Chapter 3. IDENTIFICATION OF SEISMIC SOURCES WITH THE HELP OF NEURAL NETWORKS

1. INTRODUCTION

All over the world there is recently observed a sharply increased scope of scientific researches in the field of artificial neural networks (ANN). It is firstly explained by those potentialities, which ANN gives for resolving complex, often formalized-with-difficulty problems. The use of the neural network technologies ensures the sufficiently adequate representation of a model of the process or phenomenon to be researched, an opportunity to detect laws in a large flow of contradictory information, maintenance of connections between the essential factors, compression of processed information without loss of sense, as well as associative storage of it [1, 3, 5, 6].

Problems of classification and identification of images are especially effective settled on the basis of ANN [1]. The examples of such problems are problems for processing the seismic information (data set ordered in time and to spatial distribution) with the purpose to determine the nature of its source. This problem may be fully formulated as follows: if the available number of samples (sets of experimental data corresponding to one or another class) is limited, one should develop a neural algorithm (a structure and methods for learning the neural network), which will allow him to determine the nature of an event (earthquake-explosion), initiated the written oscillations of the earth surface according to a given seismogram. Solution of this problem can be considered as constructing of a mathematical model of the phenomenon, creating the initial data, and based on it separation of points belonging to a producing source, but not included into observations being studied.

The structure of such a model is represented as a set of artificial neurons – summing elements with many inputs and a single output. Each input is characterized by its weight, i.e. by the amplification/attenuation coefficient of a signal, and output – by a nonlinear conversion from a sum total of all the neuron inputs. According to levels of weight connections neurons are combined into layers, thus forming a multi-layer structure (Fig. 1, 4). However, this structure by itself is not a finished model. The important stage of its development is a learning procedure. Its sense consists in choosing the weight coefficients for every layer, at which the designed structure of neurons will give an adequate result to any specified input effect. Learning of the neural-network structure is implemented on training samples of data, characterizing the modeled process. In this case, the more is set of such samples, the less is the possibility to make a wrong decision when processing a sample, not participating in the learning process. The built in such a way neural-network model enables to perform the matching-up of any input data sample with one or another class of the source.

At the given stage of research we suppose to realize the solution of a problem for identification of a source of seismic information on the basis of two most adapted neural-network paradigms: a multi-layer fully connected neural network [2, 5, 7] and a Kohonen neural network [3]. The use of two different methods for solving the raised problem will allow us to select the most effective one (to quantitative and qualitative indexes), and also to continue investigations on the basis of their composition and application of other neural-network paradigms.

2. DESCRIPTION OF NEURAL ALGORITHM BASED ON A KOHONEN-LIKE NEURAL STRUCTURE

In this section we consider the implementation of solving the problem for identification of a source of seismic information on a two-level (two-layer) neural network of the following structure (Fig. 1):

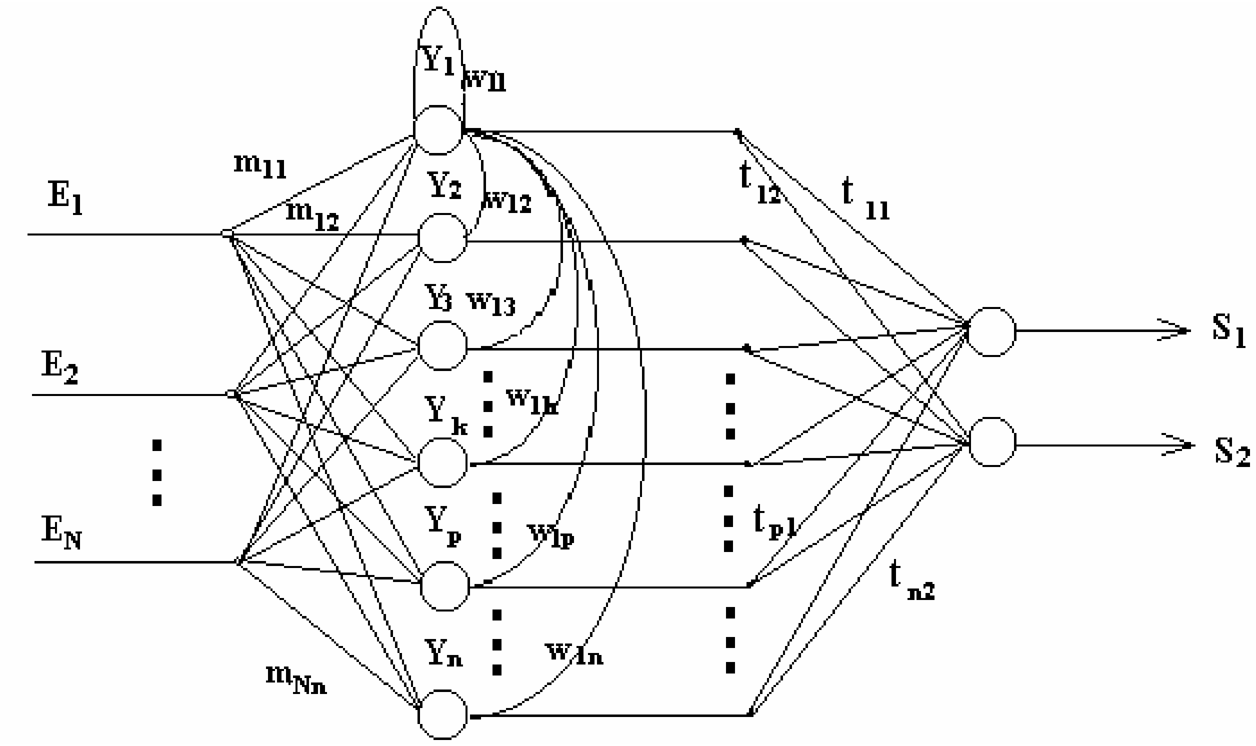

Fig. 1. Kohonen-like neural network

The first level is a two-dimensional Kohonen network, which presents a matrix of neurons Y_i $(i=1,2, \dots ,n)$, having dimension $n = k \times k$, with activation function *f(x)* (fig. 2) (matrix *Y* in Fig. 1 is shown for simplicity in the form of vector).

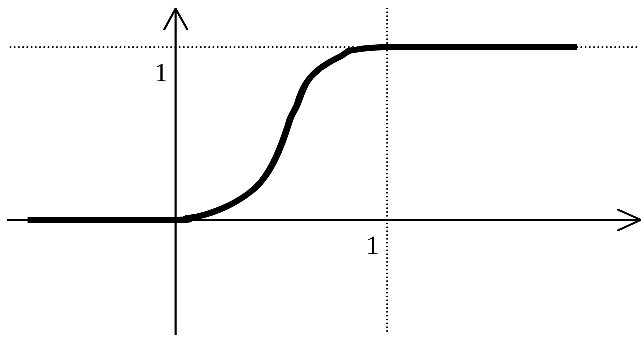

Fig. 2. Activation function

Input elements of network E_i $(i=1,2,\dots,N)$ are obtained by conversion of elements of researched sample $X = \{x_i; i=1,2,\dots,N)$ by the following formula:

$$E_i = \frac{1}{norm} exp(-\frac{x}{A}), i = 1,...,N$$

$$norm = \sum_{i=1}^{N} [x_i^T x_i]^{1/2}$$

A>1 is a constant. Such an exponential conversion gives good separation of signals, corresponding to different input situations.

For the first time coefficients m_{ij} $(i=1,2,\dots,N; j=1,2,\dots,N)$ are assigned with small random values from range (0,1), but further they are normalized:

$$m_{ij} = \frac{\mathrm{m_{ij}}}{\left(\sum_{i=1}^{N} \sum_{j=1}^{n} m_{ij} m_{ji} \right)^{1/2}}$$

The final values of weight coefficients m_{ij} agree with normalized input vectors, that is why, normalization prior to learning procedure approaches the weight coefficients to their final magnitudes, thereby reducing the learning process. Then at every step of learning process weights m_{ij} are adjusted according to expression:

$$m_{ij} = m_{ij} + \eta(E_i - m_{ij})Y_j , 0 < \eta < 1.$$

The process of modification of weights m_{ij} occurs, until the following condition is observed:

$$\left| \sum_{i=1}^{N} \sum_{j=1}^{N} m_{ij} E_j - \sum_{i=1}^{N} E_i^T E_i \right| \le \varepsilon, \varepsilon << 1$$

All the neurons in Y_i $(i=1,2,\dots,n)$ are connected with side (lateral) links. The side link means a function of distance, determined in the following way: neurons arranged nearby amplify each other, while more distinct neurons exert a braking effect. Therefore, we suppose to describe the degree of side interaction (coefficients w_{jk} $(j=1,2,\dots,n; k=1,2,\dots,n)$ by function of “Maxican hat” type (Fig. 3) [3].

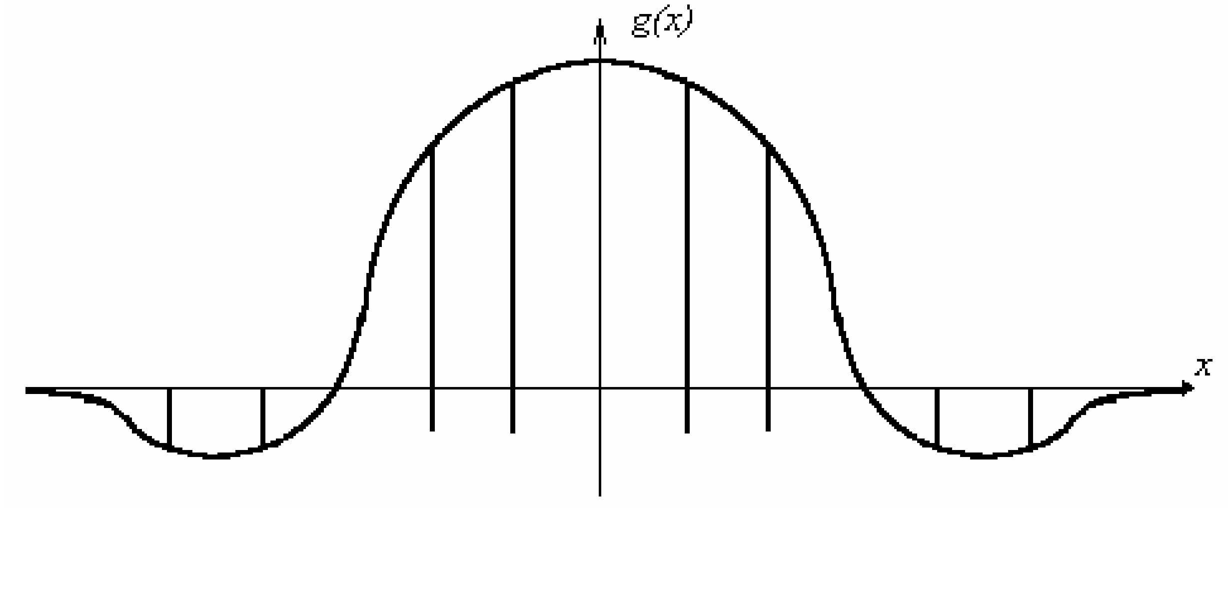

Fig. 3. Function “Maxican hat”

To avoid the “boundary effects” the first layer of neurons is made cyclic – neurons of the first and the last rows and columns of matrix Y we also connect with side links. As a result, the convolution of matrix Y into a continuous toroidal surface is conducted. Thus, the output of first-layer neurons is characterized by the following expression:

$$Y_j = f\left(\sum_{k=1}^{n} w_{jk} \sum_{i=1}^{N} m_{ki} E_i - \sum_{i=1}^{N} m_{ji} E_i - G\right), \qquad j=1,2,\ \dots\ ,n$$

In the course of learning (self-organizing) process at the output of the Kohonen layer there are formed clusters (a group of active neurons of the definite dimension, whose output differs from zero), characterizing specific categories of input vectors (groups of input vectors, corresponding to one input situation). Size of the group of active neurons depends on parameter

$$\gamma = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}\ ,$$

which defines the equilibrium between negative and positive forces of side links. Increasing γ expands the cluster, decreasing – narrows it. Dimension of the cluster also depends on threshold value *G*. The best value for this threshold can be found in the following way: *G* is lowered from the high value, at which no clusterization phenomena occurs, till a value, at which the specified cluster size is attained.

At every step of data processing input vector $\overline{E}$ matches up to one or another cluster. With the repetition of spatial situations clusters are amplified (outputs of cluster neurons grow) and, on the contrary, if spatial situations often change, they get attenuated.

The second level of the neural network is used for coding information. Weight coefficients t_{ij} *(i=1,2,..,n; j=1,2,…,n)* are calculated as follows:

$$t_{ij} = \begin{cases} 1,\ if\ Y_j\ and\ S_j\ are\,activated \\ 0,\ even\,is\,one\,of\ them\,is\,not\,activated, \end{cases}$$

i.e. after pre-training and forming of clusters in the Kohonen layer all the neurons of every gained cluster are connected at a phase of secondary learning by active (single) synapses with their output neuron, characterizing this cluster. The output of second-layer neurons is determined by expression:

$$S_j = f_1\left(\frac{1}{K_j} \sum_{i=1}^{n} t_{ij} Y_j - R\right), \qquad j = 1,2$$

where $f_1(x) = \begin{cases} 1\ ,if\ \ x > 0 \\ 0\ ,if\ \ x \le 0 \end{cases}$

K_j – dimension of j-th cluster, i.e. number of neurons of Kohonen layer, connected with a neuron of j-th output layer with coefficients other than zero,

R – a threshold value (0<R<1).

Threshold value R is selected in such a way, that, on one hand, the values of activated clusters were not lost, and, on other hand, “noise of non-activated clusters” was cut off.

As a result, at every step of processing of initial data we have at the output values S_j, that characterize the phenomenon, creating the given input situation (S_1=1 – earthquake, S_2=1 – explosion).