



Research article

Anomaly detection in ECG based on trend symbolic aggregate approximation

Chunkai Zhang*, Yingyang Chen, Ao Yin and Xuan Wang

Department of Computer Science and Technology, Harbin Institute of Technology, Shenzhen, China

* **Correspondence:** Email: ckzhang812@gmail.com.

Abstract: ECG anomaly detection is a necessary approach to detect disease Electrocardiography(ECG) signals before the detail diagnosis process in medical field to gauge the health of the human heart. Nowadays, there are many anomaly detection methods for ECG detection including supervised learning and unsupervised learning. For supervised learning, it requires the knowledge of expert and different types of Arrhythmia data for training. However, since the anomalies are less and unknown in many cases which are difficult to distinguish and be labeled, unsupervised methods are more suitable to detect the ECG anomalies. Furthermore, the existing unsupervised learning studies do not take ECG shape into account where different diseases have different shapes. In this paper, a novel simple trend aggregate approximation method is proposed, the relative binary trend representation are used to record the shape feature in original time series and to detect the anomaly heart signals by similarity comparison. We use the ECG dataset in UCR Time Series Classification Archive to obtain ECG time series data and the experiment results are assessed by means of sensitivity, specificity, false alarm rate measures which is robust and promising with high accuracy.

Keywords: similarity measurement; trend distance, electrocardiography; anomaly detection; binary String

1. Introduction

Anomaly detection is to find different patterns in data which are not due to random deviations and it is widely used in fields in network intrusion detection [1, 2, 3], fraud detection [4, 5, 6], data leakage prevention (DLP) [7], disaster warning [8, 9, 10] and so on. Especially in medical field [11, 12, 13], anomaly detection for ECG stream has been used to detect any time periods of unusual beats, which is a necessary approach and standard to detect disease ECG signals before diagnosis process. Besides, there are many types of abnormal ECG signal according to AAMI, and the patients without expert knowledge are more concerned about whether the ECG is anomaly or not. Therefore, ECG anomaly

detection has increasingly become a popular task among researchers and practitioners. Observing and analyzing these ECG changes is the basis for correctly diagnosing abnormal heart rate [14] in ECG sequences. A typical ECG wave of a normal heartbeat consists of a P wave, a QRS complex and a T wave, as shown in Figure 1. From the perspective of analyzing the ECG data, the QRS complex contains the most important feature information, with the highest deflection and the shortest duration, which can provide a lot information for heart detection.

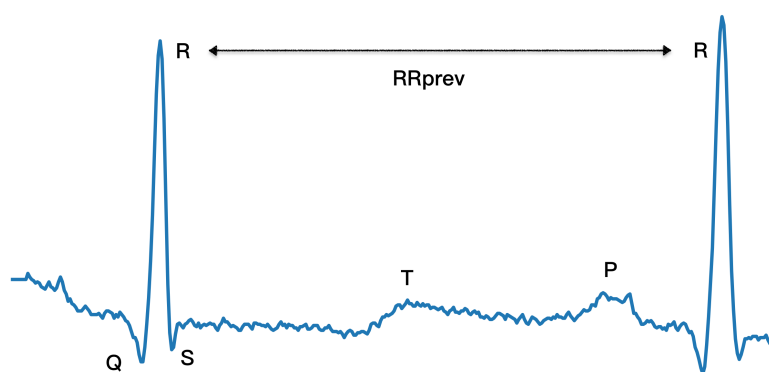


Figure 1. ECG segment by MIT-BIH 100 with two R waves.

ECG anomaly detection are divided into supervised and unsupervised methods. As for supervised learning, the existing methods [15, 16, 17] always use normal data to train the model, whereas, the accuracy is extremely depended on the precision of the features extraction and needs a lot of prior knowledge and labels. Compared to the supervised method, unsupervised anomaly detection [18, 19, 20, 21] is more suitable to ECG detection for considering only the internal structure of the data set. The work PLR-DTW [22] uses piecewise linear representation to keep important information of an ECG signal segment while using dynamic time warping to calculate the similarity measure between two signal segments. PLR transforms time series to a continuous, end-to-end line segment to approximate the original sequence, while DTW calculates the similarity between two time series by extending and shortening the time series. However, at coarser scales, details are lost in noise with PLR. Huorong Ren *et al.* [23] use piecewise aggregate pattern representations (PAPR) based on PAA, and this method divides the original ECG into several regions with equal probability and counts statistics including number, mean, variance in each region to form a matrix. HOTSAX [24] finds the distance of non-self matches to its nearest neighbor and the drawback are obvious that it involves an additional parameter and needs to be set carefully. As for particular ECG feature, J.L. Rodriguez-Sotelo *et al.* [25] preprocess and segment the ECG based on calculation of QRS complex to find the abnormal Holter Recordings which is strict to the state of feature selection. Takuya *et al.* [26] propose a "mother signal" which is the average of normal subsequences of one period length to speed up the process of anomaly detection.

The representation of time series is of great interest which determines whether there is missing information before the anomaly detection, and various approaches for ECG anomaly representation have been proposed. To address the aggregate approximation, PAA [27] divides the original time series with equal-size sliding window and counts the mean of each segment. However, this method is sensitive to the length of sliding window. If window size is larger, this sequence will be roughly regarded as mean value and all fluctuations in this sequence will be smoothed out. In order to figure out this problem, Burcu Kulahcioglu *et al.* [28] examines the suitability of symbolic PAA analysis with minimum and

maximum values to periodic signals. However, it is easy to be affected by noise. Symbolic Aggregate Approximation(SAX), proposed by Lin and Keogh [29], is a method to symbolize the time series efficiently and accurately. However it also has some disadvantages such as the dimensionality reduction, which may miss important patterns in some data sets. Battuguldur [30] propose ESAX, which uses additional symbols with two points max and min to improve representation preciseness. Youqiang Sun *at al.* [31] improved the SAX representation by using the difference between starting /ending points and average of each segment, which proposed an approach SAX_TD. However, even if SAX_TD considered the values at the beginning and the end of each subsequence to represent the trend, if the length of each segment is very long, as for ECG datasets, it is possible that their values at both ends are always similar, then adding these information is useless.

As can be seen from the method mentioned above, existing methods may distort the ECG morphology after suppression, which the important limitation is that all above these methods ignore the shape and fluctuation of the ECG signals. As we can see in the Figure 2, normal and anomaly ECG data have apparently different shape in square A and square B when time series are divided into subsequence. And in square C, we can see that once the series is replaced by the mean value, many small fluctuations in ECG data will be flattened. What's more, according to piecewise aggregate, their aggregation value are similar while the trend is apparently opposite. To address this problems, in this paper, we present an unsupervised methods to detect the anomaly ECG signals via symbol trend approximation. This method does not require prior knowledge of anomalies to work and the accuracy is high. We segment the ECG time series with sliding windows and transform the subsequence shape into binary string according to relative mean value, which will be compared the trend similarity to find out the anomalies.

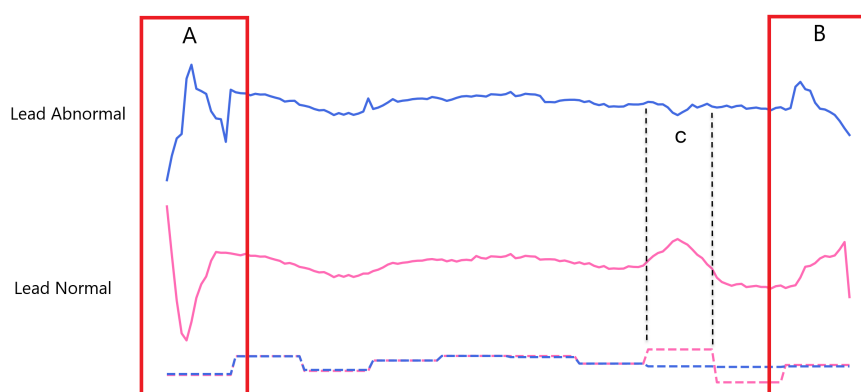


Figure 2. An example of single-lead normal and anomaly ECG data, the part circled by the red frame is the different shape in normal lead and abnormal lead while the mean value in these segment are the same. Some fluctuation in the two subsequences lead corresponding to the black dashed line(c) are smoothed out either.

The remainder of the paper is organized as follows: Section 2 provides the related work of unsupervised methods of anomaly detection in ECG. Section 3 presents our proposed method and explains the trend representation. Section 4 presents the experimental results of anomaly detection on MIT-BIH data sets. Finally, section 5 concludes the paper.

2. Model and method

In this section, we first provide a trend representation and then present the similarity calculation method and anomaly detection method. As we reviewed above, the shape of ECG beats within the same class are generally similar to each other along time axis, however, the shape are quite dissimilar between different classes. Therefore, if the shape of beat is different with higher anomaly score, the beat can be suspect as anomaly. To achieve higher performance, we use the symbolization of the trend to represent the shape of beat in our methods, which can quickly locate anomalies.

2.1. The trend representation

Given the time series $Q = \{q_1, \dots, q_n\}$, z-normalization as Equation(2.1) can remove the unit limit of the data and convert it to a dimensionless pure value, so the indicators of subsequent trend units, which will be describe in detail later, can be weighted. What's more, z-normalized time series values follows a normal distribution, which is easy to use the look-up table method to determine the linear coordinates under the normal curve and to divide the region under the Gaussian curve. In Equation(2.1) the μ and σ are the statistics of entire sequence. Then it will be reduced into equal-size segments by sliding window without overlapping, and the mean value of i th segment will be calculated as Equation(2.2):

$$q'_i = \frac{q_i - \mu}{\sigma}, \text{ where } i \in n \quad (2.1)$$

$$\bar{q}_j = \frac{w}{n} \sum_{i=\frac{n}{w}(j-1)+1}^{\frac{n}{w}j} q_i \quad (2.2)$$

where w represents as the length of segment($w \leq n$). And the new time series will be presented as $\bar{Q} = \{\bar{q}_1, \bar{q}_2, \dots, \bar{q}_w\}$.

The trend of each segment is based on binary string, which can roughly but efficiently reflect the relative trend change to mean value in each segment. We can use binary string $B = \{0, 1\}^n$ to represent the trend relative to the mean and the bits are defined as follows:

$$b_j = \begin{cases} 1, & q_j \geq \bar{q}_j \\ 0, & q_j < \bar{q}_j \end{cases} \quad (2.3)$$

in Equation(2.3), each raw data point segment is represented as $\mathbf{1}$ when the raw data is greater than the mean value of i th segment, otherwise, if the raw data is less than the mean, it is represented as $\mathbf{0}$.

We do not use the absolute trend distance because the mean of the entire time series can not record the fluctuations below the mean and above the mean. Once the time series does not fluctuate above and below the baseline of the mean, this addition will have no effect. If we record the relative trend of each segment, although the benchmark of mean value in segment is not the same in one time series, in n number of time series, their corresponding segments \bar{Q}_j and \bar{C}_j are comparable. In our method, if the two segments are similar segments, then their bit strings are the same, on the other side, if the two segments are not similar, even if their mean values are the same, their fluctuations above and below the mean can be recorded different.

For the average of each segment, the breakpoint is the ordering of the numbered regions under the $N(0, 1)$ Gaussian curve, which can be represented as $B = (\beta_1, \beta_2, \dots, \beta_{a-1})$, where $\beta_{i-1} < \beta_i$ and

$\beta_0 = -\infty, \beta_a = \infty$. The conversion of the vector of PAA coefficients \bar{C} into the string \hat{C} implemented is as follows, and a lookup table that contains the breakpoints is shown in Table 1.

$$\hat{c} * i = \text{alpha} * j, \text{ iif, } \bar{c} * i \in \beta_{j-1}, \beta_j \quad (2.4)$$

Table 1. Lookup table of breakpoints from 2 to 8.

β	2	3	4	5	6	7	8
β_1	0.00	-0.43	-0.67	-0.84	-0.97	-1.07	-1.15
β_2	-	0.43	0.00	-0.25	-0.43	-0.57	-0.67
β_3	-	-	0.67	0.25	0.00	-0.18	-0.32
β_4	-	-	-	0.84	0.43	0.18	0.00
β_5	-	-	-	-	0.97	0.57	0.32
β_6	-	-	-	-	-	1.07	0.67
β_7	-	-	-	-	-	-	1.15

2.2. Distance Measure

As we review above, the PAA distance function *Dist* and the SAX distance function *MINDIST* are defined as the follows. We still use Euclidean distance in distance metrics instead of dynamic time warps [32] or others. DTW can distort the data set to match the nearest neighbors, which is a better method for unequal sequences, but in larger data sets, we can directly match the series with Euclidean distance.

$$\text{Dist}(\bar{Q}, \bar{P}) = \sqrt{\frac{n}{w}} \sqrt{\sum_{i=1}^w (\bar{p}_i - \bar{q}_i)^2} \quad (2.5)$$

$$\text{MINDIST}(\hat{Q}, \hat{C}) \equiv \sqrt{\frac{n}{w}} \sqrt{\sum_{i=1}^w (\text{dist}(\hat{q} * i, \hat{c} * i))^2} \quad (2.6)$$

The trend distance of the binary string between two series is defined as follows, and the length of time series is n and it is divided into w segment.

$$\text{bitDist}(\bar{Q}, \bar{C}) = \sqrt{\frac{w}{n}} \sqrt{\sum_{i=1}^w \text{count}(b_{ci} \oplus b_{qi})} \quad (2.7)$$

where b_{ci}, b_{qi} are the binary string of corresponding segment of two series, and the function *count* is used to sum up the number of 1 in the binary string. Finally, we can define the BIT_Dist measure function based on trend distance and SAX as follows.

$$\text{TSAX}(\bar{Q}, \bar{C}) = \text{MINDIST}(\hat{Q}, \hat{C}) + \sqrt{\frac{w}{n} \text{count}(B_C \oplus B_Q)} \quad (2.8)$$

where B_C and B_Q are the complete series as $B_C = \{b_1, b_2, \dots, b_n\}$. From Equation(2.8), it can be seen that the effect of trend distance on the overall distance is weighted by w/n , which n is fixed. The

larger of w , the greater the proportion of trend distance and the longer length of one segment. Once the subsequence is very long, the trend among this segment will change into a parallel line with no trend change, therefore, the increase of trend distance helps distinguish between the similarity of two subsequence. On the contrary, the smaller of w , the smaller proportion of trend distance. Because if the length of subsequence is small, even contains only two time points, their trend is similar to linear, which will not lose trend information a lot. Algorithm 1 summarizes the procedure in form of pseudo-code.

Algorithm 1 TSAX: Anomaly detection with TSAX

Require: ECG series $T = t_1, \dots, t_n$

number of segment w

Ensure: anomaly score: the top k_{th} anomalies in ECG

```

1:  $T := Z\text{-norm}(T)$  // Normalize the time series
2: for  $i = 0$  to  $w$  do
3:    $mean \leftarrow Tmean(t_i)$  // segment the time series into equal length
4:    $symbol_i = \text{lookupTable}(mean)$ 
5:   // Transform the time series into binary string
6:   for  $j$  in  $t_i$  do
7:     if  $j > mean$  then
8:        $bit = 1$ 
9:     else
10:       $bit = 0$ 
11:    end if
12:     $bitstring \leftarrow bit$ 
13:  end for
14:   $binary\ string \leftarrow bitstring$  // Combine all segments
15: end for
16:  $distbit = \text{compareBit}(t_i, t_{i+1})$ 
17:  $distsax += \text{compareSax}(symbol_i, symbol_{i+1})$ 
18:  $distK = \text{SortDistance}(distsax)$ 
19:  $score = \text{KNN}(distK)$ 
20: return  $score$ 

```

3. Analytical Results

In this section, we demonstrate the effectiveness of our method over real data. The performance is evaluated in different aspects. Experiments are conducted to analyze the performance of our proposed algorithms in comparison to classical unsupervised methods in SAX [29], SAX-TD [31] and ESAX [30] on ECG dataset. All above methods are used to calculate the similarity and anomaly score among dataset and we use KNN to find the most likely anomalies. We set the $k = 3$. The experiments are conducted on a 2.5GHz processor with 16GB physical memory, running Windows 10.

3.1. Experimental setup

3.1.1. Data set

In this section, we use our method to evaluate the data sets downloaded from PhysioBank [33]. The BIDMC Dataset "chf07" is a 20-hour long ECG which have been independently classified into five types. In order to detect the anomalies, we simply divide the multi-class into two class, and the default physionet annotations are converted into AAMI recommended categories as showing in Table3, and different class has different shape as shown in Figure 3, therefore, the trend method is useful in these data sets. We simply combine the Normal beat and one of other abnormal beats together, and we name them with the representations such as "N+r","N+S", etc.

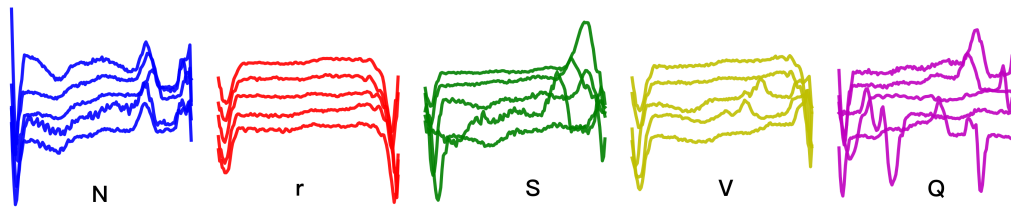


Figure 3. The shape of BIDMC Dataset ch07. Different shapes of ECG segment in five class. The last one is unknown beat so their shapes are different.

Table 2. AAMI recommended beat of five classes and number of ECG in MIT-BIH.

No.	Classes	representation	Number
1	Normal beat	N	2919
2	R-on-T Premature Ventricular Contraction	r	1767
3	Supraventricular ectopic beat	S	194
4	Premature Ventricular Contraction	V	96
5	Unknown beat	Q	24

3.1.2. Comparison methods and Metrics

As base lines, we use several anomaly detection techniques such as SAX [29], SAX_TD [31], ESAX [30]. We compare the baselines with the proposed method TSAX. SAX_TD uses the deviation numerical value between the starting/ending point and mean value as the trend information. ESAX uses the maximum and minimum values based on sax, but symbolizes both values along the way as SAX. To demonstrate the detection ability of the proposed method, True positive (TP), false positive (FP), true negative (TN) and false negative (FN) are defined in Table 3.

Table 3. Definition of statistical values.

	Annotated Anomaly	Annotated Normal
Detected Anomaly	TP	FP
Detected Normal	FN	TN

The algorithm is evaluated based on 5 measurements: sensitivity, false alarm rate, specificity, positive predictive value (PPV) and root of mean square error (RMSE).

Sensitivity describes the proportion of all positive cases identified in all positive cases. False alarm rate describes the case where the negative case is identified as a positive case as a percentage of all negative cases. Specificity describes the proportion of negative cases identified in all negative cases. Positive Predictive Value (PPV) is the part of the anomalies that can be correctly identified in all anomalies. Root of mean square error (RMSE) represents for the sample standard deviation of the difference between the predicted value and the observed value. They can be calculated as follows

$$\text{Sensitivity} = TP / (TP + FN) \quad (3.1)$$

$$\text{False alarm rate} = FP / (FP + TN) \quad (3.2)$$

$$\text{Specificity} = TN / (TN + FP) \quad (3.3)$$

$$\text{PPV} = TP / (TP + FP) \quad (3.4)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (Q_i - C_i)^2}{n}} \quad (3.5)$$

3.2. Accuracy Comparison

According to the above results, since the length of original time series is 140, we set the window length $w = 16$ and symbol number $\alpha = 3$ through these experiments, and each experiment is used by ten-fold cross-validation. To verify the accuracy of different type of mixture ECG signals, we test the four metrics with 5 different combinations. The first four combinations are the combination of a normal class and one of each anomaly class, the last one mixes 5 classes together for comparison. The first four results are shown in Table 4 and indicates that our method can detect different anomalies with low error rate for first three combination due to the shapes in their own class are similar. For the last combination (N+Q), the result is not ideal because of the beat is unknown and also the shape is hard to distinguish. Figure 4 shows the performance of the algorithm on different data sets more intuitively, although the SAX_TD has little higher accuracy than our method, our method performs better in mixture data N+Q that it recognizes the different shape better than others generally.

Table 4. Metrics comparison with one anomaly class with normal class by using our proposed method.

Data set	Anomaly Ratio	Sensitivity	Specific	PPV	Flase Alarm Rate	RMSE
N+r	37.71%	0.987	0.977	0.991	0.023	0.010
N+S	6.23%	0.812	0.625	0.978	0.375	0.013
N+V	3.18%	0.902	0.804	0.984	0.196	0.013
N+Q	0.82%	0.624	0.250	0.830	0.750	0.007

As for the five mixture dataset, we compare the four metrics with other methods. and the result is shown in Table 5. As we can see from the Figure 5, our method of TSAX achieves higher accuracy, specificity and sensitivity while lower false alarm rate than other methods. The advantages of our algorithm are better reflected in the mixed data set, because the mixed data set has many different shapes, so it is more sensitive to the shape.

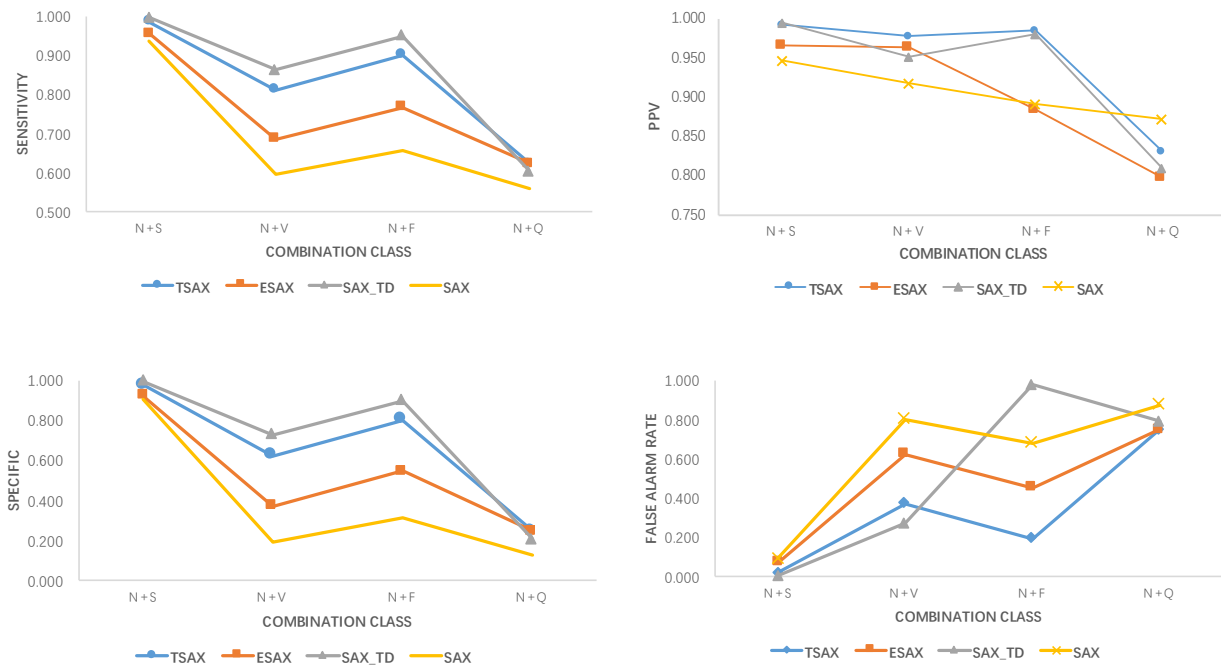


Figure 4. Four metrics comparison among one anomaly class with normal class by four methods.

Table 5. The metrics for five mixture data set by using four methods, and the best results are highlighted in bold.

Method	PPV	sensitivity	specificity	false alarm rate
TSAX	0.9879	0.9847	0.9731	0.0269
SAX	0.8364	0.8385	0.8212	0.1788
ESAX	0.8156	0.8146	0.7785	0.2215
SAX_TD	0.9800	0.9740	0.9607	0.0393



Figure 5. The comparison of four metrics with four methods, the above three are the higher value the better, and the false alarm rate is the lower value the better. The first column is TSAX.

3.3. Error Rate Comparisons

Figure 6 shows the anomaly detection results obtained for this data set when considering the TSAX method with other methods, the parameters of sliding window size is $w = 4$ and the breakpoint is $\beta = 3$. Compare with the ground truth figure below, the above four figures in red parts are detected as anomalies and the blue part are normal. As we can see, error classification of our algorithm is significantly less than others and the false alarm rate are less either.

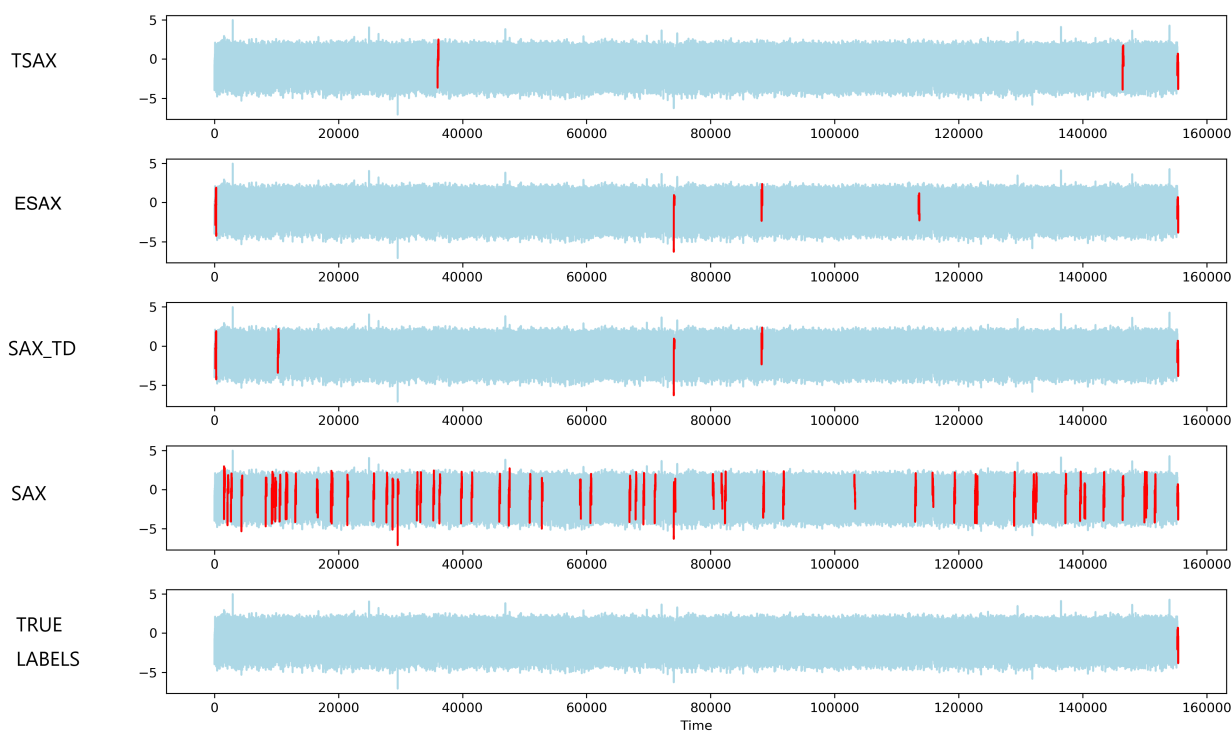


Figure 6. An example of a subsequence of ECG. The above plots are the time series detected by TSAX, SAX_TD, ESAX and SAX, and the below plot is the time series with true anomaly. The subsequence in red are anomalies and the blue subsequence is normal.

One of the advantages of SAX is the novel symbol dimensionality reduction, and the dimensionality reduction ratio is calculated by sliding window size w . In order to verify the effect of the parameter sliding window size w on the accuracy of four methods, we choose to set the breakpoint $\beta = 3$ according to [34], which is suitable for most of datasets. We change the window size range from (3,64) to test its impact on RMSE and the result are shown in Table 6. It can be figured that with the increasing of the length of window, the optimal window size is different in different ways. With the w becomes larger, the TSAX performs the better which indicates that even if this time slice is very long, we can still use the shape to judge the anomalies by our method. And other methods perform not so good when the when the sliding window size gets longer. Therefore, the results show that our methods are more robust than others and the trend representation works.

Table 6. The RMSE result with different window size from 3 to 64. The best win_size is highlight in bold.

Method	3	4	5	6	7	8	16	24	32	64
TSAX	0.0236	0.0220	0.0206	0.0168	0.0168	0.0162	0.0134	0.0144	0.0130	0.0128
SAX	0.0860	0.1012	0.0874	0.1490	0.1302	0.1374	0.1758	0.4158	0.1964	0.2096
ESAX	0.0648	0.0628	0.0604	0.0718	0.0804	0.0906	0.1550	0.2938	0.1984	0.2112
SAX_TD	0.0116	0.0124	0.0100	0.0116	0.0110	0.0130	0.0116	0.0126	0.0132	0.0214

3.4. Computation Time

As for computation performance, our method can be compared the computation time with other methods. In the Figure 7 we can see that the time cost decrease when the window size increasing in all methods. While it takes a lot of time in the process of symbolization, and ESAX cost the most because it needs to convert to symbol more than others and with the sliding window size becomes larger, the time cost decline fast. And other two method TSAX and SAX_TD are similar to SAX. Although the time cost of our TSAX method is little higher than SAX, these times have greatly improved the accuracy, so the time spent is worthwhile.

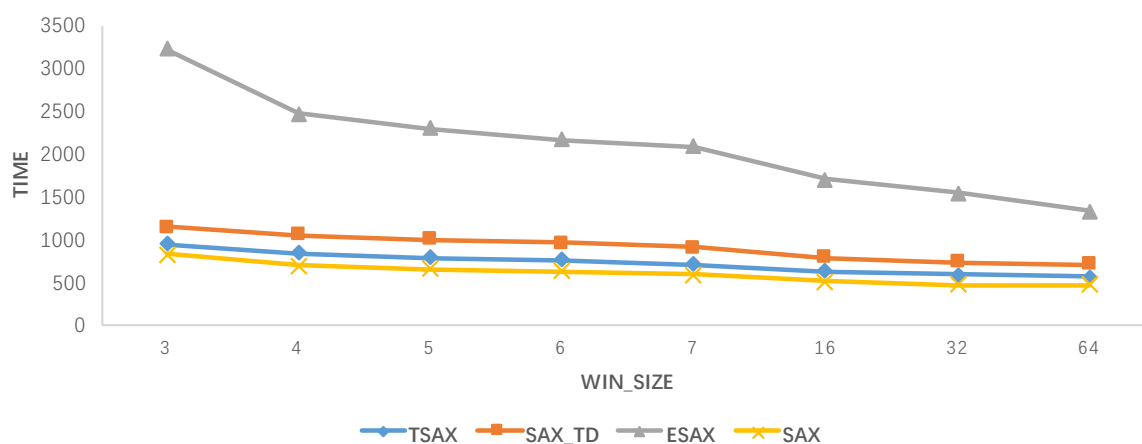


Figure 7. The computation time cost of the methods with different sliding window size, the vertical axis represented as time and the horizontal axis represented for win_size.

4. Discussion and Conclusion

We have shown that our proposed method can achieve higher accuracy with more trend information. This proves that different types of RCG signals do have different shapes which can be detected by our proposed algorithm. In this paper, we propose a trend symbolized method (TSAX) to detect the anomaly heart signals, which use binary string to record the relative trend change of a time series. The relative trend represented by this binary string can distinguish the shape characteristics well. And the similarity distance is base on SAX distance as the final distance measure. We have evaluate the the proposed method using the MIT-BIH datasets, and the result shows that even if the computation time is little bit higher than SAX, it reaches higher accuracy about 98.7% than other methods. In our

future work, we want to divide the original time series with change point to get the adaptive length segments and further improve the efficiency of the algorithm. Dividing time series of different lengths can more accurately measure the ECG anomaly while challenges also exists. In the experiment part, more extensive experiments shall be carried out near future.

Acknowledgments

This study is supported by the Shenzhen Research Council (Grant No.JCYJ20170307151518535).

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. N. Rui and N. Horta, A new sax-ga methodology applied to investment strategies optimization, in *Conference on Genetic and Evolutionary Computation*, (2012), 1055–1062.
2. C. Krgel, T. Toth and E. Kirda, Service specific anomaly detection for network intrusion detection, in *Proc. 2002 ACM Symposium on Applied Computing*, (2002), 201–208.
3. P. Garcia-Teodoro, J. Diaz-Verdejo and G. Maciá-Fernández, et al., Anomaly-based network intrusion detection: Techniques, systems and challenges, *Comput. Secur.*, **28** (2009), 18–28.
4. M. Anderka, S. Priesterjahn and S. Priesterjahn, Automatic atm fraud detection as a sequence-based anomaly detection problem, in *International Conference on Pattern Recognition Applications and Methods*, (2014), 759–764.
5. W. Zhang and X. He, An anomaly detection method for medicare fraud detection, in *IEEE International Conference on Big Knowledge*, (2017), 309–314.
6. V. Chandola, A. Banerjee and V. Kumar, Anomaly detection: A survey, *ACM. Comput. Surv.*, **41** (2009), 15.
7. J. Sigholm and M. Raciti, Best-effort data leakage prevention in inter-organizational tactical manets, in *Military Communications Conference*, (2013).
8. J. Cucurull, M. Asplund and S. Nadjmtehrani, Anomaly Detection and Mitigation for Disaster Area Networks, *Recent Advance. Int. Detect.n*, (2010).
9. Q. Yu, L. Jibin and L. Jiang, An improved ARIMA-based traffic anomaly detection algorithm for wireless sensor networks, *Taylor Francis*, (2016).
10. A. Pyayt, A. Kozionov and V. Kuserbaeva, et al., Signal analysis and anomaly detection for flood early warning systems, *J. Hydroinf.*, **16** (2014), 1025–1043.
11. M. Zhang, A. Raghunathan and N. K. Jha, Medmon: securing medical devices through wireless monitoring and anomaly detection, *IEEE J. Sel. Top Signal Process*, **7** (2013), 871.
12. D. Jiang, Z. Yuan and P. Zhang, et al., A traffic anomaly detection approach in communication networks for applications of multimedia medical devices, *Multimed Tools Appl.*, **75** (2016), 1–25.

13. O. Salem, A. Guerassimov and A. Mehaoua, et al., Anomaly detection in medical wireless sensor networks using svm and linear regression models, *IJEHMC*, **5** (2014), 20–45.
14. W. Einthoven, The string galvanometer and the measurement of the action currents of the heart, *Nobel Lecture, December*, **11**.
15. S. Chauhan and L. Vig, Anomaly detection in ecg time signals via deep long short-term memory networks, in *IEEE International Conference on Data Science and Advanced Analytics*, (2015), 1–7.
16. M. C. Chuah and F. Fu, Ecg anomaly detection via time series analysis, in *International Conference on Frontiers of High PERFORMANCE Computing and NETWORKING*, (2007), 123–135.
17. J. Ma, L. Sun and H. Wang, et al., Supervised anomaly detection in uncertain pseudoperiodic data streams, *ACM T. INTERNET. TECHN.*, **16** (2016), 4.
18. C. Zhang, A. Yin and Y. Deng, et al., A novel anomaly detection algorithm based on trident tree, in *Cloud Computing – CLOUD 2018*, (2018), 295–306.
19. C. Zhang, Y. Ao and H. Liu, Design and application of electrocardiograph diagnosis system based on multifractal theory, *Chin. J. Netw. Inf. Scy.*.
20. C. Zhang, A. Yin and Y. Wu, et al., Fast time series discords detection with privacy preserving, in *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, IEEE, (2018), 1129–1139.
21. H. Sivaraks and C. A. Ratanamahatana, Robust and accurate anomaly detection in ecg artifacts using time series motif discovery, *Comput. Math Methods Med.*, **2015** (2015), 453214.
22. J. Shen, S. D. Bao, Y. C. Yang, et al., Theplr-dtw method for ecg based biometric identification, in *International Conference of the IEEE Engineering in Medicine & Biology Society*, (2011), 5248.
23. H. Ren, M. Liu and Z. Li, et al., A piecewise aggregate pattern representation approach for anomaly detection in time series, *Knowledge-Based Systems*.
24. E. Keogh, J. Lin and A. Fu, Hot sax: efficiently finding the most unusual time series subsequence, in *IEEE International Conference on Data Mining*, (2006), 226–233.
25. J. L. Rodriguez-Sotelo, D. H. Peluffo-Ordoez and D. Lopez-Londoo, Segment clustering for holter recordings analysis, in *International Work-Conference on the Interplay Between Natural and Artificial Computation*, (2017), 456–463.
26. T. Kamiyama and G. Chakraborty, Real-time anomaly detection of continuously monitored periodic bio-signals like ecg, in *Jsai International Symposium on Artificial Intelligence*, (2015).
27. J. Lin, E. Keogh, L. Wei, et al., Experiencing sax: a novel symbolic representation of time series, *Data Min. Knowl. Disc.*, **15** (2007), 107–144.
28. B. Kulahcioglu, S. Ozdemir and B. Kumova, Application of symbolic piecewise aggregate approximation (paa) analysis to ecg signals. in *17th IASTED International Conference on Applied Simulation and Modeling*, (2008).
29. J. Lin, E. Keogh, S. Lonardi, et al., A symbolic representation of time series, with implications for streaming algorithms, in *ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery*, (2003), 2–11.

30. B. Lkhagva, Y. Suzuki and K. Kawagoe, Extended sax: Extension of symbolic aggregate approximation for financial time series data representation, *DEWS2006 4A-i8*, **7**.
31. Y. Sun, J. Li and J. Liu, et al., An improvement of symbolic aggregate approximation distance measure for time series, *Neurocomputing*, **138** (2014), 189–198.
32. M. Yoshimura and I. Yoshimura, An application of the sequential dynamic programming matching method to off-line signature verification, in *Brazilian Symposium on Advances in Document Image Analysis*, (1997), 299–310.
33. M. Kachuee, S. Fazeli and M. Sarrafzadeh, Ecg heartbeat classification: A deep transferable representation, in *IEEE International Conference on Healthcare Informatics*, (2018), 443–444.
34. U. K. Tanaka Y, Motif discovery algorithm from motion data, in *Proc of the 18th Annual Conf of the Japanese Society for Artificial Intelligence*, (2004), 2–3.



AIMS Press

©2019 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)