35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

# An effective wavelet strategy for the trend prediction of physiological time series with application to pHealth systems

*T. Rocha[1]†, S. Paredes†, P. de Carvalho‡, J. Henriques‡*

**Abstract - This work proposes a wavelet decomposition based scheme to estimate the evolution trend of physiological time series. The scheme does not involve the explicit development of a model and is essentially supported on the hypothesis that future evolution of a biosignal can be estimated from similar historic patterns. The strategy considers an a-trous wavelet decomposition, where the most representative trends are extracted from the historic similar patterns. Then, a set of distance-based measures able to assess the prediction likelihood of each representative trend, is introduced. From these measures and through an optimization process, a subset of these trends is selected and aggregated to derive the required time series evolution trend.**

**The effectiveness of the methodology is validated in the prediction of blood pressure signals collected in two telemonitoring studies: TEN-HMS and MyHeart. Additionally, Friedman and Nemenyi statistics tests are implemented to rank several methods, confirming the value of the proposed strategy.**

## I. Introduction

It is known that health is linked to behavior and lifestyle. Therefore, it is recognized by clinical professionals that the focus should be on prevention, as the best method to avoid diseases from happening. In this context, phealth and telemonitoring solutions are making a huge impact on preventive medicine. The patient is at the center of the health delivery process and, through remote monitoring and management applications, pHealth systems aim the continuity of care at all levels of health care delivery. By enabling remote patient monitoring, together with adequate diagnosis and prediction methodologies, they are of extreme significance for the conception of early prevention systems, providing professionals with adequate tools to diagnose and predict the occurrence of severe events.

The main topic of the present work is the research of methodologies for time series prediction, mainly to support the early detection of critical events. It is founded on the hypothesis that the estimation of biosignals future evolution can be supported on current and past measurements taken from historical data of a group of patients, including the patient under study. To this aim two main stages are considered: *i*) in the first, the selection of patients who display similar behaviors in their physiological time series is carried out by means of a similarity analysis process; *ii*) then, in the second stage, an estimation of the biosignal's future values is performed, based on the similar time series identified in the first phase.

Significant advances have been made in the development of methods for the determination of similarities in time series. The simplest algorithms used the Euclidean distance between raw time series of the same length. Others proposed dynamic time warping for time series of different lengths [1]. Nevertheless, due to the high dimensionality of time series, most of the approaches perform dimension reduction on data. Among them, some used discrete Fourier transform [2], principal component analysis (Karhunen-Loève transform), [3], and wavelet transform [4]. Rocha *et al* [5] presented an innovative measure able to efficiently evaluate the similarity between two physiological time series. It combines the Haar wavelet decomposition, in which signals are represented as a combination of a set of orthogonal basis, with the Karhunen-Loève transform, allowing for an optimal reduction of that set of basis. Using an iterative algorithm for computing the corresponding coefficients, the computational complexity of the method was significantly decreased.

Among prediction techniques, linear regression methods, such as autoregressive structures, have been the most used in practice. However, linear models are usually inadequate for biosignals, since, in practice, these are non-linear to some extent. Among the non-linear methods, neural networks became very popular mainly due to their universal approximation properties. Many different types of neural networks, such as time delay and recurrent neural networks, have been proven to be effective for time series modeling and prediction [6]. On the other hand, in most clinical cases, an assumption of global stationarity can not be considered. Among time-frequency methods, wavelet transform, which can produce a good local representation of the signal in both the time and frequency domains, offering an appropriate framework to deal with the non-stationarities, has been applied. Although the wavelet transform itself is not a forecasting methodology, it may be incorporated in hybrid prediction schemes involving the multi-resolution decomposition of signals [7].

This work, being the follow-up of the one proposed by the same authors [5] in the context of similarity analysis, presents a strategy based on wavelet decomposition for the prediction of biosignals. In this procedure no explicit model is involved, and the goal of the methodology is not to perform an accurate prediction, but to obtain a reasonable forecast of the future trend. Basically, from the wavelet decomposition of similar signals, the most appropriate trends at each decomposition level are identified and combined through an optimization process, directly providing an estimation of the current time series evolution.

The structure of the paper is as follows: section 2 describes the proposed wavelet multi-resolution scheme and section 3 presents its application to blood pressure signals using data collected during TEN-HMS and MyHeart telemonitoring studies. Finally, in section 4, some conclusions are drawn.

This work was partially financed by iCIS (CENTRO-07 -ST24-FEDER-002003), HeartCycle EU project (FP7-216695) and CISUC (Center for Informatics and Systems of University of Coimbra).

† Instituto Politécnico de Coimbra, Departamento de Engenharia Informática e de Sistemas, Portugal, {teresa@isec.pt, sparedes@isec.pt}.

‡ CISUC, Departamento de Engenharia Informática, Universidade de Coimbra, Coimbra, Portugal, {carvalho@dei.uc.pt, jh@dei.uc.pt}.

## II. METHODOLOGIES

The Figure 1 illustrates the concept behind the proposed prediction approach.

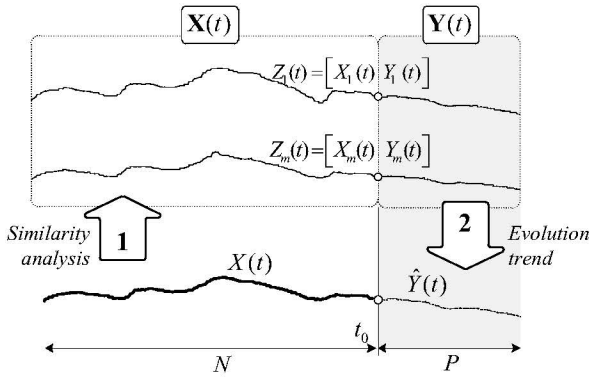


*Figure 1 – Prediction strategy based on similar signals.*

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 conditions (patterns) $\mathbf{X}(t) \equiv \{X_m(t) \in \mathbb{R}^{1,N}\}$, $m = 1,...,M$, is identified [5]. From these, the corresponding subsequent $P$ future values, $\mathbf{Y}(t) \equiv \{Y_m(t) \in \mathbb{R}^{1,P}\}$, are straightforwardly obtained (known past values from historic dataset). Then, the known "*future*" evolution of the identified patterns, $\mathbf{Y}(t) \equiv \{Y_m(t)\}$, can be used in a prediction mechanism to estimate the evolution of the current template, $Y(t)$. The global set of patterns, $\mathbf{Z}(t) \in \mathbb{R}^{M,N+P}$, is therefore composed of two components $\mathbf{X}(t)$ and $\mathbf{Y}(t)$, in the form of (1).

$$\mathbf{Z}(t) = [\ \mathbf{X}(t)\ \ \mathbf{Y}(t)\ ] \tag{1}$$

The Figure 2 depicts the main steps involved in the estimation of $Y(t)$, based on the similar patterns $\mathbf{Z}(t)$ and trough a wavelet decomposition scheme.

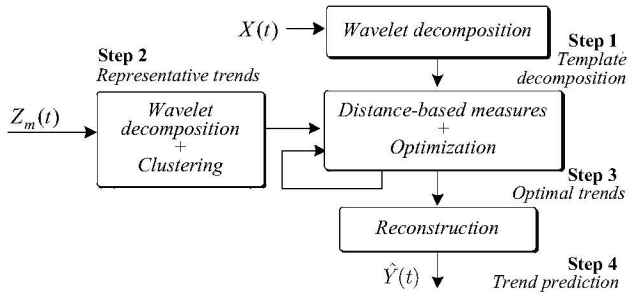


*Figure 2 – Prediction methodology.*

**Step 1. Template decomposition**

In the first step the template $X(t)$ is decomposed using the Haar à-trous wavelet transform [8], a shift-invariance transform. As result, assuming for instance three levels of decomposition, the template can be obtained at the time instant $t_0$ as the sum of details $d^l X(t)$, for scales 1 to 3, plus the trend $a^3 X(t)$, as given by (2).

$$x(t_0) = d^1 x(t_0) + d^2 x(t_0) + d^3 x(t_0) + a^3 x(t_0) \tag{2}$$

For a general case, considering $L$ levels of decomposition, the wavelet decomposition is described by $W\{X\} \in \mathbb{R}^{L+1,N}$, according to (3).

$$W\{X\} = \{d^l X(t), a^L X(t)\},\ l = 1,..,L \tag{3}$$

**Step 2. Representative trends**

The second step involves the determination, at each decomposition level, of the most representative time series (trends) from the retrieved similar historic signals. To achieve this goal, the historic signals are in a first phase decomposed using the à-trous wavelet, according to (4).

$$W\{Z_m\} = \{d^l Z_m(t), a^L Z_m(t)\},\ l = 1,..,L \tag{4}$$

The variables $a^L Z_m(t)$ and $d^l Z_m(t)$, $m = 1,\ldots,M$, represent, respectively, the approximation and the details. It is important to note that, in this case, the decomposition can be extended to the "future" (time instants from $N+1$ to $N+P$), with length $N+P$, that is, $W\{Z_m\} \in \mathbb{R}^{L+1,N+P}$. Then, in a second phase, the representative decomposition trend at each level is determined through a clustering process. In this case the subtractive method was employed [9].

$$d^l \overline{Z}(t) = subCustering\ \{d^l Z_m(t)\},\ m = 1,..,M,\ l = 1,...,L \tag{5}$$

$$a^L \overline{Z}(t) = subClustering\ \{a^L Z_m(t)\},\ m = 1,..,M \tag{6}$$

The variables $d^l \overline{Z}(t) \in \mathbb{R}^{1,N+P}$ and $a^L \overline{Z}(t) \in \mathbb{R}^{1,N+P}$ denote, respectively, the representative details and approximation.

**Step 3. Optimal trends**

In this step, the representative trends are reduced to an optimal set, that is, to a set of trends (decomposition levels) that have the potential to contribute to a consistent prediction. To this purpose, a combination process comprising the minimization of a set of distance-based measures, that assess the likelihood that a representative trend will contribute to a correct estimation, is implemented.

*i. Distance-based measures*

The distance-based measures are computed for each decomposition level $l = 1,..,L+1$, where $L+1$ stands for the approximation, using: *i*) the template $X(t) \in \mathbb{R}^{1,N}$; *ii*) the corresponding wavelet decomposition at $l$ level, $d^l X(t) \in \mathbb{R}^{1,N}$; *iii*) the wavelet decomposition of similar patterns at the same level, $d^l Z_m(t) \in \mathbb{R}^{1,N+P}$, $m = 1,..,M$; *iv*) the corresponding clustering, that is, the representative trends $d^l \overline{Z}(t) \in \mathbb{R}^{1,N+P}$. Using these signals, a set of distance-based measures $\theta_i^l$ is computed as follows, for $m = 1,..,M$:

$$\theta_1^l = S(X(t), d^l X(t)), \qquad t = 1,\ldots,N \tag{7}$$

$$\theta_2^l = S(d^l X(t), d^l \overline{Z}(t)), \qquad t = 1,\ldots,N \tag{8}$$

$$\theta_3^l = mean\left[S(d^l X(t), d^l Z_m(t))\right],\ t = 1,\ldots,N \tag{9}$$

$$\theta_4^l = exp\left[-std\ \{S(d^l X(t), d^l Z_m(t))\}\right],\ t = 1,\ldots,N \tag{10}$$

$$\theta_5^l = mean\left[S(d^l \overline{Z}(t), d^l Z_m(t))\right], \qquad t = N+1\ldots,N+P \tag{11}$$

$$\theta_6^l = exp\left[-std\ \{S(d^l \overline{Z}(t), d^l Z_m(t))\}\right], \qquad t = N+1,\ldots,N+P \tag{12}$$

As result, a vector composed of six measures is obtained, $\Theta^l = \left[\theta_1^l, \theta_2^l, ..., \theta_6^l\right]$, where each $\theta_i^l \in [0,1]$.

The measure $S(X_1(t), X_2(t)) \in [0,1]$, is a normalized similarity measure, where a value of 1 of means a total agreement between the signals. The operators *mean*$(\cdot)$ and *std*$(\cdot)$ denote, respectively, the mean and standard-deviation operators.

*ii. Selection of the optimal trends*

The optimization strategy assumes that each of the parameters $\theta_i^l$ defines a measure, that enables to assess the quality of each representative trend. Therefore, according to the obtained measures, a decision regarding the inclusion or exclusion of a specific representative trend in the optimal set can be taken. To support this decision the operators *maximum*$(\cdot)$ and *product*$(\cdot)$ are employed, respectively, as the aggregation and conjunction operators. As result, the quality of a specific prediction is assessed according to (13).

$$p(\Theta) = \max\{p(\Theta\sigma_1),...,p(\Theta\sigma_i),...,p(\Theta\sigma_n)\} \quad (13)$$

Each variable $p(\Theta\sigma_i)$ denotes the possible decomposition level combinations, resulting from the operator $C(nL, nN)$ (combinations of $nL$ taken $nN$ at a time). In this process, the conjunction of the metrics corresponding to a specific level is given by the *product*$(\cdot)$ operation (14).

$$p(\Theta^l) = \prod_{i=1}^{6} \theta_i^l \quad l = 1,...,L+1 \quad (14)$$

If two levels $L_i$ and $L_j$ are combined, the corresponding aggregation $p(\Theta\sigma_i)$ is obtained according to (15).

$$p(\Theta\sigma_i) = \max\left\{p(\Theta^{L_i}), p(\Theta^{L_j})\right\} \quad (15)$$

**Step 4. Trend prediction**

Finally, the optimal trends resulting from the optimization process are combined to obtain the trend prediction corresponding to the template $X(t)$, as (16).

$$Y(t) = a^{\sigma}\overline{Z}(t) + \sum d^{\sigma}\overline{Z}(t) \qquad t = N+1,..,N+P \quad (16)$$

where the subscript $\sigma$ denotes the optimal trends identified by the optimization process.

## III. Results

The present section focuses on the analysis of blood pressure (BP) signals daily collected by two telemonitoring platforms: TEN-HMS [10] and MyHeart [11]. The main goal is to assess the effectiveness of the presented predictive strategy to estimate the evolution trend of such signals.

### A. Prediction methods

The performance of the proposed wavelet multi-resolution scheme (*WMM*) is also compared with other typical prediction strategies, namely the autoregressive integral moving average model (ARIMA), and two non-linear regression models, the generalized regression neural network (GRNN) and the support vector regression (SVR).

***ARIMA***: Regarding the ARIMA model, experiments using BP signals with degree of differencing $d = 1,2$ and regressive and moving average orders $n_a = 1,2,3$ and $n_c = 1,2,3$, resulted in $ARIMA(n_a, d, n_c) = ARIMA(2,1,2)$.
The respective parameters were estimated using a least mean square method, by means of the *armax*$(\cdot)$ Matlab command[2].

***GRNN***: The *newgrnn*$(\cdot)$ Matlab command was used to implement the neural model, where the width of the kernels was experimentally determined as $\lambda = 0.2$.

***SVR***: The implementation of the SVR method was done through the *libsvm* framework[3]. The training data was previously normalized to the range $[-1,1]$ and the parameters were: *i*) type of SVR: *epsilon-SVR; ii)* kernel type: *radial basis function; iii)* width of kernel function, $\gamma = 0.5$ ; *iv*) cost, $C = 1$; *v*) tolerance of termination criterion, $\varepsilon = 0.001$.

***AVP***: *Average of the patterns* - the fourth method simply considers the average value of historic signals $Y_m(t)$, as the estimation of $Y(t)$.

***WMM***: Finally, the last method implements the wavelet multi-resolution strategy proposed in section 2.

### B. Comparison of the prediction methods

The accuracy of the methods can be determined using common prediction metrics, such as NRMSE and MAPE. Here, the well-known Pearson correlation (values in the range $[-1,1]$) was implemented. Additionally, the Friedman and Nemenyi tests were applied to compare the several predictors [12]. The Friedman test compares schemes' average ranks to decide whether to reject the null-hypothesis, which states that all the schemes are equivalent. If the Friedman test rejects the null-hypothesis, the post-hoc Nemenyi test can be applied to rank the prediction schemes and indicate whose performances have statistically significant differences.

### C. Experiments with the TEN-HMS and MyHeart datasets

**1. Data pre-processing**

Firstly, a pre-processing procedure is applied to the original BP signals. The result is a historic dataset with a value per day (sampling rate of 1 day). This process also involves: *i*) averaging of values (in case two measurements have been performed in the same day), *ii*) dealing with missing values and, *iii*) noise reduction.

A forecast period of approximately one week (eight days, $P$=8) was stipulated. In terms of the length of the template, that is, the past information used in the prediction, the value suggested by the physicians was about a month ($N$=32). Moreover, the values of $N$ and $P$ are powers of two, to simplify the wavelet transform operations.

The number of retrieved patterns from the historic dataset was $M$=5 and the level of wavelet decomposition was $L$=5. Only the patients for whom there were BP measurements in, at least, 90 days (3 months) were selected, resulting in a total of 51 and 100 patients, respectively for TEN-HMS and MyHeart datasets.

**2. Comparison of prediction results**

To compare the proposed prediction method against the other strategies, a set of 300 and 500 templates of length $N$=32 were randomly selected, respectively, from TEN-HMS and MyHeart datasets. The Figure 3 depicts the box-plots corresponding to the Pearson correlation computed between the actual and the predicted values by ARIMA, GRNN, SVR, AVP and WMM.

2 Matlab 7.10.0.499 (R2010a), The MathWorks Inc.

3 www.csie.ntu.edu.tw/~cjlin/libsvm/

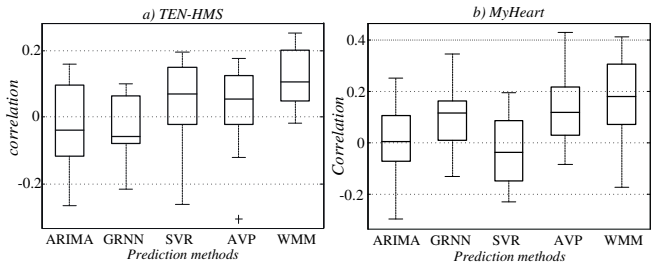


*Figure 3 − Comparison of the prediction methods.*

From the analysis of Figure 3 and, in global terms, it appears that the proposed method is superior to the others, both with TEN-HMS and MyHeart datasets. In effect, the wavelet based prediction method (WMM) presents the highest median, however, a higher variability than some of the other prediction methods. Three of the methods (ARIMA, GRNN, and SVR) compute the prediction based on an iterative approach: a one-step ahead model is iteratively applied during *P* times, being the current predictions used by the model in order to obtain the next forecast. The last two methods (AVP and WMM) do not involve the explicit computation of a model. Thus, they are, to some extent, similar to a direct approach. This fact can justify why GRNN and ARIMA present poor results, in particular for the TEN-HMS dataset.

In order to accurately compare the predictive methods, the Friedman test was implemented. The Table 1 shows the average values obtained in the ranking of the methods (each method was ranked in the range {1,..,5}, according to its capacity to predict the template future evolution)

*Table 1- Comparison of the prediction methods.*

| | ARIMA | GRNN | SVR | AVP | WMM |
|---|---|---|---|---|---|
| ***TEN-HMS*** | 2.50 | 2.10 | 3.20 | 3.00 | 4.20 |
| ***MyHeart*** | 2.15 | 3.35 | 2.1 | 3.60 | 3.80 |

From the average of ranks, the value of qui-square was $\chi_F^2 = 30.32$ and $\chi_F^2 = 21.24$ , respectively for TEN-HMS and MyHeart. As result, the null hypothesis "*Ho: all the methods behave similarly*" was rejected for both datasets with a high significance level (respectively with a *p-value*=0.0004 and 0.0003, for TEN-HMS and MyHeart). As a consequence, the Nemenyi test was performed to compare the methods based on the computed average ranks, which results are presented in Table 2.

*Table 2 - Nemenyi test.*

| | | GRNN | SVR | AVP | WMM |
|---|---|---|---|---|---|
| **ARIMA** | *TEN-HMS* | -0.40 | 0.70 | 0.50 | 1.70 *** |
| | *MyHeart* | 1.20 | -0.05 | 1.45 ** | 1.65 *** |
| **GRNN** | *TEN-HMS* | | 1.10 | 0.90 | 2.10 *** |
| | *MyHeart* | | -1.25 * | 0.25 | 0.45 |
| **SVR** | *TEN-HMS* | | | -0.20 | 1.00 |
| | *MyHeart* | | | 1.50 ** | 1.70 *** |
| **AVP** | *TEN-HMS* | | | | 1.20 |
| | *MyHeart* | | | | 0.20 |

** at a significance level of 10%, ** at a significance level of 5%, *** at a significance level of 1%*

In these particular experiments, by means of the Nemenyi test it is possible to conclude that two methods are significantly different at levels of 1%, 5% and 10%, if their average ranks differ at least the critical value, respectively, 1.9902, 1.6861 and 1.5317. Thus, from the Table 2, it can be concluded that the proposed WMM method outperforms the ARIMA and the GRNN predictors at the level of 1% for the TEN-HMS dataset. In case the of the MyHeart dataset, the WMM method outperforms ARIMA and SVR at the level of 1%, and the GRNN at the level of 10%. Thus, it can be concluded that the proposed method is globally superior to the ARIMA, GRNN and SVR, but comparable with the AVP method.

## IV. Conclusions

This work proposed a wavelet decomposition based scheme to efficiently estimate the evolution trend of physiological time series. By means of a similarity analysis procedure, a set of signals presenting a dynamics similar to the current time series, is retrieved from the historic. From the wavelet decomposition of these signals the most representative trends are extracted at each decomposition level and combined through and optimization process, from which the evolution of the current signal is straightforwardly obtained. The scheme was successfully implemented using TEN-HMS and HeartCycle telemonitoring studies, to the trend estimation of blood pressure signals.

Future work will be devoted to extend this strategy to multi-parametric settings.

## References


[1] Park S., et al; *Efficient searches for similar subsequences of different lengths in sequence databases*; Proc. of the International Conference of Data Engineering, pp 23-32, 2000.

[2] Agrawal R., et al; *Efficient similarity search in sequence databases*; Proc. of the Intl. Conf. on Foundations of Data Organizations and Algorithms; pp 69–84, 1993.

[3] Yang K., Shahabi C.; *A PCA-based Similarity Measure for Multivariate Time Series*; Proc. of the 2nd ACM international workshop on Multimedia databases, 2004.

[4] Saeed M., Mark R.; *A Novel Method for the efficient retrieval of similar multi-parameter physiologic time series using wavelet-based symbolic representations*; AMIA symp., 2006.

[5] Rocha, et al; *An Efficient Strategy for Evaluating Similarity between Time Series based on Wavelet / Karhunen-Loève Transforms*; IEEE EMBS, San Diego, pp. 6216- 6219, 2012.

[6] Haykin S.; Neural networks and learning machines; Third Edition. Prentice Hall, 2008.

[7] Fryzlewicz et al; *Forecasting non-stationary time series by wavelet process modeling*; ISM., v55, n4, 737–764, 2003.

[8] Renaud et al; *Prediction based on a multiscale decomposition*; Int. Journal of Wavelets, Multiresolution and Information Processing, V. 1, N. 2, pp. 217-232, 2003.

[9] Kriegel et al.; *Density-based clustering*; WIREs Data Mining and Knowledge Discovery, 1 (3): 231–2, 2003.

[10] Cleland et al*.; Noninvasive home telemonitoring for patients with heart failure at high risk of recurrent admission and death*; J Am Coll Cardiol, 45:1654-1664, 2005.

[11] Habetha, J.; *MyHeart - A new approach for remote monitoring and management of cardiovascular diseases*; IEEE 2006..

[12] Demsar, J; *Statistical Comparisons of classifiers over Multiple Data Sets*; Machine Learning Research, 7, pp 1-30, 2006.