A NEW METHOD FOR PAROXYSMAL ATRIAL FIBRILLATION AUTOMATIC PREDICTION

BY

HARITON COSTIN1,2,∗, CRISTIAN ROTARIU1 and ALEXANDRU PĂȘĂRICĂ1

1Grigore T. Popa University of Medicine and Pharmacy, Iași, Faculty of Biomedical Engineering
2Institute of Computer Science of Romanian Academy Iaşi Branch

Received: February 25, 2013
Accepted for publication: March 28, 2013

Abstract. The main reason of this study is the long-term effects of atrial fibrillation of the human heart, which lead to increased risk of cardiac mortality. This paper presents the use of two different methods for the prediction of the onset of paroxysmal atrial fibrillation (PAF) by means of surface electrocardiographic (ECG) signal automatic analysis. The first method is commonly used and consists in the analysis of the heart rate variability (HRV) of the ECG signal. Two significant parameters are taken into consideration: the time-domain standard deviation of average five minute window of the time series and the frequency-domain low-frequency/high-frequency ratio of the ECG RR interval. The second analysis method, which is based on the morphological timing characteristics of the QRS complex, is called morphologic variability (MV) of the ECG signal, and was not used before us for PAF prediction. Both methods are applied on 198 Holter records taken from the PAF Database found on physionet.org portal. The results show a better accuracy of the MV analysis than that obtained by means of HRV technique alone. Moreover, by using an appropriate decision rule, both methods were combined and the overall accuracy of PAF onset prediction was raised up to 90%. Experimental results indicate that our method is applicable for usual Holter recordings, is robust against noise and

∗Corresponding author; e-mail: hcostin@gmail.com
common artifacts, and is fast enough, as it uses only 5-minute signal windows. Its high prediction accuracy is comparable with that obtained by manual annotation made by experts and is suitable to be used in clinical practice.

**Key words:** atrial fibrillation prediction, surface ECG, HRV analysis, morphologic variability.

*2010 Mathematics Subject Classification:* 60G35, 94A12, 92C55.

### 1. Introduction

In recent years a relatively large number of both clinical and technical studies have been done on the topic of paroxysmal atrial fibrillation (PAF) of the heart muscle. This is defined as short duration episodes of AF lasting from 2 min. to less than 7 days, while chronic AF is defined as lasting more than 7 days. The main reason for this is not the immediate effect of the onset of atrial fibrillation over the patient’s health (AF detection) but the long-term effects: increased risk of cardiac mortality, associated respiratory problems, increase in heart muscle fatigue, increase in thromboembolic and stroke events, etc.

Paroxysmal atrial fibrillation (PAF) is a heart arrhythmia characterized by a highly irregular rhythm, swirling blood flow that leads to a high risk of stroke due to the formation of blood clots and an irregular onset that makes it hard to detect on normal ECG tests. Thus it is necessary for cardiologists to benefit from a robust and precise tool that could predict the onset of such events, in order to prevent them by defibrillation, drug treatment and anti-tachycardia pacing techniques. During recent years many attempts were made to obtain automatic and accurate prediction of PAF. Papers (Poli et al., 2003; Bollman et al., 2006; Chiarugi, 2008), and (Sörnmo et al., 2009) represent useful reviews describing different techniques for PAF or chronic AF prediction, from technical to clinical points of view. The “Computers in Cardiology Challenge 2001” revealed a maximum obtained accuracy of about 80% (Moody et al., 2001; de Chazal & Heneghan, 2001; Maier et al., 2001). Paper (Thong et al., 2004) reports a sensitivity and specificity of 89% and 91% respectively, by analysis of atrial premature complexes (that trigger 93% of PAF episodes), also papers (Chandy et al., 2004; Wiggins, 2005), and (Sovilj et al., 2010) describe studies for AF prediction after coronary artery bypass graft (CABG) surgery. P wave of ECG signal may act as a predictor for PAF (Cabasson et al., 2010), K-Nearest Neighbors, Bayes classifier and multilayer perceptrons were used in (Pourbabaee et al., 2008) for the same reason, support vector machines proved their usefulness (Graja & Boucher, 2005), PR interval of ECG signal and neural networks were tools for screening PAF in (Arvaneh et al., 2009), wavelets proved usefulness within a complex procedure (Maglaveras et al., 2002), also rough sets and decisions rules were used in (Wiggins et al., 2006).
In general, these above prediction models are able to detect the transition to PAF events with accuracies of 70–90%, by means of records of at least tens of minutes and rather complex analysis procedures. Moreover, some methods tend to employ too many ECG parameters to characterize PAF, which yield an increased computational complexity.

Paper (Syed et al., 2008) introduced the concept of morphologic variability (MV) of ECG signal (also adopted by us within this study) and used it with good results for short-term risk-stratification following acute coronary syndromes.

Traditionally, metrics in time-domain, frequency-domain, and non-linear approaches are appropriate to calculate heart rate variability (HRV) (Malik, 1996), and its use for PAF screening and prediction was extensively analyzed e.g. in (Yang & Yin, 2001; Lynn & Chiang, 2001), and (Chesnokov et al., 2007). Their main drawback is that they mainly offer information about the RR interval and its power spectrum, but do not approach the changes in the ECG morphology, which may often appear. That’s why this new method – morphologic variability (MV) – can be used to improve the results obtained with HRV and other metrics.

2. Database and Methods

The database used for this study was PAF Prediction Challenge Database 2001 from Physionet.org portal (physionet.org, 2001), mainly its long-term version (physionet.org, ltafdb). The facilities offered by this important database of ECG recordings are presented in (Golberger et al., 2000) and (Moody et al., 2001). Each record contains two simultaneously recorded Holter ECG signals, V1 and II leads, digitized at 128 Hz with 12-bit resolution over a 20 mV range. This is an automatically annotated database and consists of 3 types of record sets. The first record set which has records that begin with the letter “n” comes from 50 subjects who do not have documented atrial fibrillation, either during the period from which the records were excerpted or at any other time. The length of these records is 30 min.

The second record set contains records from 25 subjects: two 30-min records with consecutive record names (e.g., p15 and p16), and two 5-min “continuation” records with names ending in “c” (e.g., p15c and p16c) for each subject. All four records in each record set are excerpts of continuous ECG recordings of a single subject. The names of these records begin with the letter “p”. The first record (odd-number) contains 30 min of the ECG during a period that is distant from any episode of PAF (there is no PAF during the 45-min period before the beginning or after the end of the 30-min record). The second record (even number) contains 30 min of ECG signal that immediately precedes a PAF, as shown in the file that has the letter “c” at the end of the name. The first two record sets represent the learning set.
The third record set contains 99 annotated recordings. These are named t01, t02..., t99 and as the learning set contains 30 min of ECG signal. According to the database details approximately half of the test recordings come from subjects with PAF. We used both channels of the ECG signal to create our final database, summarized in Table 1.

<table>
<thead>
<tr>
<th>The Data Base Used for PAF Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Learning set</strong></td>
</tr>
<tr>
<td>non-PAF recordings</td>
</tr>
<tr>
<td>Pre-PAF areas</td>
</tr>
<tr>
<td>99</td>
</tr>
<tr>
<td>50</td>
</tr>
</tbody>
</table>

In order to accurately detect $R$ peaks in the ECG signal some preprocessing steps are necessary. In this way, false or missed peaks that appear due to noise or artifacts are corrected. Also, ectopic or supraventricular beats have to be removed. We have used an automated preprocessing technique – the “20% filter”, in which $RR$ intervals differing more than 20% of the previous interval are replaced by the average value of the 5 preceding and 5 following intervals.

Other preprocessing stages imply removal of baseline wander and interference noise reducing. The ECG (isoelectric) baseline is restored by means of a band pass filter with cut-off frequencies of 0.5 Hz and 40 Hz. For lowering interference noise, a median filter with length 8 was used.

We detected $R$ peaks from ECG signal using a software program based on the traditional Pan-Tompkins algorithm (Pan & Tompkins, 1985), found on physionet.org portal. This algorithm is credited with an average error rate of about 1%, and false detections occur mainly due to noise.

The first used prediction method is based on the heart rate variability (HRV) metrics. From the large number of possible indicators (e.g. SDNN, RMSSD, HRVF, HRVI) the time-domain metric SDANN (standard deviation of average five-minute window of the time series) and the frequency-domain metric ratio (LF/HF) were chosen to be used due to their good previous results.

The SDANN indicator is determined using the formula:

$$SDANN = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\overline{RR}_i - \overline{RR})^2}, \text{[ms]}$$

where $\overline{RR}_i$ is the mean of $RR$ intervals in “i”-th 5-min window; $\overline{RR}$ – the mean of all means of $RR$ intervals in all 5-min windows; $N$ – the number of analysis windows.
The LF/HF ratio or Symphatovagal balance Index \( (SVI) \) \( (2) \) is the ratio between the low frequency \((0.04 \text{ Hz} - 0.15 \text{ Hz})\) and high frequency \((0.15 \text{ Hz} - 0.4 \text{ Hz})\) total power components of the \( RR \) time series. They are obtained by applying the Fourier Transform to the ECG signal and this was done using the HRV analysis software available on physionet.org.

\[
HRV_{\text{- LF}} / HF = \frac{\text{Power between 0.04 and 0.15 Hz}}{\text{Power between 0.15 and 0.4 Hz}} \tag{2}
\]

The second analysis is based on a newer signal processing method named \textit{morphologic variability} \( (MV) \) \( (\text{Syed et al.}, 2008) \), which highlights the underlying physiological activity of the heart. First, as a preprocessing stage, this method uses \textit{dynamic time-warping technique} \( (DTW) \) which is mostly used in voice pattern recognition \( (\text{Rabiner, 1978}) \). The need for DTW of ECG signal comes from the fact that a reliable algorithm has to compute the \textit{energy changes} between consecutive beats, not only differences in amplitude or time axis.

As an example, Fig. 1 shows that a simple Euclidean distance is not a proper metric to compare the two beats depicted in the left part, as the T-wave areas following QRS complexes must be aligned or warped in time domain to reveal morphological differences between several beats, as shown in the right part of Fig. 1.

![Fig. 1](image1.png)

\[ \text{Fig. 1 – Using of dynamic time-warping method can align two ECG beats in order to assess morphologic differences between them.} \]

The DTW distance metric computes the distortion needed to align two time series. An alignment of two sequences, \( e.g. \) A and B, of length \( m \) and \( n \) respectively, is a sequence of integer pairs of length \( k \),

\[
(\varphi_a[1], \varphi_b[1]), (\varphi_a[2], \varphi_b[2]), \ldots, (\varphi_a[k], \varphi_b[k]),
\]

where \( \varphi_a \) and \( \varphi_b \) satisfy the boundary conditions:

\[
\varphi_a[1] = 1, \varphi_b[1] = 1 \tag{3}
\]

\[
\varphi_a[k] = m, \varphi_b[k] = n.
\]
Also, they satisfy the continuity conditions:

\[ \varphi_A[j] \leq \varphi_A[j + 1] \leq \varphi_A[j] + 1, \quad \forall j : 1 \leq j < k; \]  

\[ \varphi_B[j] \leq \varphi_B[j + 1] \leq \varphi_B[j] + 1, \quad \forall j : 1 \leq j < k. \]

Each ordered pair \((\varphi_A[i], \varphi_B[i])\) matches two elements to be aligned.

The DTW distance between \(A\) and \(B\) is the minimum sum-of-squares difference between pairs of matched elements under all allowable alignments,

\[
DTW(A, B) = \min_{\varphi_A, \varphi_B} \sum_{i=1}^{k} \left( \Phi_A[i] - \Phi_B[i] \right)^2,
\]

where \(k\) is the length of the sequence \(\varphi_A\). In this manner, DTW captures (by varying \(k\)) both amplitude and timing differences between the two signals.

DTW may be computed in an efficient way using dynamic programming, as follows. Suppose we define the function \(D(i, j)\) as the minimum cost of aligning the subsequences \(A[i ... m]\) and \(B[j ... n]\), and by convention let \(D(i, j) = \infty\) if \(i > m\) or \(j > n\). If \(i = m\) and \(j = n\), then the two inputs only have one sample each, so \(D(m, n) = (A[m] - B[n])^2\). In general, the continuity conditions yield the following recurrence relation:

\[
D(i, j) = \left( A[i] - B[j] \right)^2 + \min \left\{ D(i+1, j), D(i, j+1), D(i+1, j+1) \right\}.
\]  

One may say that the cost of aligning two subsequences is the sum of two terms: the cost of aligning their first elements, and the cost of the cheapest possible way of aligning the remainders of the two sequences. An implementation of eq. (6) can be used to compute the optimal total cost (namely, \(D(1,1)\)) in \(O(mn)\) time. In this way DTW produces the optimal alignment of the two sequences by constructing a \(m\)-by-\(n\) distance matrix \(D\). Each entry \((i,j)\) in this matrix \(D\) represents the square of the difference between samples \(A[i]\) and \(B[j]\). The general recurrence (6) allows a single sample in either sequence to be aligned with an arbitrarily large number of consecutive samples in the other sequence. This is unrealistic in many applications because such a situation usually represents an unphysical alignment. Due to this reason we have used a restrictive version of (6):
The above recurrence for ECG signals ensures that one sample in either sequence is aligned with maximum 3 samples in the other sequence.

The used algorithm aligns the current QRS complex to the previous two and the next two complexes, then determines the difference between the lengths of each consecutive QRS complex in order to obtain the morphological difference (MD) time series:

\[ MD = |QRS_i - QRS_{i-2}| + |QRS_{i-1} - QRS_i| + |QRS_{i-1} - QRS_{i+1}| + |QRS_{i+1} - QRS_{i+2}|, \]

where \( QRS_i \) is the length of the QRS complex of the \( i \) beat.

In this way, the technique described here measures changes in morphology resulting from both amplitude and timing differences between two beats and transforms the original ECG signal from a sequence of beats to a sequence of energy differences. This new signal, comprising pairwise, time-aligned energy differences between beats, is then smoothed using a median filter of length 8.

The MV for a record can be calculated from the MD time-series using metrics characteristic for HRV analysis. This method offers complementary information to the analysis of RR intervals. Again, the two metrics used are SDANN and LF/HF ratio. These indicators determine two new characteristics of MV: MV_SDANN and MV_LF/HF. The results obtained with both HRV and MV are then compared.

3. Results and Discussions

For the training set, as well as for the test set, we used first the HRV software and computed two indicators: HRV_SDANN and HRV_LF/HF. Then, by means of an algorithm implemented in Matlab® 2008, we computed the morphologic distance between QRS complexes, we obtained morphologic variability (MV) and computed similar parameters: MV_SDANN and MV_LF/HF. The sampling frequency of 128 Hz was enough for the purposes of the study, as we mainly computed RR and QRS intervals, not the exact morphology of the ECG or PAF signals.
A t-test for the parameters of HRV and MV, for the “n” recording set (characterized by a normal ECG signal) and for the “p” set (with PAF), was performed. At this stage we determined confidence intervals for the mean value, for a confidence level of 99% and $p = 0.01$.

The t-test results given in Table II show that both “n” set and “p” set do not contain data significantly different from the reference value (mean value). Moreover, for a value $p = 0.01$ lower and upper confidence level are specific for each type of HRV or MV indicator, both for normal and pre-PAF ECG signals. The t-test set from Physionet contains 142 “normal” recordings (without PAF episodes) and 56 recordings with PAF.

Every recording was evaluated and manually annotated by experienced cardiologists, with an accuracy considered to be approximately 95% ± 1%. The obtained indicators for “t” set are compared with confidence intervals limits computed by means of t-test applied on non-PAF and PAF training sets.

The recordings are labeled with “N” if they are of type normal or “A” if the analyzed signal precedes a PAF event.

### Table 2

<table>
<thead>
<tr>
<th>Record type</th>
<th>Indicator</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>$t$-statistics</th>
<th>Confidence intervals: 99% ($p = 0.01$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower limit</td>
</tr>
<tr>
<td>Normal</td>
<td>HRV_SDANN</td>
<td>0.04843</td>
<td>0.04261</td>
<td>−4.64</td>
<td>0.02982</td>
</tr>
<tr>
<td></td>
<td>HRV_LF/HF</td>
<td>2.6129</td>
<td>1.66962</td>
<td>−1.59</td>
<td>0.97977</td>
</tr>
<tr>
<td></td>
<td>MV_SDANN</td>
<td>0.35036</td>
<td>0.14076</td>
<td>1.91</td>
<td>0.31753</td>
</tr>
<tr>
<td></td>
<td>MV_LF/HF</td>
<td>1.5213</td>
<td>0.7742</td>
<td>2.99</td>
<td>1.53603</td>
</tr>
<tr>
<td>PAF</td>
<td>HRV_SDANN</td>
<td>0.0238</td>
<td>0.02215</td>
<td>−1.19</td>
<td>0.01798</td>
</tr>
<tr>
<td></td>
<td>HRV_LF/HF</td>
<td>0.40211</td>
<td>0.19789</td>
<td>−1.83</td>
<td>0.35013</td>
</tr>
<tr>
<td></td>
<td>MV_SDANN</td>
<td>0.27448</td>
<td>0.09819</td>
<td>−4.68</td>
<td>0.24869</td>
</tr>
<tr>
<td></td>
<td>MV_LF/HF</td>
<td>1.38199</td>
<td>0.39612</td>
<td>7.07</td>
<td>1.32796</td>
</tr>
</tbody>
</table>

### Table 3

Types and Number of Recordings from “t” Set

<table>
<thead>
<tr>
<th>Type</th>
<th>HRV (correct/total)</th>
<th>MV (correct/total)</th>
<th>HRV and MV (correct/total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>120/142</td>
<td>124/142</td>
<td>127/142</td>
</tr>
<tr>
<td>A</td>
<td>47/56</td>
<td>49/56</td>
<td>50/56</td>
</tr>
<tr>
<td>Total</td>
<td>167/198</td>
<td>171/198</td>
<td>177/198</td>
</tr>
</tbody>
</table>
The results in Table 3 are obtained for HRV, MV and a combined method, respectively. The last method implies a decision support module that labels a recording from “t” set as being “N” type (normal) or “A” (PAF) if both methods (HRV and MV) indicate the same signal type. If an indecision situation occurs, the system analyzes indicators individually, and a decision is taken if and only if three out of four indicators (HRV/MV_SDANN, HRV/MV_LF/HF) show a signal of the same type. The decision rules are also implemented in Matlab© 2010.

Also, statistical indexes of sensitivity and specificity, for each analysis method, for \( p = 0.01 \) were computed using formula (9), where \( e.g. \) TP are true positive test results, \( i.e. \) ECG recordings preceding PAF events:

\[
\text{Specificity} \% = \frac{TN}{TN+FP} \times 100; \\
\text{Sensitivity} \% = \frac{TP}{TP+FN} \times 100. \quad (9)
\]

The above indicators are useful to draw ROC (“receiver operating characteristic”) curves, as shown in Fig. 2. The ROC curves also show correlation between MV values and PAF onset. These curves show binary classifiers performance when discrimination threshold varies. In this respect, the more area under the curve is greater, the classifier accuracy (in our case of the prediction method) is better.

The ROC curve analysis provides the area under the curve for each used method (HRV, MV and combined HRV&MV methods), for the specified cut-off values. Both the area under the curve, that is proportional with the prediction precision, and the cut-off values are presented in Table 4.

![Fig. 2 – ROC curves for HRV, MV and both methods combined.](image-url)
Table 4

ROC Curve Analysis: Area under the Curve and Cut-Off Values

<table>
<thead>
<tr>
<th>Method</th>
<th>ROC Area</th>
<th>Cut-off value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HRV</td>
<td>0.842</td>
<td>0.0527</td>
</tr>
<tr>
<td>MV</td>
<td>0.874</td>
<td>0.2956</td>
</tr>
<tr>
<td>HRV &amp; MV</td>
<td>0.894</td>
<td>1</td>
</tr>
</tbody>
</table>

The final results of the “t” test set by means of HRV and MV analysis, individually used or as a combined method, are shown in Table 5 below.

Table 5

Obtained Specificity and Sensitivity for the Three Used Analysis Methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Specificity (p = 0.01)</th>
<th>Sensitivity (p = 0.01)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HRV</td>
<td>83.93%</td>
<td>84.51%</td>
</tr>
<tr>
<td>MV</td>
<td>87.50%</td>
<td>87.32%</td>
</tr>
<tr>
<td>HRV + MV + decision rule</td>
<td>89.29%</td>
<td>89.44%</td>
</tr>
</tbody>
</table>

4. Conclusions

In this study we applied a synergistic combination of two different automatic prediction methods for PAF prediction, one of them usually used – HRV analysis – and the other – a new technique for surface ECG analysis: morphologic variability (MV) of QRS complexes. Some efficient preprocessing methods, of which the dynamic time-warping (DTW) has a major role, ensure better results than those obtained with HRV alone. The experiments emphasized the main and new outcome of our study, that MV is a reliable electrocardiographic risk stratification measure to predict atrial fibrillation onset. In this respect, the experiments demonstrated that low levels of the MV_LF/HF measure are significantly associated with a higher risk of atrial fibrillation.

Also, the experimental work revealed a better accuracy provided by MV in comparison with HRV method. When using both methods and a decision support module, the prediction accuracy is the best, of about 90%, comparable with the best results in the literature, which were obtained by using special recordings and more complex and time consuming methods. Our method has the main advantage that uses usual Holter recordings and the analysis is a short-term one (on a 5-min window from 30 min total interval).
REFERENCES


O NOUĂ METODĂ PENTRU PREDICȚIA AUTOMATĂ A FIBRILAȚIEI ATRIALE PAROXISTICE

(Rezumat)

Motivația studiului și a lucrării provine din efectele pe termen lung ale fibrilației atriale a inimii umane, care conduc la un risc crescut de mortalitate cardiacă. Această lucrare prezintă utilizarea a două metode diferite pentru prediciția începutului fibrilației atriale paroxistice (FAP), bazate pe analiza automată a semnalului electrocardiografic de suprafață (ECG). Prima din metode este folosită uzual în anumite studii și constă în analiza variabilității frecvenței cardiace a semnalului ECG (HRV).
Din multitudinea de parametri posibili a fi calculați în cadrul HRV am folosit cele mai relevante două metrice (în sensul puterii discriminatorii), anume una temporală – deviația standard într-o fereastră de cinci minute a seriei de timp R-R – și una frecvențială, anume raportul spectru de joasă frecvență/înaltă frecvență a aceluiași semnal – intervalul R-R al ECG.

A doua metodă de analiză, recent introdusă în literatură, se bazează pe dinamica variației morfologiei complexului QRS în timp și se numește variabilitate morfologică (MV). După știința autorilor, MV nu a fost folosită până acum la predicția FAP.

Ambele metode au fost aplicate pe 198 de înregistrări ECG de tip Holter, aflate în baza de date de referință internațională din portalul physionet.org. Rezultatele analizei au indicat o precizie clar mai bună a metodei MV decât în cazul analizei HRV în domeniul frecvenței (cotată cea mai bună tehnică de analiză HRV).

În plus, am „hibridizat” cele două metode de analiză (HRV și MV) și cu ajutorul unei reguli de decizie relativ simplă am obținut o precizie a predicției episoadelor de FAP de cca. 90%, comparabilă cu cele mai bune precizii raportate în literatură, obținute în urma unor analize complexe de semnal, efectuate pe înregistrări de minimum 30 de minute. Noi am folosit înregistrări de doar cinci minute și precizia ridicată am obținut-o în urma unei atente preprocesări a seriei de timp dată de intervalul R-R al ECG, pentru reducerea zgomotului aditiv și a diverselor artefecte suprapuse peste semnalul ECG. În plus am folosit metoda deformării temporale (DTW) introduse pentru procesarea semnalului vocal, în vederea alinierii complexelor QRS.

În concluzie, metoda hibridă propusă poate fi folosită, odată implementată într-un software de aplicație, în mediul spitalicului și validarea ei clinică constituie tema unui demers științific viitor.