TIME SERIES PREDICTION USING GENETICALLY TRAINED WAVELET NETWORKS

A. Procházka, V. Sýs
Prague University of Chemical Technology
Department of Computing and Control Engineering
Technická 1905, 166 28 Prague 6, Czech Republic
Phone: 0042-2-332 4259 ∗ Fax: 0042-2-243 11082
E-mail: prochaz@vscht.cz, sysv@vscht.cz

Abstract - The paper presents a contribution to the analysis of wavelet transfer function use in neural network systems and the discussion of some possible learning algorithms of such structures. Wavelets local influence both in time and frequency domains are presented at first giving motivation for wavelet networks application and providing bases for their initial coefficient estimation described recently. The main part of the paper is devoted to the network coefficients optimization using genetic algorithms as an alternative to the gradient descent method. Principles of the evolution techniques are presented for a simple system in this part and then applied for a given time series modelling and prediction.

INTRODUCTION

Problem of nonlinear time series prediction is studied in various disciplines now including engineering, medical and economical applications [1]. Most approaches applied in this area are based upon the use of adaptive methods and multilayer neural networks [2], [3]. Their modification include wavelet transfer function study [4], [5], [6] and cascade correlation networks [7].

The paper contributes to this problem in study of links between wavelets and neural networks presented recently [8], [6]. The emphasis is in the analysis of the learning process based upon genetic algorithm use [9], [10] as an alternative to the gradient estimate applied in the backpropagation method. The goal of the paper is to show advantages of genetic algorithm approach in this case.

Methods presented in the paper are verified for simulated series to enable their further use for real signal analysis including that in the Signal processing information base [11] and physiological signals study.

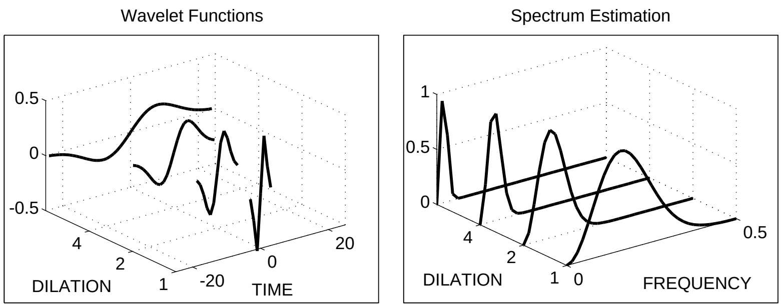


Figure 1:

An example of the wavelet set based on the given initial function $h(x) = -x\, e^{-x^2/2}$ with its spectrum estimation and relation between wavelet dilation and spectrum compression.

WAVELET NETWORK CONSTRUCTION

Wavelet functions and their application in signal processing are described in various papers published recently [12], [13], [14]. These functions in close connection with the wavelet transform can be used very efficiently both for signal analysis as an alternative to the short-time Fourier transform and for signal modelling in wavelet network structures.

Basic properties of wavelet functions include their local influence both in time and frequency domains [15]. The set of continuous wavelet functions is usually derived from the initial (mother) wavelet $h(x)$ and coefficients of dilation (d) and translation (t) defining function

$$h_{d,t}(p) = h((p-t)/d) = h(w\, p + b) \tag{1}$$

for $w = 1/d$ and $b = -t/d$. Using the signal processing point of view it is possible to consider the initial wavelet as a pass-band filter. Wavelet dilation resulting in its pass-band compression is presented in Fig. 1 for a chosen wavelet function. Properly constructed wavelets allow in this way signal decomposition into different frequency bands and their multiresolution analysis [14].

Basic two layer neural network structure used for signal prediction is presented in Fig. 2. Signal pattern is used as a neural network input and the whole network is trained to evaluate its output close to the target signal values in the mean square sense. While the output layer transfer function is usually linear in this case the transfer function of the first layer can be nonlinear and sigmoidal functions are often applied. We shall study the alternative use of wavelet functions in such a structure now.

Mathematical description of a wavelet network $R - S1 - S2$ is similar to that of sigmoidal neural network system. Summarizing network coefficients in matrices

$$\mathbf{W1}_{S1,R} = \begin{bmatrix} w1_{1,1} & \cdots & w1_{1,R} \\ & \cdots & \\ w1_{S1,1} & \cdots & w1_{S1,R} \end{bmatrix}, \quad \mathbf{W2}_{S2,S1} = \begin{bmatrix} w2_{1,1} & \cdots & w2_{1,S1} \\ & \cdots & \\ w2_{S2,1} & \cdots & w2_{S2,S1} \end{bmatrix}$$

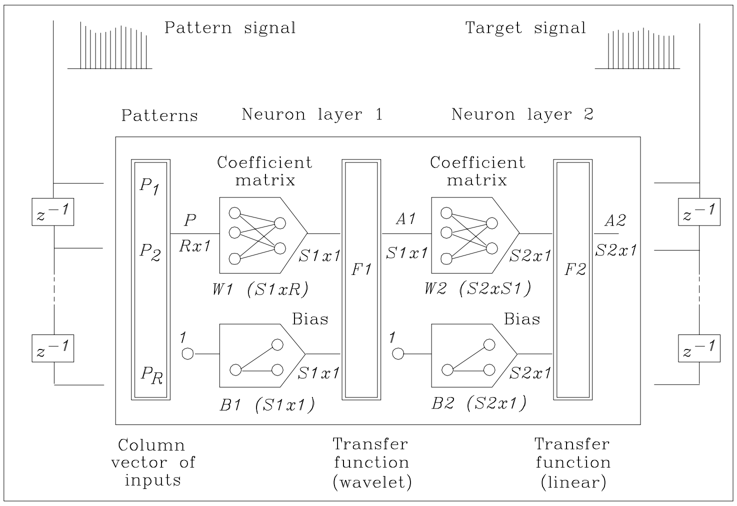


Figure 2:

Signal prediction based on the two layer neuron network having the output vector $\mathbf{A2} = F2(\mathbf{W2} * \mathbf{A1} + \mathbf{B2})$ where $\mathbf{A1} = F1(\mathbf{W1} * \mathbf{P} + \mathbf{B1})$.

and bias vectors

$$\mathbf{b1} = [b1_1, \cdots, b1_{S1}]', \qquad \mathbf{b2} = [b2_1, \cdots, b2_{S2}]'$$

it is possible to use transfer functions $F1$ and $F2$ for the first and the second layer respectively to evaluate network output

$$\mathbf{A2} = F2(\mathbf{W2} * \mathbf{A1} + \mathbf{B2}), \qquad \mathbf{A1} = F1(\mathbf{W1} * \mathbf{P} + \mathbf{B1})$$

While $F2(t) = t$ in many cases of signal prediction it is possible to choose the first layer transfer function in various ways including wavelet functions as well.

During the learning process the summed square error deviations between evaluated and target values are minimized. To shorten the iterative learning process various methods of coefficients initial estimates have been suggested [3], [8]. We shall restrict our attention to the study the genetic algorithm approach to the wavelet transfer function case.

SYSTEM COEFFICIENTS OPTIMIZATION

Analyzing basic problems of wavelet network optimization it is possible to start with signal approximation of given pattern and target values by a single wavelet function $F1(x) = -x\, e^{-x^2/2}$ for unknown dilation a translation coefficients. Error surface in the vicinity of problem solution is presented in Fig. 3. It is obvious that classical steepest descent approach and closely connected backpropagation method can be very inefficient even in such a simple case.

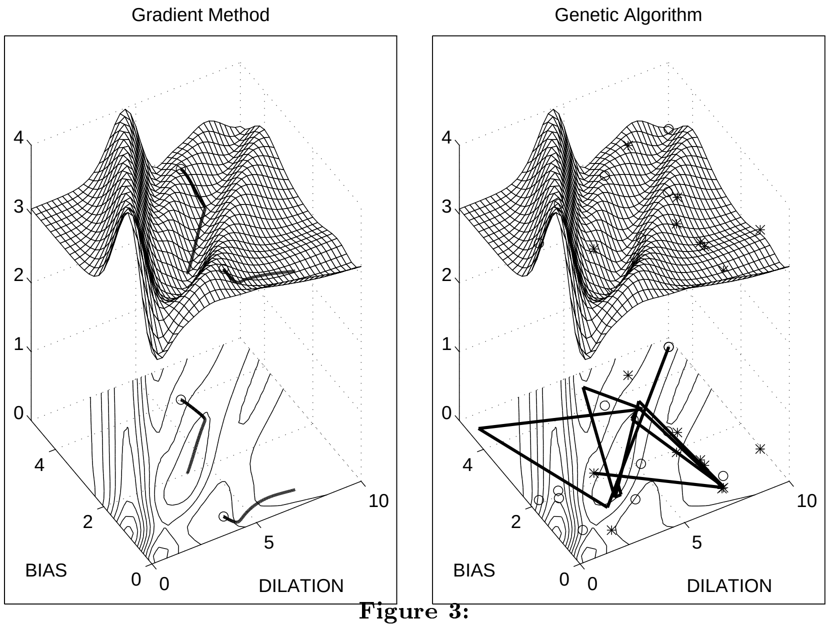


Figure 3:

Comparison of the initial part of the optimization process for the wavelet network having its structure 1−1−1 for a given nonlinear transfer function $F1(x) = -x\ e^{-x^2/2}$ using a gradient search and genetic algorithm optimization approach: **(a)** Gradient search for two different initial estimates, **(b)** Genetic search presenting the initial ($\circ$) and final ($*$) population and the best individual evolution.

On the other hand the genetic approach [9], [16] for system coefficients optimization can provide more reliable results even for such a complicated nonlinear problems. Its general principle resembles a very simplified version of the natural selection giving better chance to survive for the best individuals while others die out. Basic principles of recombination, crossing over and random mutation can be efficiently applied in various engineering problems and wavelet networks as well.

A random set of network parameters (dilations, translations) in wavelet network system presented above can be used to cover the whole region of possible solutions and to define separate individuals. All pairs of two coefficients creating a population are coded in binary form to ease the following operations. As all solutions are assumed to approximate the given signal it is possible to evaluate the mean square error for all of them standing for the fitness number and giving idea about both the population and each individual performance. Individuals with a relatively high fitness number obtain a higher probability for further reproduction and vice versa. The whole algorithm assume

- the selection process based upon the roullete wheel manner providing

Neural Network		*Mean Square Error*	
Type	*Structure*	*Learning part*	*Prediction part*
Wavelet Network	1-3-1	0.0298	0.0305
Sigmoidal Network	1-5-1	0.0358	0.0363

TABLE I:
Comparison of mean square errors during the learning and prediction parts for wavelet and sigmoidal networks.

statistically better chance for reproduction to individuals (defined by corresponding dilation and translation coefficients) with higher fitness number

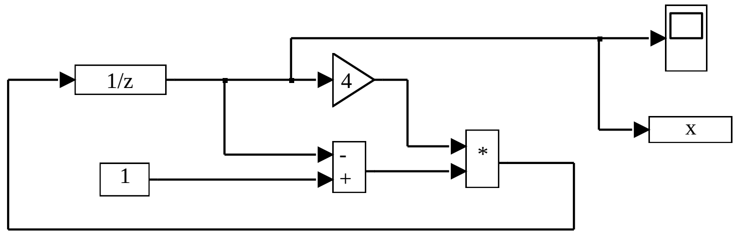
- the mating process based upon the random pairs selection and crossing-over application exchanging substring in selected pairs using a random position generator
- the mutatation process resulting in random switches between zero and one using a given probability of such a process

For a new population fitness numbers are evaluated again and the whole process is repeated.

Results for a single wavelet neuron optimization using a genetic approach are presented in Fig. 3. Comparing initial and final population it is possible to follow the result of the whole population mean square error decrease. The best individual evolution is presented as well.

SIGNAL MODELLING AND PREDICTION

Nonlinear time series modelling and prediction by wavelet networks has been tested for a common nonlinear time series defined by relation $x(i+1) = \alpha x(i)(1-x(i))$ for $i = 1, 2, 3, \cdots$, $\alpha = 4$ and $x(1) = 0.1$ and simulated using diagram given in Fig. 4. Applying network structure presented in Fig. 2 restricted to system $1 - S1 - 1$ it is possible to use values $x(i)$ as patterns and $x(i+1)$ as targets. Using the genetic algorithm approach for the learning process it is possible to obtain results presented in Fig. 4 as well. Evaluating further the mean square error in the learning and prediction part it is possible to compare results for various network structures and different wavelet functions as well. Signal used to test methods presented in the paper implied values of the mean square error summarized in Tab. I giving these values in the learning part and the prediction part for the wavelet and sigmoidal neural network. Errors of the same order were achieved with the wavelet network of significantly lower number of network variables. Genetic algorithms provide in all cases the method for system coefficient evaluation.



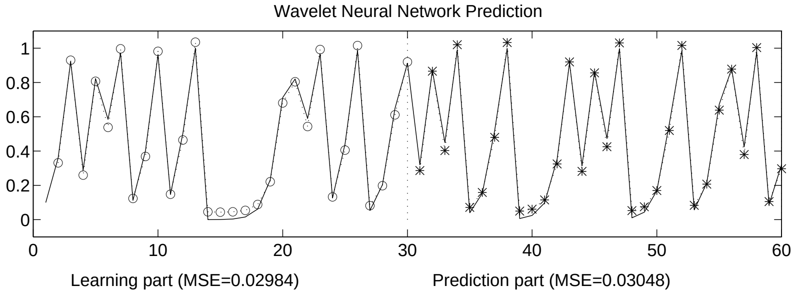


Figure 4:

Selected nonlinear chaotic time series generation for the initial value of x(1)=0.1 applied for the prediction method analysis and results of signal identification using the wavelet network of structure $1-3-1$ comparing one step ahead evaluation during the learning process and signal estimation during the prediction part.

CONCLUSION

The paper presented the application of genetic algorithms for wavelet network optimization process providing (coefficient) population of individuals with improving properties defined by their fitness numbers. Even though other algorithms including the gradient descent method may converge faster genetic methods avoid to be captured by local minima owing to the mutation process and they are very effective in converging to the global optimum. The paper presents comparison of these methods for simulated signals and provides basis for real signals modelling and prediction.

ACKNOWLEDGEMENTS

The paper has been supported by the Department of Computing and Control Engineering at the Prague University of Chemical Technology.

REFERENCES

[1] A. Lapedes and R. Farber, "Nonlinear signal processing using neural networks: Prediction and system modelling, LA-UR-87-2602", Tech. Rep., Los Alamos National Laboratory, USA, 1987.

[2] J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computing*, Addison-Wesley Publishing Company, Redwood City, California, 1991.

[3] D. Nguyen and B. Widrow, "Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights", in *Proceedings of Internationale Joint Conference on Neural Networks*, 1990, pp. III/21–III/26.

[4] B. R. Bakshi and G. Stephanopoulis, "Wavelets as basis functions for localized learning in multi-resolution hierarchy", in *Proceedings of Internationale Joint Conference on Neural Networks*, Baltimore, 1992, pp. II/140–II/145.

[5] Sathyanarayan S. Rao and Ravikanth S. Pappu, "Nonlinear time series prediction using wavelet networks", in *Proceedings of World Congress on Neural Networks*, Portland, Oregon, July 1993, pp. IV/613–IV/616.

[6] Q. Zhang and A. Benveniste, "Wavelet networks", *IEEE Trans. on Neural Networks*, vol. 3, no. 7, pp. 889–899, 1992.

[7] S. Fahlman et. al, "The cascade-correlation learning architecture, CMU-CS-90-100", Tech. Rep., School of Computer Science, Carnegil Mellon University, USA, 1992.

[8] Q. Zhang, "Regressor selection and wavelet network construction", Tech. Rep., IRISA, France, 1993.

[9] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Publishing Company, Reading, Massachusetts, 1989.

[10] C. H. Chu and C. R. Chow, "A genetic algorithm approach to supervised learning for multilayered networks", in *Proceedings of World Congress on Neural Networks*, Portland, Oregon, July 1993, pp. IV/744–IV/747.

[11] D. H. Johnson and P. N. Shami, "The signal processing information base", *IEEE SP Magazine*, pp. 36–42, October 1993.

[12] O. Rioul and M. Vetterli, "Wavelets and signal processing", *IEEE SP Magazine*, pp. 14–38, October 1991.

[13] G. Strang, "Wavelets and dilation equations: A brief introduction", *SIAM Review*, vol. 31, no. 4, pp. 614–627, December 1989.

[14] M. Vetterli, "Wavelets and filter banks: Theory and design", *IEEE Trans. Signal Processing*, vol. 40, no. 9, pp. 2207–2232, September 1992.

[15] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis", *IEEE Trans. Inform. Theory*, vol. 36, pp. 961–1005, September 1990.

[16] B. Müller and J. Reinhardt, *Neural Networks*, Springer-Verlag, 1991.