TRAINING OF A WAVELET PREPROCESSED NEURAL NETWORK FOR DETECTION OF ANOMALY SIGNAL PATTERNS BY USING A MODIFIED EVOLUTIONARY ALGORITHM

Res.Assist. Mehmet Siraç ÖZERDEM
Assist.Prof.Dr. Osman Kaan EROL
Assist.Prof.Dr. Berk ÜSTÜNDAĞ
Istanbul Technical University, Electrical-Electronics Faculty, Computer Engineering Dept., 34390 Maslak, Istanbul, Turkey

***Abstract*: Earthquake is a natural event that has no linear property in time and space and cannot be modeled exactly. Estimation of an earthquake by means of pattern learning and recognition property of wavelet - artificial neural networks (ANN) is the essential goal of this study. The values obtained through the sensor to measure earthquake-related monopolar electric field constitute time dependent patterns. These patterns are analyzed and evaluated by the assistance of wavelet - ANN. Phases of evaluation process can be explained briefly: wavelet is used as a feature extraction and classification of the patterns are carried out with wavelet. In the second stage of the detection of anomalies, clustered patterns are trained with multi-layer feed forward neural network architecture. A modified Genetic Algorithm with unified reproduction and crossover is used to train the Neural Network. At the output layer of network, a mechanism is realized that produces results related with pattern's being normal or abnormal and this result depend on the input pattern's content. As a result of applications covered in this study, the measurement of electric field resulted from regional tectonic stress and the earthquake forerunner pattern that resides in this field is detected with a high success rate. In addition, it is observed that learning is improved and success is increased by the increase in the number of pattern that includes abnormal change.**

I INTRODUCTION

The change in many physical parameters such as, the change in water level in wells [1], temperature change of spring waters, radon emission [2], determination of long term regional displacement using GPS data [3], the change in earth magnetic field have been studied in prediction of earthquakes. A network of 16 stations, having a specifically developed monopolar electric field (MEF) probe has been established to predict earthquakes in a joint project by the Faculty of Electrical Engineering and Faculty of Mines of Istanbul Technical University [4]. Filtering property of wavelet and learning-recognization properties of artificial neural networks methods are also used in evaluation of MEF data, which is believed to be more correlated with earthquakes than other physical precursors. Automatic decomposition of anomaly patterns and their usage in prediction are possible in the models having known parameters. However, the artificial neural network is trained according to related pattern, whenever the physical event occurs but in case of seismological events, which have complicated models, with many uncertain parameters.

An artificial neural network is a hierarchical organization, in which artificial neurons interacts with each other [5]. In contrast to classical expert systems, where knowledge is interpreted as a set of rules, artificial neural networks learn from given examples and established their own set of rules [6]. Kohonen pointed out that artificial neural network are hierarchical structures consisting of densely parallel-connected adaptive components and having the ability to interact with real world like the biological systems do [7].

A decisive property of artificial neural network is its learning ability. Learning of artificial neural network is realized through changing the connection weights of network, depending on input samples or in addition to them, depending on related outputs [8]. Two types of learning strategies in literature can be classified as supervised learning and unsupervised learning [9]. The basic difference between them is the existence or nonexistence of desired outputs during learning procedure. The more and different samples are used in training, the more objects, artificial neural networks will recognize and the less error will be expected to occur.

II DATA AND METHODS

This study concerns monopolar electric field measurement (MEFM) which is assumed to be related to change in fault stress and evaluation of patterns shaped by acquired data from different locations. Patterns are constructed at a central computer that receives on line

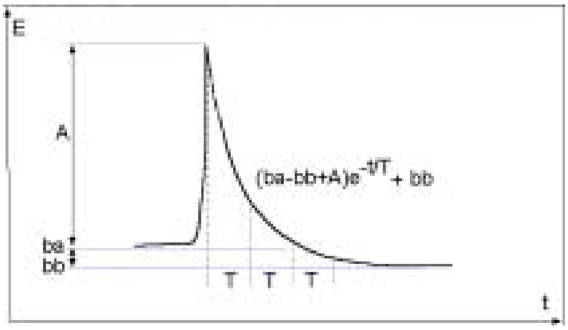


Figure 1. Pre-earthquake anomaly pattern

data packages from 16 stations through FTP over the Internet in Marmara Region (northwestern Turkey). Figure 1 shows us voltage as a function of time for a basic template of pre-earthquake anomaly pattern. Basic main characteristics of the pattern can be explained as follows: *A* variable denotes us the amplitude difference between the peak of normal value and anomaly, whereas *ba*, the corresponding amplitude value in which the normal behavior begins in time domain and *bb*, anomaly value in which the abnormal behavior begins. $\Delta\tau$ shows us decaying time constant of abnormal behavior.

Although the geological validation of the hypothesis derived from monopolar electrode measurements remains unclear, our research project group was able to predict the location, intensity and time of large earthquakes by investigating amplitude (A), *ba-bb* difference and $\Delta\tau$ parameters of the patterns acquisied from electrode. Those patterns are implemented as training vectors as input to the neural network architecture.

Analog signals obtained by a monopolar electric field (MEF) probe, having sensitivity in Femto coulomb range are digitized in measurement stations and after processing in a digital adaptive filter sent to the data storage center via Internet. Patterns shaped by the data, collected from 16 stations in Marmara Region (northwestern Turkey) are applied to the input of wavelet-artificial neural network shown in figure 2.

As shown in figure 2, wavelet coefficients of measurement data are obtained with using wavelet analysis. Then the coefficients are used as input vectors for multi layer perceptron.

III FEATURE EXTRACTION

The measurement based on the monopolar probe is non-stationary signals. It means that these signals cannot be analyzed with using classical methods such as Fourier transformation. To overcome this problem, wavelet transformation is used.

The Wavelet Transform (WT) is a technique for analyzing signals. It was developed as an alternative to the short time Fourier Transform (STFT) to overcome problems related to its frequency and time resolution properties. More specifically unlike the STFT that provides uniform time resolution for all frequencies the discrete WT provides high time resolution and low frequency resolution for high frequencies and high frequency resolution and low time resolution for low frequencies [10].

There are some types of wavelet methods. In this paper, Daubechies 1, 2, 3, 4, 5, Symlet 2, 3, 4, 5, Haar, Coiflet 1, 2, Bior 1.1,1.3,1.5,2.2,3.1 are used for getting the wavelet coefficients of anomalies. These transformation processes are implemented with using MATLAB. It is observed that Symlet 5 wavelet type has the most capability to identification of anomalies among the implemented types. Because of that the changes of Symlet 5 wave is approximatily same as the anomalies. An anomaly and its 5.th datail wavelet coefficients are shown in figure 3.

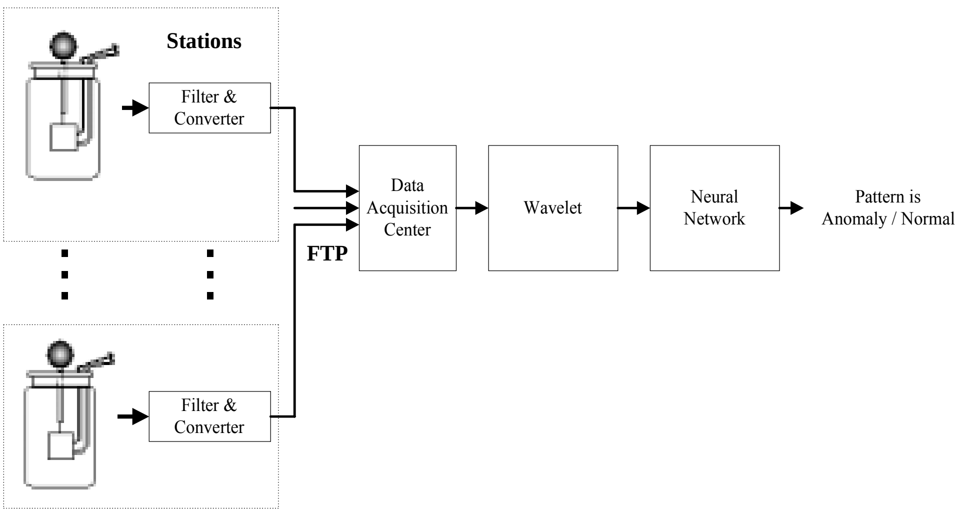


Figure 2. General architecture of earthquake pattern recognition system

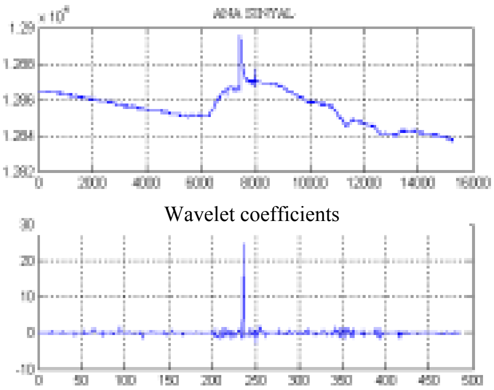


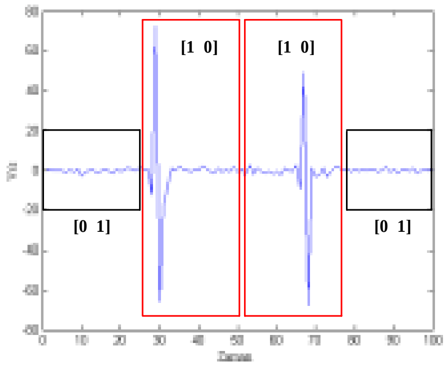
Figure 3. 5.th detail coefficients of the anomaly

III NEURAL NETWORK CONSIDERATIONS

In this paper, our model is based on (25-35-5-2) multi layer perceptron structure. A modified Genetic Algorithm with unified reproduction and crossover is used to train the Neural Network.

The training patterns are pre-processed in two stages. In the first stage, high frequency components of changes are filtered. Wavelet coefficients are normalized between 0.1 and 0.9 by using linear scaling in the second stage. Totally, the number of patterns is 120 for training neural networks.

The desired output vector of network constitutes two different classes. One of them includes anomaly patterns and the other one includes ordinary changes. These classes are coded as follows: the class that includes anomaly patterns is coded as [1 0] and the class that includes ordinary changes is coded as [0 1]. Hamming distance is 2 between the classes' codes. An example of coding classes is shown in figure 4.



[1 0] : Anomaly code - [0 1] : Normal code

Figure 4. Coding of classes

A modified Genetic Algorithm is used in the training of neural network and the detail of this process is explained in the next section.

IV MODIFIED GENETIC ALGORITHM

The optimization problem can be summarized as finding the set of parameters $\{x_n\}$ that minimizes an objective function $f(x_0,\ldots,x_n)$, which is also referred to as the fitness function in evolutionary algorithms. GA [11]–[14], is a global optimization tool for the minimization of the multi-modal fitness functions. The main difference between GAs and the other classical search algorithms is that GAs operates on a "population" of strings (chromosomes in genetic language) instead of one input parameter.

The reason for using a modified version of the GA is that classical GAs when coupled with an highly elitist selection property have the potential to converge all of the population to the fittest member; that is, GAs may fail to converge to the global optimum point depending upon the selective capacity of the fitness function (large variation of the fitness function for small variation of its input variables). The global convergence property of the algorithm is assured only through the mutation operator, which in turn creates new chromosomes at the expense of extra computational effort. In order to prevent the local optima entrapment, some "more intelligent" rules and/or hybrid techniques have been added to the GA. Tabu search procedure for example [15], [16], has led to the use of GA at beginning of the algorithm due to quick convergence to the fittest individual, then diverting to other search algorithms after some prescribed number of iterations in order to converge to the global optimum faster than GA alone. Some other hybrid techniques such as evolutionary-gradient search (EGS) procedures were also presented to increase the convergence speed of GAs [17], [18]. However, all of these precautions cause the algorithms to be inflated by extra number of calculations.

In this study, we have concentrated on the operators of the GAs and tried to improve their search capacity while avoiding entrapment in local optima. Furthermore, an improvement in the speed has been gained by eliminating unnecessary calculations in the reproduction stage. Since the GA algorithm presented in this study does not need the generational approach, it can be classified as a kind of steady-state GA [19]. The selection and crossover operators are unified. The new approach to GAs does not require roulette-wheel selection. It uses random selection, which does not require the evaluation of fitness values for selection purposes. Therefore, the calculation of the fitness values of every individual has not been required.

The crossover is replaced by a "partial overwrite" operator that has been applied with some probability depending on the difference between fitness values of two randomly selected candidates for the optimal solution. These two randomly selected candidate solutions are called "combat chromosomes" and for this reason, the algorithm can also been named as combat algorithm (CA). The partial overwrite operator is then applied to the less-fit chromosome with a probability that increases as the difference between the combat chromosomes becomes larger. If there are not any significant differences between the fitness values of the selected chromosomes, normal crossover is been applied. The following basic steps can better explain the partial overwrite operation:

Step.1: Choose randomly two chromosomes as candidates for the solution

Step.2: Calculate their fitness values

Each “bits” or genes of the most fit chromosome is selected or not with an equal probability. This is referred as “uniform bit selection”.

Step.3: Change the genes of the less-fit chromosome with the selected bits of the most-fit chromosome in the previous step. Therefore, this step can be considered as a one directional crossover since the genes of the most-fit chromosome remain untouched.

This procedure eliminates the dominance of the fittest individual over all other members in the population. The roulette-wheel selection rejects the less-fit chromosomes while the partial overwrite operator rejects only randomly selected genes of it. Therefore the dominance of the most-fit chromosomes over the less-fit ones is smoothened. As a consequence, the new algorithm is successful in all kinds of fitness or objective functions without having the danger of being entrapped in a false optimum.

The new genetic algorithm can be summarised as follows:

Step.1: Form an initial generation of m chromosomes in a random manner.

Step.2: Choose two chromosomes randomly from the initially generated pool.

Step.3: Calculate the fitness values for the two randomly selected chromosomes and then form the relative difference by using

$$\Delta_r = \frac{|f_1 - f_2|}{f_1 + f_2} \tag{1}$$

where f_1 and f_2 are the fitness values of the chromosome$_1$ and chromosome$_2$, respectively.

Step.4: Compare the fitness values of the two selected chromosomes. If there is a relatively high difference in the fitness values, then the partial overwrite operator is applied. This means that the less-fit chromosome loses some information. If the difference is relatively small then the classical uniform crossover operation is applied. In mathematical terms:

If $f_1 < f_2$ and $\lambda < \Delta_r$, where λ represents a random number drawn between [0,1] then apply only partial overwrite on crossover$_2$. However, if $f_1 < f_2$ and $\lambda > \Delta_r$, then apply uniform crossover.

If $f_1 > f_2$ and $\lambda < \Delta_r$, then apply partial overwrite on crossover$_1$. However, if $f_1 > f_2$ and $\lambda > \Delta_r$, then apply uniform crossover.

Step.5: Apply mutation operator with a probability of $1/m$.

Step.6: Return to the Step 2 until a stopping criteria has been accomplished.

Unless otherwise stated, the population size m, the number of parameters or search space dimension n, and the resolution or the number of bits for each parameter is chosen to be equal to 30, 1055, and 16, respectively. The maximum value that a parameter can take is limited to 1. In this study, the performance measure is taken as the number of false outputs in a training process of 120 different signal types. Since the process output is a bit of data, it is expected the initial number false outputs is around 60, since there are equal number of alarming or ordinary signals. The training performance has been averaged over 20 independent trials and the average of the best fitness values is taken. The uniform crossover probability for the classical GAs is set to unity. The mutation probabilities for both algorithms are set to be equal. During the training process, decrement in the number of false outputs is given in Fig. 5.

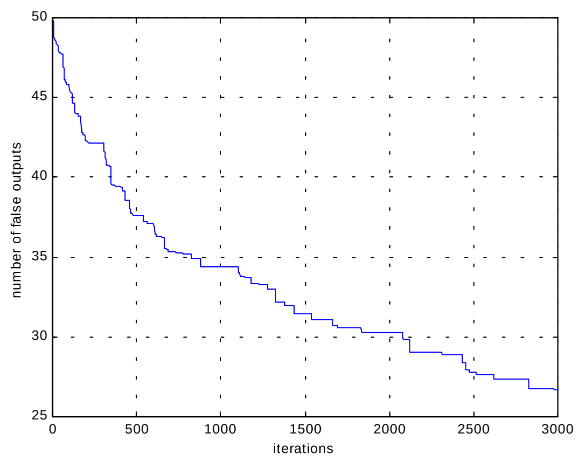


Figure 5. The decrease of false outputs by using GA

V RESULTS AND DISCUSSION

It has been demonstrated that combined wavelet–multi layer perceptron method can be used as a tool for exploratory earthquake prediction based on monopolar electric field measurement network data. This paper also has shown us that with the combined implemented analysis and conditioning of monopole electrode data, one can recognize anomalies as precursory signs of earthquake.

Our analysis has come up with the following results a) the measurement of electrical field patterns assumed to be depending on rock stress change and with the occurring pre-earthquake sign patterns emanated from this electrical field were recognized with a high-success probability. b) In this study, with the increasing number of the anomaly patterns, the success rate of capturing those patterns was being improved gradually. c) this method can also be used in recognizing other extraordinary events in nature.

VI REFERENCES

[1] TIECHENG, W., BAOYIN, W., 'The Tidal Feature of Well Water Level and Its Anomalous Variations Before Earthquakes', *Journal of Earthquake Prediction Research*, **4**(3), 1995.

[2] YONG, S., WAI, Z.: 'The Correlation Between Radon Variation and Earth Solid Tide Change in Rock-Groundwater System', *Journal of Earthquake Prediction Research*, **4**(3), 1995.

[3] REILINGER, R., BARKA, A.A.: 'GPS Constraints on Slip Rates in the Arabia-Africa-Eurasia Plate Collision Zone', NATO ASI Series, **28**, pp:91-108, 1997

[4] CANYARAN L., ÜSTÜNDAĞ B.,: 'Earthquake Forecast System', Turkish Patent Institute, Application No: 1999/02911, 1999.

[5] ZURADA, J.M., 'Introduction To Artificial Neural Networks', West Publishing Com., 1992.

[6] ROJAS, R., 'Neural Networks, A Systematic Introduction', Springer, Germany, 1996.

[7] KOHONEN, T., 'Content Addressable Memories', Springer-Verlag, New York, 1980.

[8] KARLIK, B., 'Çok Fonksiyonlu Protezler için Yapay Sinir Ağları Kullanılarak Miyoelektrik Kontrol', PhD Thesis, Yildiz Technical University, 1994.

[9] EFE, M.Ö., KAYNAK, O., 'Yapay Sinir Ağları ve Uygulamaları', Boğaziçi University, 2000.

[10] ENGLEHART, K., HUDGINS, B.: 'A Wavelet-Based Continuous Classification Scheme for Multifunction Myoelectric Control', *IEEE Trans. Biomed. Eng*., 2001, 48, pp.302-311

[11] GOLDBERG, D. E.: 'Genetic Algorithms in Search, Optimization and Machine Learning' (Addison-Wesley, 1989)

[12] HOLLAND, J. H. 'Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence' (MIT Press, Cambridge, 1992 2nd edn.)

[13] ED, L. D.: 'Handbook of Genetic Algorithms' (Van Nostrand Reinhold, New York, 1991)

[14] BÄCK, T., HAMMEL, U., and SCHWEFEL, H.P.: 'Evolutionary computation: comments on the history and current state', *IEEE Trans. On Evolutionary Computation*, 1997, **1**(1), pp. 3-17.

[15] GLOVER, F.: 'Tabu- search – part I' *ORSA Journal on Computing*, 1989, **1**(3), pp. 190-206

[16] KARABOĞA, D., KAPLAN, A.: 'Optimizing multivariable functions using tabu search algorithm'. The 10th Int. Symposium On Comp. And Inf. Sciences, ISCIS X, October 1995, Turkey, vol.2, pp. 793-799

[17] KIM, J.H., MYUNG, H.: 'Evolutionary programming techniques for constrained optimization problems' *IEEE Trans. On Evolutionary Computation*, 1997, **1**(2), pp. 129-140

[18] SALOMON, R.: 'Evolutionary algorithms and gradient search: similarities and differences' *IEEE Trans. On Evolutionary Computation*, 1997, **2**(2), pp. 45-55

[19] VAVAK, F., FOGARTY, T.C.: 'Comparison of steady state and generational genetic algorithms for use in nonstationary environments' *Proceedings of the 1996 IEEE Conference on Evolutionary Computation,* IEEE Press, 1996, pp. 192-195