IJDCST @November Issue- V-1, I-7, SW-46
ISSN-2320-7884 (Online)
ISSN-2321-0257 (Print)

# Classification of ECG Heart beats using Dynamic Time Warping

[1] K.Sandhya, [2] K.VijayKumar
[1] Student, PVPSIT,Kanuru, Vijayawada ANDHRAPRADESH.
[3] Associate Prof, PVPSIT,Kanuru, Vijayawada, ANDHRAPRADESH.

**Abstract:** Time series clustering has become an increasingly important research topic over the past decade. Recent existing methods for time series clustering rely on distances calculated from the entire raw data using the Euclidean distance or Dynamic Time Warping distance as the distance measure. To identify the abnormalities in ECG heart beats through Clustering and Validation by using RST complexes of ECG heartbeats. Our work shows that by using only some *local patterns* and deliberately ignoring the rest of the data. We can mitigate the above problems and cluster time series of different lengths. A novel Time series clustering and Analysis Method for ECG (Electro Cardiogram) heartbeat Analysis is proposed using K-medoids Clustering with Dynamic Time Warping (DTW) distance. The main objective of this paper is to identify the abnormalities in ECG heartbeats through Clustering and Validation by using RST complexes of ECG heartbeats.

**Index terms:** clustering, time-series, shapelets, unsupervised.

## I. INTRODUCTION

The development of bio-signal analysis systems has become a major research field due to technological progress in signal processing. ECG had been amongst the most studied type of bio-signals for several decades. Electrocardiogram records heart contractile activity using electric potentials. Cardiologists in diagnosing heart abnormalities analyze the hard copy ECG diagnostic reports. A decade old empirical comparison by Keogh and Kasetty reveals the somewhat surprising fact that the simple Euclidean distance metric is highly competitive with other more sophisticated distance measures. Dynamic Time Warping (DTW) can both solve this problem and handle the difficulty of clustering time series containing out-of-phase similarities. We argue that the apparent utility of Euclidean distance or DTW for clustering may come from an over dependence on the UCR time series archive.

The problem is that the data in this archive has already been hand-edited to have equal length and alignment. Consider the famous Gunpoint dataset that has been used in hundreds of studies for both clustering and classification. This dataset was contrived to have perfect alignment/length by audible cues that both signaled the actor and started/stopped the video recording. We show the result of clustering these time series if we are allowed to *ignore* some of the data as shown in the fig.1. We can do this without any human intervention and introducing this ability is

the core contribution of this work. We believe for most datasets we must ignore some data is a critical insight in allowing time series clustering in realistic problem settings.

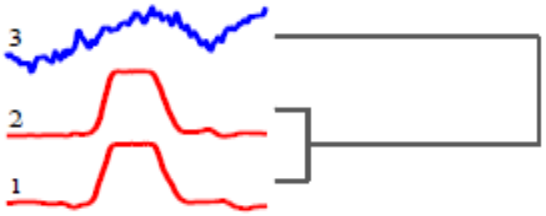


**Figure 1: Clustering Gunpoint after ignoring some data.**

Time series shapelets are small, local patterns in a time series that are highly predictive of a class.

## II. RELATED WORK

**Motivation:**

ECG abnormality detection is a challenging task in short term or long-term recordings. Medical practitioner relies on physical verification to detect abnormalities in patient's ECG hard copy recordings. However, significant amount of work is done on ECG based detection; even tough research is growing on the clustering ECG data. Since there are very few instances of clustering and in all of them depend on extracting the features and clustering based on the features. A method without extraction of remaining features is investigated in the current work by using Dynamic Time Warping distance measure by using RST complexes.

**Time Series Clustering and Classification:**

Clustering as an unsupervised classification of objects identifies structure in an unlabeled data set by objectively organizing data into homogeneous groups where the intra-group-object similarity is minimized and the intergroup-object dissimilarity is maximized, based on a similarity measure chosen.

**ECG Heart-beat analysis:**

Different waves in ECG heartbeat occur due to polarization and depolarization of auricles and ventricles. The natural pacemaker of Heart is caused by the action potential generated by Sino-Atrial (SA) node. This polarization and depolarization causes the auricles or ventricles to relax and contract respectively. RST complex/wave occurs by the depolarization of the ventricles. Depolarization causes the ventricles to contract so that the right ventricle and left ventricle pumps the impure blood and pure blood to lungs and arteries respectively. The five types of classes in ECG are used:

- Normal (N)
- Left bundle branch blocks (LBBB)
- Right bundle branch blocks (RBBB)
- Premature ventricular contraction (PVC)
- Atrial premature contraction (APC)

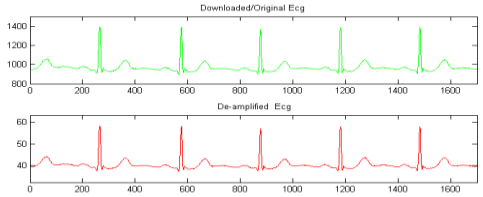


**Figure 2: MIT-BIH ECG 103 a) Original ECG (top) b) After Deamplified by gain (bottom)**

**Dynamic Time Warping Distance Measure:**

Dynamic time warping (DTW) algorithms compute distance of two unequal length time series segments by aligning the signals using dynamic programming based constraints. The Apart from the time series it can warp other dimensions like an angle for shape. DTW does not satisfy Triangle Inequality that is a condition of a Distance metric. Assume two time series A and B of lengths J and I respectively. Let A= {$A_1$, $A_2$… $A_J$} be a sequence of M samples of a 1-dimensional signal and B= {$B_1$, B2….$B_I$} be a sequence of I samples of another 1-dimensional signal. To compute the warping path between these 2 signals, DTW uses 2 passes:

1. Results a matrix of local distances C (x, y) of size IxJ using Minkowski distance as Local distance matrix

   C (x, y) = || ($A_X$-$B_Y$) ||

2. Results a matrix of global distances D(i, j) of size IxJ using local distances as global distance matrix

   - Cumulative sum of local distances of the row

   D (1, y)= $\sum_{z=1}^{y} C(A_1, B_k) for$ ;

   - Cumulative sum of local distances of the column

   D(x, 1) = $\sum_{z=1}^{x} C(A_k, B_1) for$ .

A warping path W is a contiguous set of matrix elements that defines a mapping between two time series' A and B.

## III. PROPOSED SYSTEM

First RST waves are detected and extracted using Difference Operation Method (DOM). The proposed method uses K- medoids clustering algorithm using dynamic time warping as a distance measure.

The five types of data classes used are found that either the models considering only three clusters performed better than models with fewer or more clusters. First, the data is clustered using K- medoids creating three clusters to find abnormalities of ECG

## IV. RESULT ANALAYSIS

We cluster the heartbeats using K-medoids s clustering algorithm using dynamic time warping as a distance measure to find abnormalities of ECG .

**RST complex Demarcation:**

RST waves are demarcated and extracted using Difference Operation Method (DOM). As shown in the fig.2 the original downloaded ecg103 record from MIT/BIH database and after de-amplified by ADC gain. To find the M no. of peaks the differences between the signals is calculated. After finding M points of all beats of ECG file, next step is finding the points B and E.

**K-medoids Clustering:**

K-medoids clustering algorithm using Dynamic Time Warping as distance measure is used. It returns Centers and index as output parameters. One center for each of k clusters, centers contains k centers. The below figure shows the k-medoids algorithm.

1. Initialize randomly the k cluster centers (medoids).
2. Assign each object to appropriate cluster centers by calculating DTW distance between each object and the centers, by choosing the nearest center to that object.
3. Calculate distances among the objects in each cluster to find new cluster center, and assign minimum distance object.
4. If any object changes clusters, repeat from step 2 to 4.

**Figure 3: K-MEDOIDS ALGORITHM**

## V. CONCLUSION

RBBB and PVC are difficult to classify because DTW distances between RST complex of RBBB or PVC and RST complex of NORMAL is less for some beats when compared to other distances. Future we can enhance unsupervised shapelets and showed their utility. It can select representative u-shapelets from a time series without any human intervention.

## VI. REFERENCES

[1] Jagadeeswara Rao Annam, Sai Sudheer Mittapalli, Bapi Raju S, Time Series Clustering and Analysis of ECG Heart-beats using Dynamic Time Warping, DCIS, University of Hyderabad, Hyderabad, India.

[2] E. Backer and A. K. Jain, A Clustering Performance Measure Based on Fuzzy Set Decomposition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-3, No. 1, pp. 66-75, January 1981.

[3] Chazal P. De, Reilly, R.B.: Automatic classification of ECG beats using waveform shape and heart beat interval features. In: Proceedings of the IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP03), Hong Kong, China, (vol 2), pp. 269-272.(2003).

[4] Caiani EG, Porta A, Baselli G, Turiel M, Muzzupappa S, Pieruzzi F, Crema C, Malliani A, Cerutti S (1998), Warped-average template technique to track on a cycle-by-cycle basis the cardiac filling phases on left ventricular volume. IEEE Comput Cardiol 25:7376, (1998)

[5] D. J. Berndt and J. Clifford, "Using Dynamic Time Warping to find Patterns in Time Series," AAAI-94 Workshop on Knowledge Discovery in Databases, pp. 229–248.

[6] D. Cerra, J. Bieniarz, J. Avbelj, P. Reinartz, and R. Mueller, "Compression-based unsupervised clustering of spectral signatures," Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS 2011).

**About Authors**

Sandhya received the B.tech (CSE) Degree from PVP Siddhartha Institute of Technology, Vijayawada and affiliated to JNTU Kakinada University, India. She is currently pursuing M.Tech (CSE) Degree at the Dept of Computer Science Engineering, PVP Siddhartha Institute of Technology, Vijayawada.

KANCHARLA Vijay kumar received the B.tech (CSIT) Degree from Gudlavalleru Engineering College, Gudlavalleru. He received the M.Tech (Software Engineering) Degree under Jntu Hyderabad University. He has more than 8 years teaching experience. He is currently working as an Asst Professor in the Dept of Computer Science Engineering, PVP Siddhartha Institute of Technology, Vijayawada and affiliated to JNTU Kakinada University, India. His interests are Data Mining, software engg, operating system.