Detection of cardiac events in the context of a rehabilitation platform

T. Rocha, S. Paredes
Instituto Superior de Engenharia de Coimbra, Portugal {teresa, sparedes}@isec.pt

J. Henriques, P. de Carvalho
Centro de Informática e Sistemas, Universidade de Coimbra, Portugal {jh, carvalho}@dei.uc.pt

Abstract

Abstract – This work presents the strategy for the early detection of cardiovascular events that was integrated in a rehabilitation platform developed during the FP7 European project HeartWays. The project addressed the development of methodologies able to predict the evolution trend of biosignals time series collected by telemonitoring systems, in order to support the early detection of critical events (such as hypertension and arrhythmic episodes).

The approach is based on the hypothesis that current and past measurements taken from a historical dataset can be the support for the estimation of biosignals future evolution. Two main phases are involved: a similarity analysis procedure to find a set of similar patterns in the historical dataset, and a prediction scheme that makes use of the obtained patterns to forecast the future evolution trend. The validation of the algorithms was performed using blood pressure and heart rate signals collected during the myHeart telemonitoring study.

Keywords – Early detection, trends prediction, similarity measures, wavelets transform.

I. INTRODUCTION

Approximately 5% of all deaths in Europe are due to cardiovascular diseases, and more than 20% of all European citizens suffer from a chronic cardiovascular disease, such as arrhythmias, congestive heart failure and coronary artery disease [1]. Coronary artery disease (CAD) is caused by an accumulation of plaques within the walls of the arteries that supply the myocardium with oxygen and nutrients. After long periods of progression, some of these plaques may rupture and, along with the limited blood supply to the myocardium, may result in an heart attack (myocardial infarction), requiring urgent hospitalization.

Although it is clinically recognized that cardiac rehabilitation provides many benefits for patients, its effective implementation after hospitalization remains low [2]. The use of remote patient monitoring and treatment (RMT) systems, based on telemonitoring services that enable professionals to access and evaluate symptoms and status progression, offers a huge potential in the context of cardiac rehabilitation. In effect, through the continuous monitoring of vital signs (such as blood pressure, heart rate, and body weight) the status and the quality of life and care of the patient can be assessed, and the prediction of aggravations and exacerbations of its chronic condition evaluated.

The FP7 European project HeartWays, (FP7-SME-315659) [3], aims the development of advanced modular solutions for supporting cardiac patients in rehabilitation outside an hospital centre, with the aid of wearable sensors and intelligent algorithms that personalize the management and the follow up for patients and professionals. Figure 1 depicts the approach followed in the project, which consists of three main technological workpackages: WP1: Smart Monitoring Layer; WP2: Multiparametric Analysis Layer; WP3: Patient Support and Healthcare Management Layer.

Figure 1: Schematic diagram of the HearWays project

Included in WP2, one of the modules comprises the development of methodologies for biosignals prediction, mainly to support the early detection of critical events. The specific biosignals to be addressed are daily collected by the system, namely blood pressure (BP) and heart rate (HR). The particular
events to be detected are hypertension and arrhythmic episodes, based on the evolution of BP and HR, respectively.

In terms of strategy, this module is founded on the hypotheses that the estimation of biosignals’ future evolution can be supported on current and past measurements, captured from a historical dataset.

As result, from the research and practical perspectives, two major topics are addressed: i) how the trends in time series can be captured and compared; ii) how these trends can be used in the prediction process.

For the assessment of similarity between time series two main groups of algorithms can be identified: time domain and transform-based methods [4]. In the context of this work, the time-frequency analysis methods, included in the second group of algorithms, assume a particular interest. This is the case of the wavelet transform that produces features that describe properties of the time series both at various locations and at numerous time granularities, which is particularly important when dealing with the similarity assessment problem [5]. In particular, a specific Haar wavelet was used in the proposed similarity search scheme. Given its characteristics, the application of the wavelet transform was also considered for trend extraction and time series prediction. In effect, it provides a formal method to de-noise, de-trend, and decompose time series, capturing useful information at various resolution levels, so that the capacity of a forecasting model can be improved [6].

The structure of this work is as follows: section 2 presents the similarity analysis and prediction schemes. Section 3 discusses its application to heart rate and blood pressure signals for the detection of arrhythmic and hypertension episodes, using data collected during myHeart tele-monitoring study. Finally, in section 4, some conclusions are drawn.

II. METHODOLOGY

This work is founded on the hypothesis that the estimation of biosignal’s future evolution can be supported on current and past measurements taken from a historical dataset. In effect, the proposed methodology involves two main stages, as illustrated in Figure 2.

![Figure 2: Prediction of biosignals evolution](image)

**Similarity analysis procedure:** firstly, by means of a similarity analysis process, the selection of patients who display similar behaviors in their physiological time series is carried out;

**Estimation of future evolution trend:** then, the estimation of the biosignal’s future values is performed, based on the similar time series identified in the first stage.

Basically, the process starts by considering the current signal to be predicted, designated here as the template, \( X(t) \in \mathbb{R}^{1,N} \). Using the template and from a similarity analysis procedure, the set of the M most similar patterns \( X(t) = \{ X_m(t) \in \mathbb{R}^{1,N} \} \), \( m = 1, \ldots, M \), is identified. From these, the corresponding subsequent P future values, \( Y(t) = \{ Y_p(t) \in \mathbb{R}^{1,P} \} \), are straightforwardly obtained (known past values from historical dataset). Then, the known “future” evolution of the identified patterns, \( Y(t) = \{ Y_p(t) \} \), can be used in a prediction mechanism to estimate the future evolution of the current template, \( \hat{Y}(t) \in \mathbb{R}^{1,P} \).

A. Similarity Analysis Procedure

As previously referred, the first step of the prediction strategy consists in selecting patients that display similar behaviors in their physiological time series (historical dataset). To this end, algorithms able to find the segments of a time series that present the same dynamics of a given temporal template (indexing process) should be developed. The proposed methodology for evaluating the similarity between two physiological time series combines the Haar wavelet decomposition, in which signals are represented as linear combinations of a set of orthogonal basis, with the Karhunen-Loève transform, that allows for the optimal reduction of that set of basis. The similarity measure is based on the Euclidean distance, which is indirectly calculated by means of the linear combination coefficients of both time series. Furthermore, using an iterative algorithm for computing the referred coefficients, computational complexity of the method significantly decreases. Figure 3 illustrates the procedure.

![Figure 3: Similarity analysis scheme](image)
Follows a brief description of each step identified in the above scheme.

- **Step 1** - Vertical shift removal: to guarantee that similarity assessments are independent of variations in the vertical position, a vertical shift removal procedure is employed.
- **Step 2** - Wavelet decomposition of the template: to be compared with the time series is achieved by means of a set of orthogonal wavelet basis.
- **Step 3** - Optimal dimension reduction: based on the localization property of the wavelet basis, the ones that significantly reflect the dynamical patterns of the template are chosen to compose a reduced set of basis.
- **Step 4** - Sequence description: a subsequence of the signal to be compared with the template is described by means of the previous reduced set of basis. It is important to refer that this description does not involve a wavelet decomposition, but a simple computation of coefficients.
- **Step 5** - Similarity measure: the coefficients obtained from the template and subsequence description using the reduced set of basis, are employed to derive a similarity measure. This measure allows the interpretation as a trend evolution, as well as a percentage of the amplitude difference between the time series.
- **Step 6** - Subsequence indexing: based on the previous similarity measure, and using the particular Haar wavelet, an efficient iterative similarity indexing algorithm is proposed.

The parameters to be selected, $\varepsilon \in \mathbb{R}^+$ and $\eta \in \mathbb{R}^+$ correspond to: $\varepsilon$ : controls the approximation error by determining the number of basis to be considered in the template decomposition; $\eta$ : establishes if two signals that present the same behaviour are or not similar by thresholding the difference in amplitudes of the two series under comparison.

### B. Estimation of Future Evolution Trend

Figure 4 depicts the global scheme for the trend estimation of biosignals. It is composed of three main distinct phases: i) similarity analysis to identify patients who display similar behaviours in their physiological time series (previously presented); ii) multi-resolution decomposition of the time series retrieved from such patients, $X(t) = \{X_n(t) \in \mathbb{R}^X\}$ and $Y(t) = \{Y_n(t) \in \mathbb{R}^Y\} ;$ iii) projection of the current patient data (template), $X(t)$, into the future, $\hat{Y}(t)$, by combining the optimal decomposition levels of the historic patterns $Y(t)$.

It is important to refer that the methodology presented here does not explicitly involve a model. In effect, it is based on the wavelet decomposition of the similar historical patterns, in order to derive an optimal future trend for the template. To achieve this goal two main steps are need: i) computation of the representative trends and; ii) computation of the optimal trends.

**Representative trends**: The first step involves the wavelet decomposition of the similar historical time series signals. Then, at each decomposition level, the obtained decompositions are combined to obtain a representative trend. The subtractive clustering method is used for this purpose.

**Optimal trends**: The second step aims at the selection of a subset from the representative trends, designated as optimal trends, to be used in the prediction of the current time series. To achieve this goal, an optimization process involving the minimization of a set of distance-based measures is proposed. From this process, it is possible to quantify the aptness of each individual representative trend to integrate the optimal subset. Finally, the optimal trends are then straightforwardly extended to the future and aggregated to derive the global trend, $\hat{Y}(t)$.

### III. RESULTS

#### A. Dataset

For the experiments, a private dataset resulting from the MyHeart European project was used [7]. The MyHeart project consisted in a home telemonitoring system that followed the health of heart failure patients, enabling intervention when appropriate. This was done by monitoring vital body signs with wearable technology, processing the measured data and giving recommendations (when appropriate) to the patient and professional users of the system. Using the measured data to give user feedback, the system “closed the loop” of measurements and therapy. This system was used in a clinical observational study carried out with 148 patients from six clinical centres in Germany and Spain. The trial had an enrolment phase of 9 months with 12 months of patient follow up. During the clinical study patients were requested to daily measure weight, blood pressure, and, using a vest, heart rate and bio impedance, as well as breathing rate and activity during the night by means of a bed sensor. Moreover, they were requested to complete each day two questionnaires of symptoms and mood/general well-being. From the 148 patients recruited, 102 (69%) were considered analysable, that is, with more than 30 days of telemonitoring measurements.

#### B. Prediction Methodologies

Two groups of experiments were carried out. The first group assesses the capacity of the proposed wavelet multi-resolution scheme (here designated by WMM) in the trend prediction of
BP and HR signals. Moreover, the performance of this scheme is compared with other typical prediction strategies, namely a linear regression model, the autoregressive integral moving average model - ARIMA, and a non-linear regression model, the generalized regression neural network – GRNN. Other prediction method (AVP) simply considers the average value of predictive signals \( Y_m(t) \), as an estimation for the prediction of \( Y(t) \).

The second set of experiments selects patients with BP/HR values in a critical range (around the threshold of hypertension/tachycardia), and uses the previously estimated trend to determine the risk of hypertension/tachycardia. Specifically, the goal is to evaluate whether during the following week the BP/HR signal of a given patient evolves towards hypertension/tachycardia values or, on the contrary, is maintaining or decreasing to normal values.

i. Parameters

With respect to the ARIMA model, the examination of the autocorrelation and partial autocorrelation functions of the differenced series, was used in the estimation of the order of the model \( ARIMA(n_a,d,n_p) \). The parameters \( n_a \), \( d \) and \( n_p \) identify, respectively, the number of autoregressive terms, the degree of differencing and the number of lagged forecast errors in the prediction equation. As result, the ARIMA structure was \( ARIMA(2,1,2) \). The estimation of parameters was carried out with the \textit{armax}() Matlab command.

Regarding GRNN structure, it can be seen as normalized radial basis function networks, were there is a hidden unit centred at every training case. These units are called "kernels" and, usually, are probability density functions, such as Gaussian functions. The weights from the hidden to output layer are just the target values, so the output is simply a weighted average of the target values of the training cases, close to the given input case. As a consequence, the only parameters to be learned are the widths of the units. In the experiments using BP and HR signals, the width of the kernels was experimentally determined as \( \lambda = 0.2 \). The \textit{newgram}() Matlab command was used to implement this neural model. Moreover, a different neural network had to be trained for each template.

With respect to AVP, the average prediction \( \bar{Y}(t) \), of the identified patterns was computed using an weighted average, taking into account the similarity measure evaluated for each pattern.

The last approach (WMM) put into practice the proposed wavelet strategy, considering the following parameters:

\begin{itemize}
  \item \textit{Similarity analysis:} \( N=32, \ P=8 \), where \( N \) and \( P \) denote, respectively, the time intervals before and after the current time instant; \( M=5 \), number of patterns retrieved from the historical dataset; \( L=5 \), wavelet decomposition level.
  \item \textit{Selection of the optimal trends:} Number of decompositions considered in the optimal trend selection \( l=3,4,5,6 \) (the details are the levels \( l=3,4,5 \); the approximation is the level \( l=6 \); the first two levels of detail \((l=1,2)\) were neglected; conjunction and aggregation operators were, respectively, the \textit{maximum}() and the \textit{product}() operators.
\end{itemize}

\begin{enumerate}
  \item \textit{ii. Metrics}
  
  The accuracy of the forecasting methods was determined in terms of four performance metrics: \( i \) the proposed similarity measure based on the wavelet decomposition+KLT (SWK), \( (1) \); \( ii \) the Pearson’s correlation coefficient (CORR), \( (2) \); \( iii \) the normalised root mean squared error (NRMSE), \( (3) \) and \( iv \) the mean absolute percentage error (MAPE).

  \begin{equation}
  SWK = S \ Y(t)Y(t) \quad t = N+1,\ldots,N+P \quad (1)
  \end{equation}

  \begin{equation}
  CORR = \frac{\sum_{t=N+1}^{N+P} Y(t) - \bar{Y}(t) \ Y(t) - \bar{Y}} \sqrt{\sum_{t=N+1}^{N+P} Y(t) - \bar{Y}(t) \ Y(t) - \bar{Y}} \quad (2)
  \end{equation}

  \begin{equation}
  NRMSE = \frac{1}{P} \sum_{t=N+1}^{N+P} \frac{(Y(t) - \bar{Y})^2}{Y(t) - \bar{Y}} \quad (3)
  \end{equation}

  \begin{equation}
  MAPE = \frac{1}{P} \sum_{t=N+1}^{N+P} \left| \frac{Y(t) - \bar{Y}(t)}{Y(t)} \right| \quad (4)
  \end{equation}

  In the previous equations, \( Y(t) \) is the actual BP/HR value, \( Y(t) \) is the forecasted BP/HR, \( \bar{Y} \) and \( \bar{Y} \) are, respectively, the means of the actual and the estimated signals. The metrics NRMSE and MAPE were transformed into \( NRMSE=\exp(-\kappa_N NRMSE) \) and \( MAPE=\exp(-\kappa_M MAPE) \), in order to guarantee that their values are in the range \([0,1]\). The parameters \( \kappa_N \) and \( \kappa_M \) are constants, respectively, \( \kappa_N = 0.25 \) and \( \kappa_M = 10 \).

  \item \textit{iii. Statistical validation}
  
  Among the available parametric and nonparametric tests, the Friedman test is a nonparametric one that enables to perform multiple comparisons in experimental studies. This test (Friedman, 1937), (Friedman, 1940) is equivalent to ANOVA and is particularly adequate for machine learning studies when the assumptions (independency, normality and homoscedasticity) do not hold or are difficult to verify for a parametric test [8].

  The objective of the Friedman test is to determine if it is possible to conclude, from a set of results, that there is a difference among the several methods. Basically, the Friedman test compares the average ranks \( R_i \) of each method, to decide about the null hypothesis, which states that “\text{Ho: all the algorithms behave similarly and thus their ranks } \ R_i \text{ should be equal}”. The Friedman statistics, is distributed according to \( \chi^2 \), with \( k-1 \) degrees of freedom. From the computation of the corresponding \( p-value \), the null hypothesis can be or not rejected at a given level of significance.
The Nemenyi test enables a pairwise comparison of the methods, based on the average ranks computed in the Friedman test. Basically, by means of the Nemenyi test, two methods can be significantly different at several levels, namely $\alpha=1\%$, $\alpha=5\%$, or $\alpha=10\%$, if their average ranks differ at least the critical value. In this case ($k=4$) the thresholds for the critical values are, respectively, $CD_1 = 1.4675$, $CD_2 = 1.2110$ and $CD_{10} = 1.080$.

C. Trends Prediction

This section summarizes the results obtained from the application of the predictive WMM strategy to the prediction of blood pressure and heart rate signals.

1. Prediction of blood pressure

![Graph showing comparison of prediction methods for blood pressure](image)

Observing all the similarity measures in Figure 5 it is not possible to identify a method that clearly achieves superior results in comparison with the others. In global terms, it appears that the proposed method (WMM) is comparable or slightly superior to the others.

2. Prediction of heart rate

![Graph showing comparison of prediction methods for heart rate](image)

Regarding the prediction of heart rate signals, Figure 6 shows the box-plots with the results corresponding the prediction, respectively. As in the blood pressure case, by the observation of all similarity measures it is not possible to elect the best method. In global terms, it is possible to conclude that the performance of WMM is comparable to some of the methods and superior to others.

D. Detection of Events

1. Risk of developing hypertension episodes

A group of experiments was particularly applied to patients whose blood pressure values were in a critical range (around the threshold of hypertension). The main goal was to employ the trend prediction results to assess the hypertension risk of a patient. Specifically, the aim was to determine whether during the following week the blood pressure signal of that patient would evolve towards hypertension values or, on the contrary, would be maintained or decrease to normal values. Figure 7 illustrates this idea.

![Graph illustrating assessment of hypertension risk](image)

The procedure started by identifying the patients that had recently shown blood pressure values in a critical range, more specifically, that had presented blood pressure values in the range $[-5\%,+5\%]$ of the limit value of 135 mmHg during 3 consecutive days. Then, for those patients, the blood pressure values of the following week were predicted. According to the percentage of values that were above the limit threshold (135 mmHg), the risk of the patient was assessed: if the percentage was higher than 75%, the patient was considered to be at risk of developing an hypertension episode condition; in the other case (less than 75%), the patient was considered to have no risk. Table I shows the discrimination capability of the method.

<table>
<thead>
<tr>
<th>Actual class</th>
<th>No risk</th>
<th>In risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted class</td>
<td>No risk</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>In risk</td>
<td>6</td>
</tr>
</tbody>
</table>

Table I. Hypertension Events
Using MyHeart dataset, 150 experiments were performed using random templates, in which 35 BP signals exhibited values in the critical range. From these experiments, the obtained values of sensitivity (SE) and specificity (SP) were of, respectively, 85.7% and 91.8%. These values demonstrate the potential of the trend prediction strategy.

2. Risk of developing of arrhythmic episodes
A set of experiments was carried out particularly applied to patients whose heart rate values were in a critical range (around the threshold of tachycardia). The main goal was to determine whether during the following week the heart rate signal of a patient would evolve towards tachycardia values or, on the contrary, would be maintained or decrease to normal values. The procedure is the same that was applied to hypertension case (HR limit value of 100 bpm during three consecutive days). The effectiveness of the proposed strategy was tested by selecting, from a set of 600 random templates, the ones that verified the referred requirement (to be in the critical range). In effect, 58 verified this condition: in 26 cases the patient presented risk of developing a tachycardia episode, and in 32 cases the patient revealed no risk. Table II shows the discrimination capability of the method.

<table>
<thead>
<tr>
<th>Actual class</th>
<th>No risk</th>
<th>In risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted class</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No risk</td>
<td>28</td>
<td>10</td>
</tr>
<tr>
<td>In risk</td>
<td>4</td>
<td>16</td>
</tr>
</tbody>
</table>

To quantify the validity of the method, the sensitivity and specificity were determined, resulting in a SE of 61.5% and a SP of 87.5%.

Although it was not possible to compare these results with other works, considering that the prediction involved fully random templates, the obtained SE and SP values for both cases were very satisfactory. In effect, these metrics demonstrate the potential of the trend prediction strategy.

4. CONCLUSIONS
The main goal of Heartways project, (FP7-SME-315659), was the development of an advanced modular solution for supporting cardiac patients in rehabilitation outside an hospital center, with the aid of wearable sensors and intelligent algorithms that personalize the management and the follow up for patients and professionals. One of the objectives at the algorithm level (Multi-parametric Trends Analysis and Events Prediction Algorithms) was the development of methodologies for biosignal prediction, mainly to support the early detection of critical events (such as hypertension episodes based on the evolution of blood pressure, and arrhythmic episodes based on the evolution of heart rate).

This paper proposed a strategy based on wavelet decomposition for the prediction of biosignals, which goal was to estimate signals’ future evolution trend. The capability of the proposed methodology was in a first phase compared with other common prediction mechanisms. Then, using the predicted values, the scheme was tested in the assessment of hypertension and arrhythmic events in patients whose blood pressure/heart rate values were in a critical range. For the effect, real data collected by the tele-monitoring study under myHeart project was used. The obtained values of sensitivity and specificity reveal the capacity of the strategy.

REFERENCES