A Wavelet Based Prediction Method for Time Series

Cristina Stolojescu[1,2] Ion Railean[1,3] Sorin Moga[1] Philippe Lenca[1] and Alexandru Isar[2]

[1] Institut TELECOM; TELECOM Bretagne, UMR CNRS 3192
Lab-STICC; Université europèenne de Bretagne, France
(e-mail: `firstname.lastname@telecom-bretagne.eu`)
[2] Politehnica University of Timisoara, Romania
Faculty of Electronics and Telecommunications
(e-mail: `firstname.lastname@etc.upt.ro`)
[3] Technical University of Cluj-Napoca, Romania
Faculty of Electronics, Telecommunications and Information Technology

Abstract. The paper proposes a wavelet-based forecasting method for time series. We used the multi-resolution decomposition of the signal implemented using trous wavelet transform. We combined the Stationary Wavelet Transform (SWT) with four prediction methodologies: Artificial Neural Networks, ARIMA, Linear regression and Random walk. These techniques were applied to two types of real data series: WiMAX network traffic and financial. We proved that the best results are obtained using ANN combined with the wavelet transform. Also, we compared the results using various types of mother wavelets. It is shown that Haar and Reverse biorthogonal 1 give the best results.
Keywords: time series, Stationary Wavelet Transform, forecasting.

1 Introduction

Forecasting, or prediction, is the process of estimation in unknown situations, based on the analysis of some factors that are believed to influence the future values, or based on the study of the past data behavior over time, in order to take decisions. Time-series forecasting is an important area of forecasting where the historical values are collected and analyzed in order to develop a model describing the behavior of the series. When the time series is non-stationary, it is very difficult to identify a proper global model, [3]. To overcome this problem, an efficient way is to use the wavelet decomposition technique in the preprocessing step. The Wavelet transform (WT) provides a useful decomposition of time series, in terms of both time and frequency, permitting us to effectively diagnose the main frequency component and to extract abstract local information from the time series.

WT has been frequently used for time series analysis and forecasting in the recent years, [1,2]. Models that accurately catch the statistical characteristics of the actual traffic play a significant role in studying the network, in understanding its dynamics, in designing and controlling the network. For financial time series prediction, sales forecasts are very useful in the economic domain because they are used to optimize inventory levels. Several models have been proposed for time-series forecasting such as pure statistical or based on Artificial Neural Networks (ANN). Traditional linear time series models including ARIMA (Auto Regressive Integrated Moving Average) model proved to be good at capturing the behavior of the time series. To deal with the non-linear nature of time-series, the ANN model is probably the most popular method. It can capture any kind of relationship between the output and the input theoretically.

In this paper, we analyze the influence of different mother wavelets on the performance of forecasting. We compared the results trying to find out which is the best of the mother wavelets to be applied and, using this wavelet, which method gives the best forecasts. The rest of the paper is organized as follows: in Section 2 we present some theoretical considerations regarding WT and multi-resolution analysis. In Section 3 we describe the forecasting framework. The experimental results are presented in the fourth Section and finally, Section 5 is dedicated to the conclusions.

2 The wavelet analysis

As stated before one of our goals is to compare the forecasting accuracy by using the wavelet transform in the preprocessing step. The transform of a signal is just another form of representing it, which does not change the information content present in the signal. A linear time-frequency transform correlates the signal with a family of waveforms that are well concentrated in time and in frequency. Multi-resolution analysis (MRA) is a signal processing technique that takes into account the signal's representation at multiple time resolutions. Using wavelet MRA, the collected measurements can be smoothed until the overall long-term trend is identified. Fluctuations around the obtained trend are further analyzed at multiple time scales. The level of decomposition depends on the length of the data set (the number of values). At each temporal resolution two categories of coefficients are obtained: approximation coefficients and detail coefficients. Generally, the MRA is implemented based on Mallat's algorithm [7], which corresponds to the computation of the Discrete Wavelet Transform (DWT). The disadvantage of Mallat's algorithm is the decreasing of the length of the coefficient sequences with the increasing of the iteration index due to the utilization of the decimators. Another way to implement a MRA is the use of the trous methodology, also known as Shensa's algorithm [6], which corresponds to the computation of the Stationary Wavelet Transform (SWT). In this case the utilization of decimators is avoided, but at each iteration different low-pass and high-pass filters are used. There is a variety of mother wavelets [7] such as Daubechies, Symlet, Meyer, Morlet, etc., and the choice of the mother wavelets depends on the characteristics of data. The Daubechies wavelet transforms have been increasingly adopted by signal processing researchers. Haar wavelet transform, which is also the simples Daubechies wavelet is a good choice to detect time localized information. In this work we propose to use some mother wavelets belonging to Daubechies family, but also other orthogonal wavelet families such as Symmlets, also known as the Daubechies least asymmetric mother wavelets, and Coiflets also designed by Ingrid Daubechies to be more symmetrical than the Daubechies mother wavelet, and biorthogonal respective reverse biorthogonal wavelets. Biorthogonal wavelets exhibit the property of linear phase, which is needed for signal reconstruction. If, instead of a single wavelet, two wavelets are used (one for decomposition and the other for reconstruction), interesting properties are derived, [7]. Different types of mother wavelets will be used in the data preprocessing step of our forecasting framework presented in the next Section.

3 Forecasting framework

The main idea of the prediction method using wavelets is to decompose the original signal into a range of frequency scales and then to apply the forecasting methods to these individual components. Our forecasting framework, which belongs to the supervised paradigm, is presented in Figure 1 and implies the following steps:

1. Preprocessing step which includes data clearing, such as identification of the potential errors in data sets, handling missing values, and removal of noises or other unexpected results that could appear during the acquisition process. At this stage the input data is also analyzed in order to find if it contains large spikes and valleys indicating periodicities.
2. Use the SWT to decompose the data separately for the training set and the test set. Each component represents the real data in a frequency range that is easier to predict than the original series. A good predictor should be able to identify the separate scale-related components of the series, in order to produce models that give accurate forecasts. So, our approach is to decompose the original time series into scale or frequency related components and model each component separately, in order to obtain more accurate models.
3. After obtaining the wavelet decomposition, we select the information from each level of decomposition for building the model.
4. In the training phase we design predictive models for each of the decomposed components of the original series. In the test phase the developed forecasting models are used to predict future values for each component. The Inverse SWT is used in the testing phase in order to obtain the forecasted signal from the predictions of the components.

The four models used in this work are presented below:

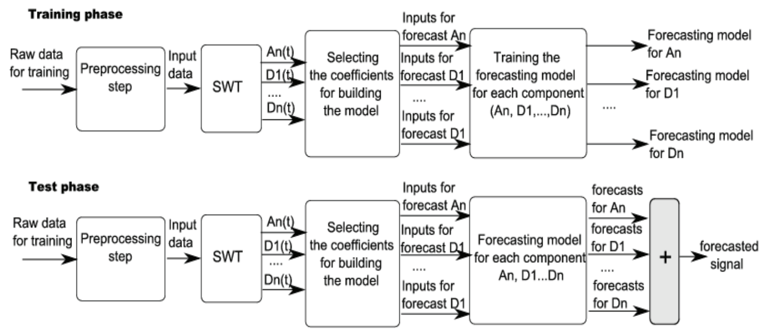

Fig. 1. The forecasting framework.

1. *ANN* models [5] represent a wide class of flexible nonlinear models which have been very used recently in the area of forecasting. The main advantage of an ANN that makes it suitable for various applications is that it learns from the past experiences. So, the basic idea is to train the ANN with past data and then use it to predict future values. Although many types of architectures have been proposed, the most popular one for time series forecasting is the feed-forward neural network [9]. In this work we used a Time-delayed neural networks(TDNN), detailed in [4].
2. *ARIMA* processes [8,10], are the natural generalizations of standard ARMA processes. This class of models is based on Box-Jenkins methodology [10] which is used to build the time series model in a sequence of steps which are repeated until the optimum model is achieved. More details about this method and how it was applied in our case are presented in [11].
3. *Linear regression (LR)* [8] is a simple statistical tool for modeling the output as a linear combination of inputs. The model's parameters are usually estimated using the least-squares method.
4. *Random walk (RW)* method [8] is based on the hypothesis that from one period to the next, the original time series takes a random "step" away from its last recorded position. The prediction of the future values is based on the previous values plus a constant that represents the average change between the two periods.

4 Experiments

4.1 Data sets

In this work we used historical data obtained by monitoring the traffic from 67 Base Stations (BS) composing a WiMAX network. The period of collection is of eight weeks, from March 17th till May 11th, 2008. Each BS has its own data set which is composed of numerical values representing the total number of packets from the uplink channel. Each value is recorded every 15 minutes. It can be easily deduced that for a given BS we have the following number of samples: 96 samples/day, 672 samples/week, and a total number of 5376 samples. So, the WiMAX data base can be seen as formed by 67 matrices (one for every BS) that have eight columns (the number of weeks) and 672 lines (the moments of time when the number of packets are recorded in a week).We also used one time series of financial data representing the total number of EUR-USD currency exchanges (the volume of data is similar to the number of packets from WiMAX. The period of collection is of fifteen weeks and the values are recorded every 15 minutes. We will have 96 samples/day, 672 samples/week and a total number of 10080 samples. In this case only one matrix will correspond to each of the two sets and it will be formed by fifteen columns (the number of weeks) and 672 lines. The objective of our work is to compare the influence of different mother wavelets used in the preprocessing step on the prediction accuracy. Also, using the best mother wavelets, we propose to evaluate some prediction models, such as pure statistical or based on neural networks.

4.2 Evaluation criteria

In order to evaluate the performance of prediction using different types of wavelets, we considered the most used statistical measures of error: the Mean absolute error (MAE), the Mean Square Error (MSE), the analysis of variance (ANOVA), the Symmetric Mean Absolute Percent Error (SMAPE), and the Root Mean Square Error (RMSE). We have also calculated SMAPE L, MAPE L and MAE L, between the mean of the original signal and the mean of the forecasted signal, because ARIMA and LR cannot be used to obtain forecasts for every moment of time as ANN and RW can. For linear models the trajectory of the forecasts is represented through sloping line which represents the weekly increase.

4.3 Results and discussions

Regarding the WT, we propose various types of mother wavelets such as Daubechies (db), Coiflet (coif), Symlet (sym), Biorthogonal (bior), and Reverse Biorthogonal (rbio). In Table 1 and Table 2, for each type of mother wavelets and every type of error, excepting SMAPE L, MAPE L and MAE L, we present the average value corresponding to the three types of ANN and RW. SMAPE L, MAPE L and MAE L correspond to all the proposed methods. We do not take into consideration the results given by using the linear regression because the wavelet transform does not have any influence on the predictions. In this case the mean value of the details is zero, and the prediction obtained for the details will be also zero. In the case of WiMAX traffic (Table 1), the results are represented as the average value for all 67 BS. According to Table 1, the wavelet of Haar (db1), which is the simplest of the Daubechies family and rbio1.1 give the best prediction performance. The results also indicate that with the increase of the filters' length (support of the mother wavelets), the performance of the wavelet transform deteriorates. The results represent the mean values for all the forecasting methods and all the 67 BSs with the observation that in the case of ARIMA only SMAPE L, MAPE L and MAE L could be calculated.

Wavelet	RSQ	SMAPE	MAPE	MSE	RMSE	MAE	SMAPE L	MAPE L	MAE L
coif 1	1.445	1.09	**0.2113**	11.72	2.80	1.0304	0.890	0.0020	0.9599
coif 2	1.493	1.22	0.2285	12.95	2.83	0.8748	0.837	0.0019	**0.7191**
db 1	**1.168**	**1.08**	0.2367	**8.06**	**2.43**	**0.7685**	**0.812**	**0.0016**	0.7327
db 2	1.364	1.15	0.2451	10.52	2.69	0.8408	0.855	0.0019	0.7768
db 3	1.358	1.12	**0.2117**	**9.76**	**2.64**	0.8193	0.857	0.0018	0.7678
db 4	1.490	1.11	0.2159	10.61	2.58	0.7985	0.834	0.0018	0.7563
db 5	1.435	1.11	0.2190	12.56	2.75	0.8339	0.823	0.0019	0.7730
bior 3.1	0.695	1.13	0.3152	9.86	2.52	84020	0.860	0.0018	**0.7071**
rbio 1.1	**1.200**	**1.08**	0.2215	10.00	2.61	**0.7948**	**0.820**	**0.0017**	0.8947
rbio 2.2	1.482	1.19	0.3202	10.29	2.71	0.8747	0.891	0.0018	0.7690
rbio 3.3	1.952	1.21	0.2623	10.33	2.88	0.9509	0.907	0.0022	1.0690
sym 2	1.365	1.26	0.2146	13.20	2.89	0.8854	0.895	0.0019	0.7412

Table 1. Comparison between wavelets, WiMAX traffic.

In the case of financial data, for the set containing the EUR-USD exchange currency, the results are shown in Table 2. We can observe that the best forecasting performance is obtained using the mother wavelets coif2 and sym2.

For the purpose of forecasting methods comparison, we propose the following variants: three types of methods based on ANN (ANN No Sliding, ANN Known Sliding, and ANN UnKnown Sliding), ARIMA, LR, and RW model for two weeks prediction. The first method using ANNs, ANN No Sliding, is the simplest one: we train the ANN once for each decomposition level. For inputs, we have the first (n-2k) weeks, where n is the total number of weeks, and k is the number of weeks we want to forecast. The target

Wavelet	RSQ	SMAPE	MAPE	MSE	RMSE	MAE	SMAPE L	MAPE L	MAE L
coif 1	0.688	0.556	0.1152	1.627	1.220	0.3586	0.516	0.0821	0.4240
coif 2	0.455	0.522	**0.0793**	**1.089**	**1.038**	**0.2967**	**0.453**	0.0732	0.3799
db 1	0.625	0.520	**0.0839**	1.356	1.126	0.3175	0.454	**0.0713**	**0.3690**
db 2	0.715	0.578	0.1088	1.610	1.219	0.3550	0.497	0.0812	0.4199
db 3	0.586	0.585	0.1156	1.499	1.188	0.3618	0.531	0.0864	0.4461
db 4	**0.871**	0.600	0.1114	1.604	1.239	0.3745	0.527	0.0863	0.4459
db 5	**0.808**	0.587	0.1121	1.557	1.225	0.3700	0.546	0.0912	0.4710
bior 3.1	0.628	0.534	0.1137	1.552	1.173	0.3390	0.433	**0.0712**	**0.3681**
rbio 1.1	0.615	**0.519**	0.0958	1.286	1.096	0.3208	0.457	0.0716	0.3704
rbio 2.2	0.541	0.555	0.1087	1.440	1.149	0.3406	0.491	0.0790	0.4081
rbio 3.3	0.772	0.595	0.0993	1.418	1.167	0.3480	0.455	0.0716	0.3705
sym 2	0.476	**0.499**	0.0890	**1.117**	**1.037**	**0.2962**	**0.453**	0.0736	0.3813

Table 2. Comparison between wavelets, EUR-USD currency exchanges.

consists of the data taken from the weeks (n-2k+1) to (n-k). The data used for ANN's inputs during the testing phase is the information from the weeks (k+1) to (n-k). The output signal is compared to the real data of the last k weeks. The next method (ANN Known Sliding) uses sliding, retraining the network with the real information. The entire signal is divided into smaller parts. Each of these sequences will predict a small part of the final forecasted signal. The information for ANNs retraining is always taken from the real data. The last method, ANN UnKnown Sliding, proposes a forecasting using sliding with unknown data. The only difference consists in the fact that the information used for the next simulation and retraining is taken not from the original signal, but from the previously predicted one. For more details see [4]. The use of ARIMA is detailed in [11].

In the case of WiMAX traffic, the comparison was made using db1 mother wavelets. The results presented in Table 3 prove that ANN performs better than the other prediction techniques. Also the linear regression model gives very good forecasting results.

Forecasting Model	SMAPE L	MAPE L	MAE L
ANN No Sliding	**0.472**	0.0011	0.4428
ANN Known Sliding	0.509	**0.0009**	0.4241
ANN UnKnown Sliding	0.722	0.0017	0.6681
ARIMA	0.772	0.0027	0.9990
Linear Regression	0.523	0.0031	**0.3868**
Random Walk using Wavelets	4.440	0.0030	1.3633

Table 3. Forecasting techniques comparison for WiMAX traffic.

For the financial data base, we used the coif2 mother wavelets. The results are presented in Table 4. We found that the suitable model is as well the one using ANNs.

Forecasting Model	SMAPE L	MAPE L	MAE L
ANN No Sliding	0.169	0.020	0.1079
ANN Known Sliding	**0.153**	**0.0178**	**0.957**
ANN UnKnown Sliding	0.267	0.0344	0.1812
ARIMA	1.135	0.3245	1.7054
Linear Regression	0.191	0.0243	0.1109
Random Walk using Wavelets	0.940	0.2670	1.4173

Table 4. Forecasting techniques comparison for financial data.

5 Conclusion

Regarding the Wavelet transform, our results show that Haar, which is the simplest of Daubechies family, and Reverse biorthogonal 1 improve the performance of the prediction technique.

An important conclusion is that as much the support of the mother wavelets increases, the performance of the wavelet transform deteriorates. In addition, using the best mother wavelets in data preprocessing step, we proved that ANN outperforms the other forecasting methods. Also, our results confirms the results in [Papagiannaki, et al, 2005] and point out that if we are interested in tendency prediction, for more than one month ahead, than linear models are suitable for this type of forecasting. We should also point out that we have applied our algorithm on two different data sets which are not comparable. The financial data (the EUR-USD currency exchanges) exhibit an almost constant tendency, while WiMAX traffic presents a strong variability and its tendency (long term trend) represents a sloping line. However, our algorithm is applicable to both types of data and the obtained predictions are accurate. As a future work we propose to apply our algorithm on other time series, for example transportation data, including highway traffic, aircraft flights, traffic data of cars in tunnels, traffic at automatic payment systems on highways, traffic of individuals on subway systems, etc.

References

1. X. Wang, X. Shan, A wavelet-based method to predict Internet traffic, in *Communications, Circuits and Systems and West Sino Expositions*, vol.1, pp. 690-694, (2002).
2. K. Papagiannaki, et al, Long-term forecasting of Internet backbone traffic, in *IEEE* Transactions on Neural Networks, vol.16, pp. 1110-1124,(2005).
3. Zhang et al, Multiresolution Forecasting for Futures Trading Using Wavelet Decompositions,in *IEEE* Transactions on neural networks, vol. 12, no. 4, (2001).
4. I.Railean et al, WIMAX Traffic Forecasting based on Neural Networks in Wavelet Domain, submitted to RCIS 2010 (2010).
5. P. Mehra and B.W.Wah, Artificial Neural Networks: Concepts and Theory in *IEEE* Computer Society Press Tutorial, Los Alamitos, CA, (1992).
6. M.J.Shensa, *Discrete Wavelet Transform. Wedding the a trous and Mallat algorithms, IEEE* Transactions and Signal Processing, 40, pp. 2464-2482,(1992).
7. S. Mallat, *A Wavelet Tour of Signal Processing*, Second Edition, (1999).
8. B. Abraham and J. Ledolter, Statistical Methods for forecasting, in *Wiley Series in Probability and Mathematical Statistics*, (1983).
9. G. P. Zhang, M. Qi, Neural network forecasting for seasonal and trend time series, in the *European Journal of Operational Research* 160, pp. 501-514,(2005).
10. G. Box, G. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, CA, (1970).
11. C. Stolojescu et al, Forecasting WiMAX BS Traffic by Statistical Processing in the Wavelet Domain, in *Proceedings of the International Symposium on Signals, Circuits and Systems*, Iasi, Romania, pp. 177-183, (2009).