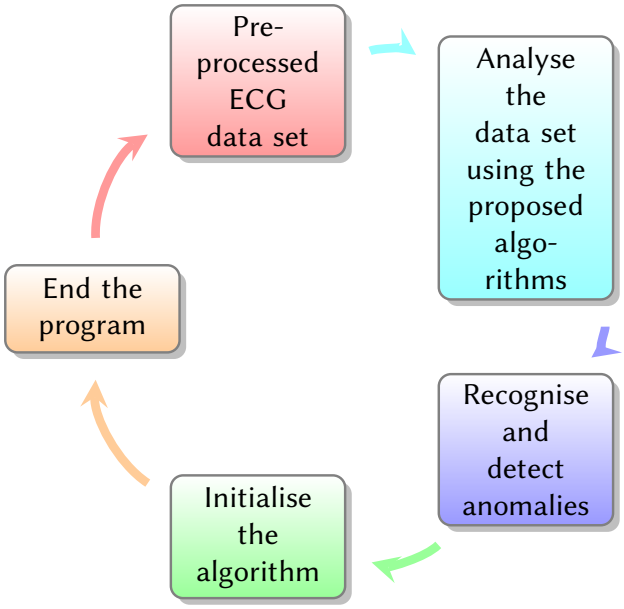

Figure 1: Proposed system

before, during and after each activity of the participants. The data obtained are streamed to a computer via the IEEE 802.15.1 Bluetooth communication protocol using the BioMOBIUS windows-based application development platform[23][24]. To process the ECG data captured, a fast Fourier transform (FFT) [25][23] is used to transform the ECG signal to determine its frequency components. Secondly, we use R-peak filtering techniques to remove and filter low-frequency noise. Finally, the average per interval heart rate for every activity is computed obtained after the filtering process. The labelled ECG data obtained from the sit to stand scenario can be seen in Figure 5 which also shows the changes, their duration and intensity.

3.2. Martingale technique

The following techniques aim to discover abnormal changes, change intensity and duration in ECG time series. In the next paragraphs, we shall focus on the martingale concept.

A martingale is a succession of a stochastic process, for which, at a specific time, the conditional expectation of the next value given all previous points is equal to the present value.

Definition 1: [26] A sequence of random variables $\{M_i : 0 \leq i < \infty\}$ is a martingale regarding the sequence of random variables $\{X_i : 0 \leq i < \infty\}$, if for all $i \geq 0$, the following conditions hold:

- The martingale M_i is a function that is measurable of $X_0, X_1, ..., X_i$,
- $E(\mid M_i \mid) < \infty$ and

- $E(M_{n+1}|X_0, ..., X_n) = M_n$.

Ho and Wechsler [26] suggested a fundamental unit of the martingale framework by defining a metric called strangeness. Strangeness measures how much a new data point diverges from the previous one in a time series.

Let us consider a sequence $Z = \{z_1, ..., z_{i-1}\}$, where there is a newly recorded point z_i. Let us also consider that the data points in Z have been clustered into k disjoint sets $Y_1, ..., Y_k, (k \leq i-1)$ [27].

Definition 2: The strangeness of z_i is defined as

$$s_i = s(Z, z_i) = \| z_i - C_r \| , \tag{1}$$

where C_r is the centroid of the cluster Y_r, for some $r \in \{1, ..., k\}$ such that $z_i \in Y_r$. $\| . \|$ denotes the chosen distance.

The strangeness of z_i is used to compute a "probability" time series where its points are named $\hat{p}_i$. If for $j = 1, 2, ..., i$, s_j is the strangeness of z_j and θ_i is a fixed value in $[0, 1]$ [26][28], $\hat{p}_i$ is computed as follows:

$$\hat{p}_i(Z \cup z_i, \theta_i) = \frac{\sharp\{j : s_j > s_i\} + \theta_i \sharp\{j : s_j = s_i\}}{i}. \tag{2}$$

Intuitively, $\hat{p}_i$ measures the probability of being more estrange than z_i. It should be noted that $\hat{p}_i$ can be seen as an unusual case of the statistical notion of p-value[26]. The set of $\hat{p}_i$ can be used to compute a new random variable that will create a new sequence known as the randomised power martingale.

Definition 3: [26] The randomised power martingale (RPM) is indexed by $\epsilon \in [0, 1]$ defined at each time-point as

$$M_n^{(\epsilon)} = \prod_{i=1}^{n} (\epsilon \hat{p}_i^{\epsilon-1}). \tag{3}$$

Fixed an $\epsilon \in [0, 1]$, once we have computed $M_n^{(\epsilon)}$, the model will detect a change in the $n-th$ timepoint if

$$M_n^{(\epsilon)} > t, \tag{4}$$

where the threshold t is chosen in a probabilistic way based on Dobb's Inequality [26].

In the following section, we introduce an approach that aims to improve the overall performance of the previously described martingale approach.

3.3. Moving Median of a martingale sequence (MMMS)

A moving median (MM) is a robust and effective smoothing technique to detect a transition in a data stream [29]. The moving median computes the median of a sequence using a sliding window. For our method, once we have computed the martingale sequence, we apply MM over that sequence to smooth the martingale points. We use median, rather than a mean for the analysis of ECG time series because it is more robust against extreme values as it is not determined by the individual points of the ECG sequence, but only by their order. This behaviour suggests that the median tends to smooth the time series, thereby, reducing the effect of noise.

Let us consider a martingale succession $\mathcal{M} = \left\{ M_i \, : \, 0 \leq i < \infty \right\}$ and fix a window length $l > 0$. We define D_k as the MM of the $k - th$ window in martingale sequence $\mathcal{M}$. Although we could use this technique to look for anomalies in the data, we will refine the sequence D_k before using it for that purpose. In the following Section, we will describe dynamic time warping.

3.4. Moving Average of the Martingale Sequence (MAS)

For our baseline method, we implement the moving average technique on the martingale sequence. The moving average computes the mean of a sequence using a sliding window. The moving average is given as;

$$MA_n = \frac{M_{n-k+1} + M_{n-k+2} + ... + M_n}{k}, \tag{5}$$

$$MA_n = \frac{1}{k} \sum_{i=n-k+1}^{n} M_i, \tag{6}$$

where M_i is the martingale point and n = is the length of the data sequence. In a later stage, we will apply PSO to the MA_n sequence to obtain the optimal parameter that maximises F1.

3.5. Dynamic time warping (DTW)

Heart rate undergoes minor changes between successive heartbeats and thereby produces a linear functional to ECG readings. One way of handling this challenge is to use dynamic time warping to specify a nonlinear heart rhythm [30][31]. Dynamic time warping (DTW) measures the affinity between the original ECG data set and our proposed algorithm output. DTW discovers the minimum path by producing a non-linear alignment between the two sequences[32]. DTW computes the optimal match between the two sequences with certain rules and conditions:

- Each one of the indexes from the initial sequence must correspond with one or more indices of the other succession and vice versa
- The first index of the initial time series must match with the first index of the other sequence
- The last index from the first time series must be similar to the last index of the other succession
- The aligning of the indices from the first sequence to indices from the other sequence must always be increasing and not becoming constant or decreasing vice versa. For instance, if $j > i$ are indices obtained from the first time-series then there should never be any two indices $l > N$ in such as way that i corresponds with index l and index j is matched to index N and vice versa

Given two sequences X and Y, we will say that each tuple (i,j) is the alignment between $X[i]$ and $Y[j]$. We define the mapping path DS as the map that minimises the distance between the sequences X and Y. We implement DTW on $MMMS$ and the original ECG sequence to obtain a new succession known as the aligned moving median of the martingale sequence (AMS). The process is repeated using the MAS and the original ECG sequence to obtain a new sequence known as the aligned moving average of the martingale sequence ($AMAS$) respectively. In the next Section, we discuss the implementation of the PSO on our new sequences.

3.6. PSO optimisation

PSO algorithm is an optimisation technique that uses a search process based on swarm exploration. In this type of exploration, each individual retains the optimal location in the swarm. For each generation, the information accumulated by the particle is then used to adapt the new location of the particle. The particles are constantly evolving in a multi-dimensional search capacity until an optimal condition is found. Each particle adjusts its position depending on its present velocity, its preceding best location (P_{best}) and the global best location (G_{best}) of the whole swarm.

To tune and explore the direction of the swamp, the velocity and location of the particles at iteration k are accomplished using the following steps:

1. Particles are initialised with arbitrary location and velocities according to the search range or space
2. Estimation of each particle using the fitness function
3. Particle are updated with individual best and global best fitness values and their location
4. The candidate solution's location and velocity are renewed
5. If the convergence criterion is satisfied then the algorithm is halted and the final solution output is presented otherwise the process will progress to step 2

Table 1
PSO component values

Parameters	value
InertiaRange	**[0.10000, 1.1000]**
InitialSwarmSpan	**200**
MaxIterations	**200 * NumberOfVariables**
MaxStallIterations	**20**
MinNeighboursFraction	**0.250**
SwarmSize:	**100**
SelfAdjustmentWeight	**1.4900**
SocialAdjustmentWeights	**1.4900**

PSO locates the optimal parameters (ϵ, window size) using fitness function (FF) to maximize F1. The fitness function is given as:

$$F1_{max} = max_{(\epsilon, windowsize)}(F1_{(AMS)}), \tag{7}$$

where ϵ ranges from 0 to 1 and window size from 2 to 20 for each activity.

PSO components implemented to maximise the fitness function are shown in Table 1. PSO implementations on the methods RPM, AMS, AMAS will be called RPM(PSO), AMS(PSO) and AMAS(POS) respectively. Furthermore, the proposed approach is illustrated in Figure 2.

3.7. Threshold computation

While Ho and Weschler [26] proposed a probabilistic way of computing threshold, we suggest a threshold based on the median absolute deviation (MAD) of the martingale sequence. MAD is a robust technique for analysing ECG time series because it measures the variability of the univariate ECG points. For a univariate sequence, $S = \{S_1, S_2, ..., S_n\}$ (in our case, we use the new PSO sequences) MAD is the median of the absolute deviations of the sequence. MAD is given as follows:

$$MAD(S) = median(\{|S_i - \tilde{S}|, i = 1, ...n\}), \tag{8}$$

where $\tilde{S} = median(S)$. MAD shows how spread out the data is. Ley et al. [33] proposed, based on outlier detection, a threshold for change detection of $ME \pm MeAD$, where ME is the mean of the data points and $MeAD$ is the mean absolute deviation. We used a similar approach using the median to compute a threshold t for our methods. Therefore, this model will detect a change when :

$$H_k \geq t. \tag{9}$$

Data: Input (F): ECG univariate data set
Result: Output: AMS(PSO) points
Initialise: $M(0) = 1$; $i = 1$; $F = \{\}$;
Set values for cluster group k, ϵ value, window size;
while do
  A new example of normalised z_i is discovered;
  if $F = [\,]$ **then**
    Set the strangeness of $z_i := 0$
  else
    Compute the strangeness of z_i and the data points in F
    Compute the $\hat{p}_i$ of z_i;
    Compute the RPM points using equation (4);
    Compute the MMMS points;
    Compute the AMS points;
    Compute the AMS (PSO);
    Compute the threshold t
  end
  if $MMS \geq t$ **then**
    Discover abnormalities
    Estimate the AHI
    Estimate the AHD
    Re-initiate $M_i = 1$
  else
    Add z_i into F ;
  end
  if $i = i + 1$; **then**
  end
end

Figure 2: The algorithm

where H_k can represent RPM(PSO), AMS(PSO) and AMAS(PSO). If H_k exceeds the given threshold t, a change has been detected. When the analysis of data of D_k is finalised, the algorithm is restarted.

The evaluation performance for the approaches is measured using robust evaluation metrics (EM)[34] such as accuracy, precision, recall(sensitivity), harmonic mean ($F1$).

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} * 100, \qquad Precision = \frac{TP}{TP + FP} * 100,$$

$$Recall = \frac{TP}{TP + FN} * 100, \qquad F1score = \frac{2 * Recall * Precision}{Recall + Precision} * 100,$$

$$Specificity = \frac{TN}{TN + FP} * 100,$$

where TP, TN, FP, FN are true positive, true negative, false positive and false negative respectively. The performance metric provides a proper estimation of the suggested method, especially on imbalanced time series [35].

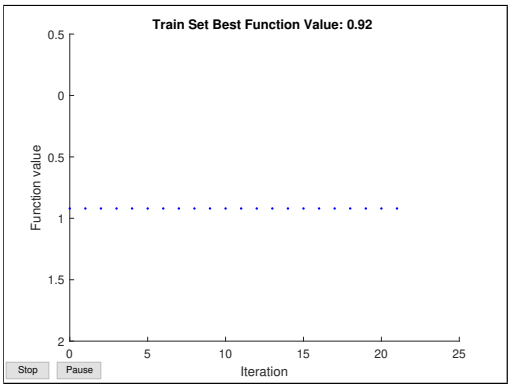

Figure 3: PSO iteration using AMS method

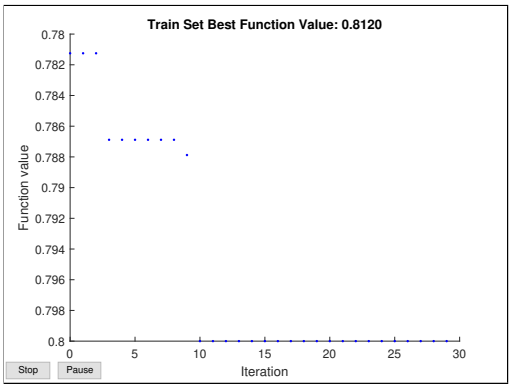

Figure 4: PSO iteration using AMAS method

Table 2
Confusion matrix using optimised parameters for training set

Approach	Training set	ϵ	window size	TP	TN	FP	FN
RPM(PSO)	N1	0.9	-	10.0(63.0%)	35.0(83.3%)	17.0(63.0%)	7.0(16.7%)
AMAS(PSO)	N1	0.9285	19.0	26.0(96.3%)	30.0(71.4%)	1.0(3.7%)	12(28.6%)
AMS(PSO)	N1	0.5741	19.0	23.0(85.2%)	42.0(100.0%)	04(14.8%)	0.0(0.0%)

3.8. Abnormal heart rate intensity and duration

Our proposed method can measure the abnormal heart rate intensity (AHI) and abnormal heart rate duration (AHD). To compute the AHI and AHD, we first find the threshold and then subtract it from the highest algorithm output point. The HAI is given as:

$$AHI = M - T, \tag{10}$$

where $M = max\{H_k \mid k = 1, ..., n\}$, being H_k the output time series of the used algorithm and T as the used threshold. Furthermore, we can also compute the duration of the abnormality in heart rate data by computing the time (sec) of the changes. AHD is the length of time in seconds, the changes take place. AHD is given as:

$$AHD = Seconds\,(TP), \tag{11}$$

In the next Section, we shall discuss the experimental results of our proposed algorithm.

4. Experimental results

The section exposes an overview of the different approaches adopted to detect anomalies in the ECG data set. The ECG data used for this experiment was discussed in Section 3.

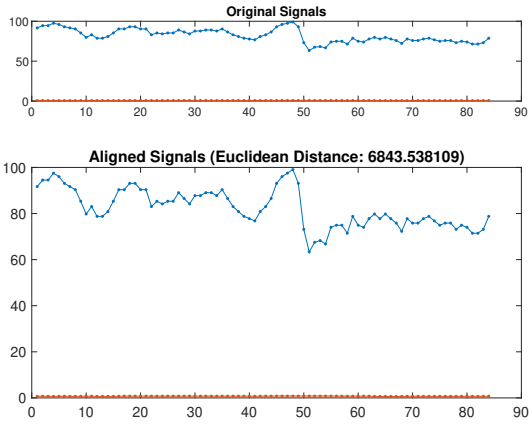

Figure 5: Original and aligned Signals

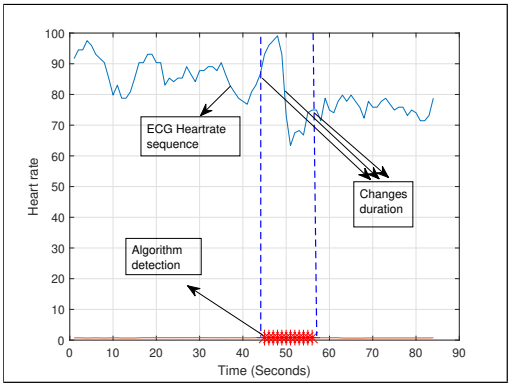

Figure 6: Test data change detection

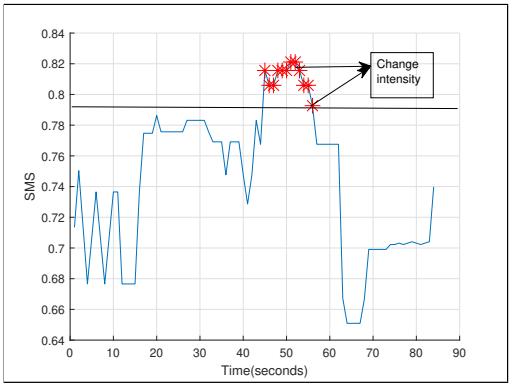

Figure 7: AMS(PSO) output

Table 3
Confusion matrix using optimised parameters for test set

Approach	Test set	ϵ	window size	TP	TN	FP	FN
RPM(PSO)	N2	0.9	-	05(45.5%)	62(84.9%)	06(54.5%)	11(15.1%)
AMAS(PSO)	N2	0.9285	19.0	11(100.0%)	53(72.6%)	00(0.0%)	20(27.4%)
AMS(PSO)	N2	0.5741	19.0	11(100.0%)	72(98.6%)	00(0.0%)	01(1.4%)

4.1. Cross validation technique

To show the rational potential, we evaluate the proposed method using different training ($N1$) and test set ($N2$). Both datasets are obtained from an individual performing similar activities. The dataset captured the participant heart rates. Our objective is to detect when there is a change in these heart rates and the actual duration of this change.

We implement our proposed method to obtain the optimal parameters. In this specific work, we used ϵ and the window size as the parameters to be optimised. This technique is used to

Table 4
Evaluation metrics of the proposed and previous approaches

Approaches	N1	N2	ϵ	window size	Sensitivity(%)	Specificity(%)	Accuracy(%)	Precision(%)	F1-Score(%)
RPM(PSO)	70	84	0.9	-	41.3[37.04, 45.5]	84.1[83.3, 84.9]	72.5[65.2, 79.8]	45.1[58.8, 31.3]	41.3[45.5, 37.0]
AMAS(PSO)	70	84	0.9285	19.0	97.7[96.3, 99.0]	72.0[71.4, 72.6]	78.7[81.2, 76.2]	52.0[68.4, 35.5]	66.0[80.0, 52.0]
AMS(PSO)	70	84	0.5741	19.0	92.1[85.2, 99.0]	98.8[99.0, 98.6]	96.6[94.2, 98.9]	95.4[99.0, 91.7]	93.9[92.0, 95.7]

Table 5
Estimation of algorithm performance

Approach	Ave. iteration time	AHI	AHD
RPM(PSO)	0.3801	0.2716 RPM	5(sec)
AMAS(PSO)	0.5236	0.0349 RAS	11(sec)
AMS(PSO)	0.4329	1.5275 AMS	11(sec)

Table 6
Evaluation performance comparison

Approach	Accuracy(%)	Sensitivity(%)	Specificity(%)	Precision(%)	F1(%)
RPM(PSO)	72.5	41.3	84.1	45.1	41.3
AMAS(PSO)	78.7	97.1	72.0	52.0	66.0
AMS(PSO)	96.6	92.1	98.8	95.4	93.9

analyse both the previous and newly proposed algorithms, that is RPM, AMS, AMAS.

Once we found optimal parameters for our data set N1, we used that configuration to compute the evaluation metrics for N2. Both confusion matrices can be found in TABLE 2 and 3. In addition, the information from the confusion matrices is then used to compute the performance metrics. This information can be seen in TABLE 4.

4.2. Performance of ECG detection algorithm

Our proposed algorithm (TABLE 4) produces better results compared to conventional approaches such as RPM(PSO) and AMAS(PSO) respectively. These comparison are summarised (TABLE 6) for better evaluation. The performance comparison (TABLE 6) shows that the suggested approach gives an accuracy rate of over 15% compared that of AMAS(PSO) and RPM(PSO) independently. Also, our proposed approach gives a specificity of over 10% compared to the AMAS(PSO) and RPM(PSO) independently. Overall our suggested technique produces a preferable output of over 40% in terms of precision and F1-score. However, the AMAS(PSO) method is slightly sensitive compared to our proposed approach. This can be attributed to slightly higher TP detected by the AMAS(PSO). This might not be a major issue at the moment, but we aim to address it further in future work. Our proposed algorithm (TABLE 5) also produces a better AHI (1.5275 AMS) and a lower average iteration run time of 0.4329 seconds. The AHD time for our proposed AMS method is 11 seconds greater than that of the RPM method.

5. Conclusion and future work

This paper discusses persistent heart failure and its impacts on humanity. This predicament inspires us to introduce a method (that uses the martingale framework) to detect abnormality in heart rate measuring devices such as ECG. Experimental results show that our proposed technique outperforms previous martingale approaches. Furthermore, the proposed algorithm can measure the heart rate intensity and duration of abnormality in ECG sequence. Future work is required to confirm this hypothesis using big data streams in specific populations areas and age categories especially the elderly who are more prone to heart failure and attack.

Acknowledgments

This work was supported by a University of Ulster Vice-Chancellor's Research Studentship. The authors would like to thank anonymous reviewers for their constructive suggestions.

References

[1] N. Albert, K. Trochelman, J. Li, S. Lin, Signs and symptoms of heart failure: are you asking the right questions?, American Journal of Critical Care 19 (2010) 443–452.
[2] A. Bundkirchen, R. H. Schwinger, Epidemiology and economic burden of chronic heart failure, European Heart Journal Supplements 6 (2004) D57–D60.
[3] K. Sutherland, Bridging the quality gap: heart failure, Health Foundation, 2010.
[4] A. J. S. Coats, Ageing, demographics, and heart failure, European Heart Journal Supplements 21 (2019) L4–L7.
[5] W. M. Jubadi, S. F. A. M. Sahak, Heartbeat monitoring alert via sms, in: 2009 IEEE Symposium on Industrial Electronics & Applications, volume 1, IEEE, 2009, pp. 1–5.
[6] K. Swedberg, J. Cleland, H. Dargie, H. Drexler, F. Follath, M. Komajda, L. Tavazzi, O. A. Smiseth, A. Gavazzi, A. Haverich, et al., Guidelines for the diagnosis and treatment of chronic heart failure: executive summary (update 2005) the task force for the diagnosis and treatment of chronic heart failure of the european society of cardiology, European heart journal 26 (2005) 1115–1140.
[7] Y. Ostchega, Resting pulse rate reference data for children, adolescents, and adults: United States, 1999-2008, 41, US Department of Health and Human Services, Centers for Disease Control and ..., 2011.
[8] R. D. White, G. Flaker, Smartphone-based arrhythmia detection: Should we encourage patients to use the ecg in their pocket?, Journal of atrial fibrillation 9 (2017).
[9] S. Chauhan, L. Vig, Anomaly detection in ecg time signals via deep long short-term memory networks, in: 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA), IEEE, 2015, pp. 1–7.
[10] M. Thill, S. Däubener, W. Konen, T. Bäck, P. Barancikova, M. Holena, T. Horvat, M. Pleva,

R. Rosa, Anomaly detection in electrocardiogram readings with stacked lstm networks, in: Proceedings of the 19th Conference Information Technologies-Applications and Theory (ITAT 2019), CEUR-WS, 2019, pp. 17–25.
[11] M. I. Owis, A. H. Abou-Zied, A.-B. Youssef, Y. M. Kadah, Study of features based on nonlinear dynamical modeling in ecg arrhythmia detection and classification, IEEE transactions on Biomedical Engineering 49 (2002) 733–736.
[12] G. D. Clifford, F. Azuaje, P. Mcsharry, Ecg statistics, noise, artifacts, and missing data, Advanced methods and tools for ECG data analysis 6 (2006) 18.
[13] D. J. Berndt, J. Clifford, Using dynamic time warping to find patterns in time series., in: KDD workshop, volume 10, Seattle, WA, USA:, 1994, pp. 359–370.
[14] D. Wang, D. Tan, L. Liu, Particle swarm optimization algorithm: an overview, Soft Computing 22 (2018) 387–408.
[15] H. Lu, J. Chen, L. Guo, Energy quality management (2018).
[16] S. Shabir, R. Singla, A comparative study of genetic algorithm and the particle swarm optimization, Int. J. Electr. Eng 9 (2016) 215–223.
[17] C. M. Martínez, D. Cao, Integrated energy management for electrified vehicles, Ihorizon-Enabled Energy Management for Electrified Vehicles (2019) 15–75.
[18] C. Varon, A. Caicedo, D. Testelmans, B. Buyse, S. Van Huffel, A novel algorithm for the automatic detection of sleep apnea from single-lead ecg, IEEE Transactions on Biomedical Engineering 62 (2015) 2269–2278.
[19] J. Pereira, M. Silveira, Unsupervised representation learning and anomaly detection in ecg sequences, International Journal of Data Mining and Bioinformatics 22 (2019) 389–407.
[20] H. Sivaraks, C. A. Ratanamahatana, Robust and accurate anomaly detection in ecg artifacts using time series motif discovery, Computational and mathematical methods in medicine 2015 (2015).
[21] J. Etumusei, J. M. Carracedo, S. McClean, Novel martingale approaches for change point detection, in: International Conference on Intelligent Systems Design and Applications, Springer, 2020, pp. 457–467.
[22] A. Burns, B. R. Greene, M. J. McGrath, T. J. O'Shea, B. Kuris, S. M. Ayer, F. Stroiescu, V. Cionca, Shimmer™–a wireless sensor platform for noninvasive biomedical research, IEEE Sensors Journal 10 (2010) 1527–1534.
[23] S. Zhang, L. Galway, S. McClean, B. Scotney, D. Finlay, C. D. Nugent, Deriving relationships between physiological change and activities of daily living using wearable sensors, in: International Conference on Sensor Systems and Software, Springer, 2010, pp. 235–250.
[24] M. J. McGrath, T. J. Dishongh, A common personal health research platform-shmmer and biomobius., Intel technology journal 13 (2009).
[25] M. Garrido, M. L. López-Vallejo, S.-G. Chen, Guest editorial: Special section on fast fourier transform (fft) hardware implementations, Journal of Signal Processing Systems 90 (2018) 1581–1582.
[26] S.-S. Ho, H. Wechsler, A martingale framework for detecting changes in data streams by testing exchangeability, IEEE transactions on pattern analysis and machine intelligence 32 (2010) 2113–2127.
[27] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, A. Y. Wu, An

efficient k-means clustering algorithm: Analysis and implementation, IEEE transactions on pattern analysis and machine intelligence 24 (2002) 881–892.

[28] V. Fedorova, A. Gammerman, I. Nouretdinov, V. Vovk, Plug-in martingales for testing exchangeability on-line, arXiv preprint arXiv:1204.3251 (2012).

[29] A. Kawala-Sterniuk, M. Podpora, M. Pelc, M. Blaszczyszyn, E. J. Gorzelanczyk, R. Martinek, S. Ozana, Comparison of smoothing filters in analysis of eeg data for the medical diagnostics purposes, Sensors 20 (2020) 807.

[30] B. Vahabzadeh, R. Sameni, The notion of cardiac phase and its applications in electrophysiological studies, Biomedical Engineering 9 (2012).

[31] R. Sameni, C. Jutten, M. B. Shamsollahi, Multichannel electrocardiogram decomposition using periodic component analysis, IEEE transactions on biomedical engineering 55 (2008) 1935–1940.

[32] M. Müller, Dynamic time warping, Information retrieval for music and motion (2007) 69–84.

[33] C. Leys, C. Ley, O. Klein, P. Bernard, L. Licata, Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median, Journal of Experimental Social Psychology 49 (2013) 764–766.

[34] C. Goutte, E. Gaussier, A probabilistic interpretation of precision, recall and f-score, with implication for evaluation, in: European Conference on Information Retrieval, Springer, Meylan, France, 2005, pp. 345–359.

[35] M. S. Orendurff, J. A. Schoen, G. C. Bernatz, A. D. Segal, G. K. Klute, How humans walk: bout duration, steps per bout, and rest duration., Journal of Rehabilitation Research & Development 45 (2008).

PaperID : paper7

Title :
A Robust Martingale Approach for Detecting Abnormalities in Human Heartbeat Rhythm

Authors :
Jonathan Etumusei, Jorge Martinez Carracedo, Sally McClean

Pages in Proceedings :
99 - 115

created 04 Dezember 2022 @ 23:31