33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

A Wavelet-based Approach for Time Series Pattern Detection and Events Prediction Applied to Telemonitoring Data

T. Rocha, S. Paredes, P. Carvalho, J. Henriques

Abstract - **This work aims the development of a predictive strategy able to estimate future events with relevant impact in the cardiovascular status.**

Based on wavelet transform, a new time series similarity metric is introduced, which is capable to detect a pre-defined pattern in time series data. In addition, a methodology combining a wavelet scheme with state space multi-models is proposed to achieve the prediction of future signal values.

Blood pressure signals, collected by a telemonitoring platform (TEN-HMS), are used to detect the occurrence of future hypertension events.

I. Introduction

Cardiovascular (CV) diseases are the leading group of conditions that cause death worldwide. The World Health Organization estimates that 17.5 million people died of cardiovascular diseases in 2005, representing 30% of all global deaths. High blood pressure or hypertension, one of the leading public health problems, is among the top most factors associated with cardiovascular diseases.

In fact, uncontrolled and prolonged elevation of blood pressure (BP) can lead to a multiplicity of alterations in the myocardial structure, coronary vasculature, and conduction system of the heart, which can lead to the development of left ventricular hypertrophy, coronary artery disease, myocardial infarction, cardiac arrhythmias, heart failure (HF), among others [1]. HeartCycle European project aims to improve the quality of life for patients with coronary artery disease or heart failure, by monitoring their condition and involving them in the daily management of their disease [2]. Integrated in HeartCycle, the Medical Risk Assessment module is the responsible for the research and development of models to assess CV risk and status of the referred patients. Basically, these models assume that CV status *i*) is continually updated using measurements, parameters and symptoms, collected during daily home monitoring process, and *ii*) it may be characterized based on specific cardiovascular conditions. Examples of these are hypertension, myocardial ischemia, arrhythmias, pulmonary edema, etc., which are themselves defined through literature or by clinical expertise. In this context, relevant conditions have already been addressed by the authors, such as arrhythmias [3], [4] and myocardial ischemia [5]. Moreover, analysis of blood pressure signals, in particular the prediction of acute hypotensive episodes in intensive care units, has been recently addressed by the authors [6].

T. Rocha and S. Paredes - Departamento de Engenharia Informática e de Sistemas, Instituto Superior de Engenharia de Coimbra, Portugal, {teresa, sparedes}@isec.pt.; P. Carvalho and J. Henriques - CISUC, Departamento de Engenharia Informática, Universidade de Coimbra, Portugal, {carvalho, jh}@dei.uc.pt.

The present work focus on the analysis of BP signals, collected by means of a telemonitoring platform, applied to the prediction of hypertension events. To achieve this goal, a generic methodology consisting of two phases is considered. In the first, a time series pattern detection methodology is developed to identify past similar situations. In the second, these similar situations are used to derive a prediction scheme that will estimate future values of BP and, consequently, will predict hypertension events occurrence.

In terms of pattern detection in time series, several approaches have been proposed. The simplest time-domain algorithms used Euclidean distance to calculate a similarity metric between time series. Others [7] proposed the warping distance which deals with changes and shifts in time scale. Nevertheless, due to the high dimensionality of time series, most of the approaches perform dimension reduction on time series data. In effect, some works used discrete Fourier transform, others [8] employed the singular value decomposition transform, while others [9] used piecewise aggregate approximation. Works based on discrete wavelet transform (DWT) have also been proposed [10]. In effect, DWT provides an attractive means for developing multiresolution representations of signals suitable for patterns detection algorithms. Part of its success can be justified by the inherent ability of wavelet representations to reveal the superposition of different structures occurring in signals on different time scales at different times.

In terms of prediction, where the basic idea involves the development of models that estimate future values of a signal based on its past values, linear autoregressive models, like autoregressive (AR) and moving average mappings, are well known examples of such [11]. Besides these, a large number of non-linear regressive mappings have been proposed for prediction tasks, namely fuzzy systems, neural networks, multi-models and phase space reconstruction techniques [12]. Several types of transform have also been applied for time series forecasting, such as principal component analysis [13], independent component analysis [14], Fourier transform and wavelet transform [15]. In fact, wavelet analysis allows the extraction of features that characterize the dynamics of the signal trends at several scales, enabling to efficiently deal with the prediction problem [16].

The major contributions of the present work are: (*i*) a new metric based on wavelet methodology that efficiently assesses similarities between time series; (*ii*) a prediction scheme that combines wavelet decomposition with state space multi-models to estimate the evolution of signals. Contrasting with classical autoregressive representations, multi-model schemes do not recursively use model outputs as inputs for step ahead predictions. As result, prediction

errors are not propagated over the forecast horizon and long-term predictions can be accurately estimated. On the other hand, wavelet decomposition introduces robustness (denoising) as well as improves accuracy since prediction is independently considered for each frequency sub-band of the signal [17]. The effectiveness of the proposed strategy was validated by applying it to blood pressure signals, collected as part of the TEN-HMS study [18], in the prediction of hypertension events.

The remainder of this paper is organized as follows. In the next section, the proposed methodology is described. In section 3, the achieved results using the TEN-HMS dataset are presented and discussed. Finally, in section 4, some conclusions are drawn.

II. PROPOSED METHODOLOGY

The proposed methodology consists of two main phases, as depicted in Figure 1.

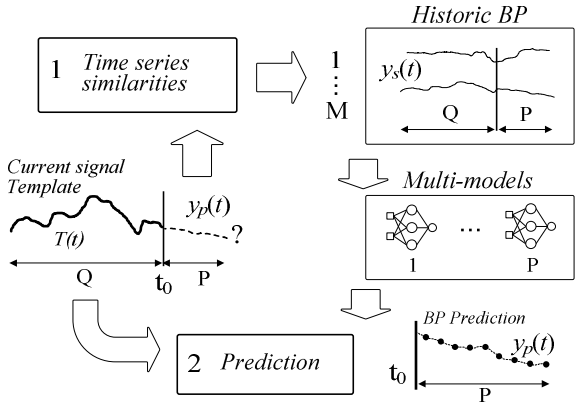

Figure 1. Hypertension detection scheme: time series similarities and prediction.

In the first phase, the last Q collected values of the BP signal are considered as a template $T(t)$ to be compared with the historic BP signals. Using the proposed wavelet-based time series similarity metric, a set of M segments, $y_S(t)$, that best match the template is identified. In the second phase, these M segments and the respective following P samples are employed to obtain the parameters of a state space multi-model structure. In this structure, a different model is used to independently predict each P sampling instant.

A. Time series patterns detection

1. Wavelet template decomposition

In a first stage the template signal $T(t) \in \Re^{1,Q}$, where t represents the discrete temporal index, is decomposed using DWT. Basically, this hierarchically decomposes a time series sequence in terms of an approximation of the original sequence, plus a set of details that range from coarse to fine. The main trend of the input sequence is preserved in the approximation part, while the localized changes are kept in the detail parts. The original signal can be reconstructed as described by (1), where $D_L(t)$ and $A_L(t)$ represent the detail and approximation coefficients at level L, respectively.

$$\begin{aligned} T(t) &= A_1(t) + D_1(t) = A_2(t) + D_2(t) + D_1(t) \\ &= \ldots \\ &= A_L(t) + D_L(t) + D_{L-1}(t) + \ldots + D_1(t) \end{aligned} \quad (1)$$

Taking into account the last expression of (1), approximation $A_L(t)$ and details $D_i(t)\ \ i = 1,..,L$ can be seen as a set of basis $\varphi_i(t)\ \ i = 1,..,L+1$ from which is possible to describe the original signal. In effect, $T(t)$ can be represented as a linear combination of these basis functions by means of a set of (L+1) coefficients $\hat{C} \in \Re^{1,L+1}$.

$$T(t) = \sum_{i=1}^{L+1} \hat{c}_i \ \varphi_i(t) \quad (2)$$

Considering equations (1) and (2), template's coefficients are $\hat{c}_i = 1$.

2. Similarity measure

The similarity searching procedure makes use of a windowing scheme to compute the correlation between the template $T(t)$ and the signal being analyzed $y(t) \in \Re^{1,N}$.

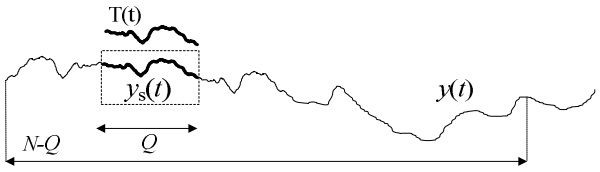

Figure 2. Similarity measure estimation.

The similarity measure is estimated for each segment, $y_s(t) \in \Re^{1,Q}$, thus a set of (N-Q) correlation values are obtained. First, each segment $y_S(t)$ is described using the set of $\varphi_i(t)$ basis functions that were derived from wavelet decomposition of the template $T(t)$, using (3).

$$y_s(t) = \sum_{i=1}^{L+1} c_i \ \varphi_i(t) \quad (3)$$

The coefficients c_i can be easily calculated based on a Least Mean Square Error problem formulation. In fact, equation (3) can be written as

$$y_s(t) = C \ \Phi(t) \quad (4)$$

Where $C \in \Re^{1,L+1}$ is a vector composed of the (L+1) coefficients and $\Phi(t) \in \Re^{L+1,Q}$ is a matrix composed of the basis functions. Using the pseudo-inverse $(\dagger)$ of $\Phi(t)$, coefficients can be obtained as follows:

$$C = y_s(t) \ \Phi(t)^{\dagger} \quad (5)$$

In the present work it is assumed that the differences (distance) between the coefficients $\hat{C}$ and C can be used as a similarity measure. Moreover, when the basis functions are orthogonal these coefficients are unique. Therefore, orthogonal wavelets, such as the Haar wavelet, are suitable to be employed in this context. Given this, the similarity between template $T(t)$ and segment $y_S(t)$ is computed as a distance between the two vectors of coefficients,

$$correlation\ \left(T(t), y_s(t)\right) = \text{dist}(\hat{C},C) \quad (6)$$

Although any distance can be used, in the present work it is computed as

$$\text{dist}(\hat{C},C)=\frac{1}{L+1}\sum\left\{\exp^{-k|\bar{C}-Cl|}\right\}=\frac{1}{L+1}\sum\left\{e^{-k|1-Cl|}\right\} \quad (7)$$

This approach generates a normalized correlation measure in the interval [0..1], which enables to easily identify the patterns that best match the template. The value of zero indicates a very low similarity, while the value of one corresponds to the maximum similarity (same coefficients). The k value, equation (7), controls the aperture of the Gaussian, enabling to exponentially weigh distances in order to better discriminate correlations.

It should be noted that this technique has the advantage of being very efficient. Remind that the inversion of the basis functions matrix is performed only once for all searches and, as a result, the coefficients are simply obtained through a matrix multiplication. Thus, it is possible to obtain correlation values for all N-Q segments of a signal $y(t)$, efficiently and accurately.

B. Prediction Models

1. Multi-models scheme

For the prediction of time series two main strategies can be followed: direct and recursive approaches. The last only use one model to recursively predict all future instants. Direct predictors imply the construction of a different model for each prediction instant (multi-models). Following a direct approach a future value at time instant P, $y_P(t+P)$, can be estimated by equation (8), where the mapping $f_P:\Re^n\rightarrow\Re$ identifies the particular P^{th} model.

$$y_p(t+P\,|\,t)=f_P\left(y(t),y(t-1),...,y(t-n+1)\right) \quad (8)$$

2. Sub-model structure

The selection of a specific sub-model involves the characterization of the function $f_i(\cdot)$ and of the number of past observations to be considered. Using wavelet transform a scalar time series signal can be decomposed into several W scale sequences, thus a time series vector can be obtained by grouping for each instant the scales values. This time series vector can be interpreted as a state, used for both modeling and prediction tasks. Then, the final forecast of the original scalar time series can be obtained through the inverse wavelet transform. As researched by some authors [19], the accuracy of the final prediction results using this approach is superior than predicting the original time series directly.

Using this formulation (state space) the prediction of a P step ahead value can be directly obtained by (9)

$$x(t+P)=f_P\left(x(t)\right) \quad (9)$$

where $x(t)\in\Re^{1,W}$ is the time series vector with the same dimension as the number of decomposition levels (W) at discrete-time t. In the simplest case, if a linear relationship is assumed, the function $f_P(\cdot)$ can be replaced by a matrix $A_P\in\Re^{W,W}$. In this case, it is possible to estimate the A_P matrix based on well-known system theory identification techniques, such as Least Mean Squares Error or Kalman filter strategies.

III. Results

A. TEN-HMS dataset

The Trans-European Network Homecare Monitoring Study (TEN-HMS) was designed to assess whether home based telemonitoring could reduce morbidity and mortality in patients with heart failure, compared with usual care or regular telephone contact. In this study, a total of 426 patients with a recent admission for HF and Left Ventricular Ejection fraction <40%, were assigned randomly to home telemonitoring (168), nurse telephone support (173) and usual care (85). Particularly, home telemonitoring consisted of twice-daily self-measuring of weight, blood pressure, heart rate and rhythm, with automated devices linked to a cardiology center, during the period of two years.

For the present work, a subsection of the complete dataset containing data from 83 patients was made available. In terms of strategy validation, only blood pressure signals were employed. Furthermore, only patients for whom there were BP measurements in 150 consecutive days (5 months) were selected for this purpose, resulting in a total of 33 patients.

B. Patterns detection

In order to determine the size of the template to be searched (Q), several experiments were conducted in which different values for Q were considered, namely 7, 14, 21 and 28 days. According to the average errors, the selected value was Q=21 (three weeks). Given its orthogonally and simplicity properties, Haar wavelet was chosen for assessing signals similarity. The level of decomposition was L=4.

On the other hand, the number of segments (M) that best match the template was adaptively computed for each case, ranging from 1 to a maximum number of 10. In a first phase, 10 segments that presented the higher correlation with the template were obtained, as well as the respective future values. In a second step, this future values were compared themselves using the same correlation measure described above (7). The future segments that were more correlated were selected. This way, segments which evolution was dissimilar from the typical ones were discarded.

C. BP signal prediction

For each BP signal 9 different predictions were carried out, establishing as starting points of the forecast window the 40^{th}, 50^{th}, 60^{th}, …, 120^{th} days. Thus, a total of 33×9=297 prediction tasks were performed, considering a forecast horizon of P=7 days (one week). The Daubechies wavelet "db4" with a level of decomposition W=4 was used to generate the time series vectors (state space models). Each state space matrix model, A_P, p=1,…,7 was computed using a LMS approach based on the segments identified in the signal similarity detection phase.

Figure 2 shows the results using the proposed prediction procedure. In this case, given a template $T(t)$ to be predicted from instant t0 to instant t0+7, 4 segments were identified in a first phase. Using these segment (values before and after t0) P=7 different state space multi-models were afterward estimated, enabling to predict BP signal over the next week $y_p(t)$.

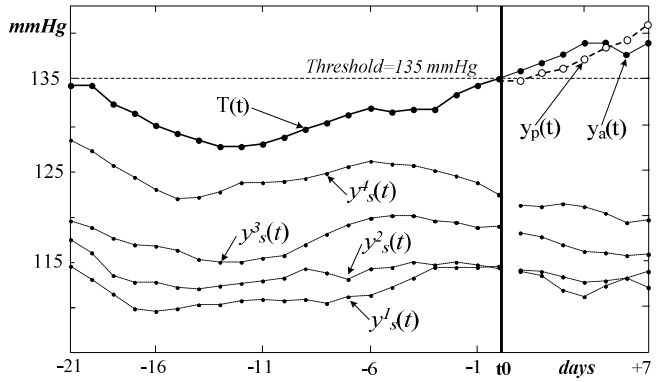

Figure 2. Prediction results using the proposed methodology.

The mean absolute percentage error (Mape) (10) and the well known Pearson correlation (Pc) between the estimated $y_p(t)$ and actual $y_a(t)$ signals, were computed.

$$Mape = \frac{1}{P}\sum_{i=1}^{P}\frac{\left|y_a(i)-y_p(i)\right|}{y_a(i)}\times 100 \qquad (10)$$

For the present example these values were Mape=1.72 and Pc=0.70. Considering all the predictions (297), the global values were Mape=(2.67±1.95) and Pc=(0.52±0.28) (average±standard deviation), revealing the capability of the methodology.

D. Hypertension events detection

Finally, hypertension events were detected, whenever patient had more than five days with arterial blood pressure higher than 135 mm Hg during the forecast period (a week). To assess the potentially of the strategy, a total of 138 examples were previously selected from the dataset: 48 corresponding to hypertension events and 90 to normal situations. Figure 2 depicts an example where a hypertension event has occurred (six days of BP higher than 135 mmHg have been predicted). Applying this algorithm to the 138 examples, a global sensitivity of 85.4% and a global specificity of 92.2% were achieved.

IV. CONCLUSIONS

This work proposed a methodology to predict hypertension events over a specific time period. Using arterial blood pressure time series, a wavelet-based strategy together with state space multi-models was implemented, enabling to estimate future values over a forecast horizon. Applied to BP signals, collected as part of the TEN-HMS study, the referred strategy allowed to adequately detect time series patterns and, then, to use these patterns to predict the occurrence of hypertension events.

Future work will address the extension of the strategy to more complex events, obtained from multi-parametric combination of several measurements. A particular topic to be researched is the HF decompensation episodes, to be validated using an available telemonitoring database.

Acknowledgements

This work was supported by HeartCycle [2], a European Community's Seventh Framework Programme, FP7-216695, and by CISUC, Center for Informatics and Systems of University of Coimbra, Portugal.

REFERENCES

[1] Riaz, K., Ahmed, A.; Hypertensive Heart Disease; Department of Internal Medicine, Section of Cardiology, Wright State University, http://emedicine.medscape.com/article/162449-overview

[2] Reiter, H., Maglaveras, N.; HeartCycle: Compliance and effectiveness in HF and CAD closed-loop management; EMBS 2009, Minneapolis, MN, pp 299 - 302, 2009.

[3] Rocha, T., Paredes, S., de Carvalho, P., Henriques, J., Antunes, M.; Phase space reconstruction approach for ventricular arrhythmias characterization; EMBS 2008, Vancouver, pp5470 - 5473, 2008.

[4] Henriques, J., P. Carvalho, M. Harris, M. Antunes, R. Couceiro, M. Brito, R. Schmidt; Assessment of Arrhythmias for Heart Failure Management; Phealth2008, Valencia, May 21-23, 2008.

[5] Rocha, T. S. Paredes, P. Carvalho, J. Henriques, M. Harris, J-Morais, M. Antunes; A lead dependent ischemic episodes detection strategy using Hermite functions; Biomedical Signal Processing and Control 5, 271–281, 2010.

[6] Henriques, J., Rocha, T.; Prediction of acute hypotensive episodes using neural network multi-models; Computers in Cardiology 2009, pp 549 - 552, 2009.

[7] Park S., Chu, W., Yoon J, Hsu C. Efficient searches for similar subsequences of different lengths in sequence databases. Proc. of the International Conference of Data Engineering, 23-32, 2000.

[8] Wu D., Agrawal, D., Abbadi, A., Singh A, Smith T.; Efficient retrieval for browsing large image databases; Proc. 5th international conference on Information and knowledge management, 11–18, 1996.

[9] Yi B., Faloutsos C.; Fast time sequence indexing for arbitrary Lp norms; Proc. of the 26th International Conference on Very Large Data Bases, 385-394, 2000.

[10] Popivanov I., Miller R.; Similarity Search Over Time-Series Data Using Wavelets.; Proc. of the 18th International Conference on Data Engineering, 212, 2002.

[11] Gang D., Shi Z., Yang L.; Time series prediction using wavelet process neural network; Chinese Physics B, 17:6, 2008.

[12] Camilleri, M.; Forecasting Using Non-Linear Techniques In Time Series Analysis: An Overview Of Techniques and Main Issues, CSWA04, Computer Science Annual Research workshop, 2004.

[13] Hiden H., Willis M., Tham M, Montague G.; Non-linear principal components analysis using genetic programming; Computers and Chemical Engineering, 23: 413-425, 1999.

[14] Roberts S, Roussos E, Choudrey R.; Hierarchy, priors and wavelets: structure and signal modeling using ICA. Signal Processing; 84: 283-297, 2004.

[15] Chong T.; Financial Time Series Forecasting Using Improved Wavelet Neural Network; Master's Thesis, 2009.

[16] Zhang, B, Coggins, R.; Multiresolution Forecasting for Futures Trading Using Wavelet Decompositions, IEEE transactions on neural networks, vol. 12, no. 4, 2001.

[17] Yevgeniy B, Lamonova M,; An adaptive learning algorithm for a wavelet neural network. Expert Systems, 22: 235-140, 2005.

[18] Cleland, JGF; Network-Home-Care Management System (TEN-HMS) study Risk of Recurrent Admission and Death: The Trans-European Noninvasive, J. Am. Coll. Cardiol., 45;1654-1664, 2005.

[19] Zheng, Y., Zhiping, L, Tay, D; State-dependent vector hybrid linear and nonlinear ARMA modeling: Applications; Circuits, Systems, and Signal Processing; Vol. 20, N. 5, 575-597, 2001.