Probalistic Neural Network Application
to Warship Radio Station Identification

Tomasz Praczyk

Naval University
Śmidowicza 69, Gdynia, Poland
e-mail: T.Praczyk@amw.gdynia.pl

Abstract: The article presents use of PNN to identify ship's radio stations. Two methods of PNN acceleration are included. The first one is a combination of PNN and Kohonen neural network. The task of Kohonen neural network is to roughly classify radio station. The final identification is performed by PNN. The second speeding up approach consists in using average data instead of original one. Moreover, a modification of PNN decision rule is applied.

Key words: identification, neural networks

I. INTRODUCTION

There are a lot of identification methods, e.g. decision trees, minimal distance methods, Bayes classification methods, learning classifier systems and neural networks. Neural networks that have been successfully applied not only to pattern classification problems but also to function approximation, optimization, pattern matching, associative memories and robot control problems, possess many varieties which could serve as classifiers. The most popular of them are MLP, RBFNN, HRBFNN, PNN, KohonenNN and fuzzyNN. Moreover, there exists a lot of learning methods of neural classifiers e.g. backpropagation algorithm – BP, learning vector quantization algorithm – LVQ [16], family of constructive algorithms [17], genetic algorithms application to construct networks [4, 9, 14].
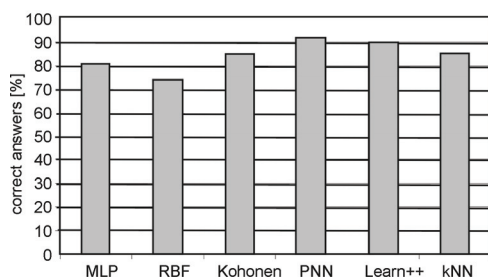

Fig. 1. Results of preliminary experiments

Majority of neural networks previously mentioned (MLP, RBFNN, PNN, KohonenNN) and additionally kNN method and Learn++ algorithm with set of MLP applica-

tion [18] were used during preliminary experiments to identify warship radio stations. The best results during the tests conducted accomplished PNN.

Unfortunately, in addition to unquestionable advantages, exhibited during the tests mentioned, PNN also possess one great disadvantage that could make it impossible to apply it in military systems. The drawback mentioned is slowness. In the paper two PNN acceleration methods are analyzed. Their application to the problem of identification of ship radio station is also presented in the paper.

The article is organized as follows. In the subsequent section PNN is outlined. In the further section two PNN acceleration methods applied during the experiments are introduced. Section IV describes radio signals and their representation used during the tests. Comparison between classical PNN and its modifications in identification of radio stations is included in section V and conclusions are made in section VI.

II. PROBABILISTIC NEURAL NETWORK

PNN classifies based on the decision function describing a membership degree of unknown pattern to given class. The role of this function plays the probability density function. Given that the density function for an object class is very often unknown its estimator is usually determined.

$$g(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}\sigma^p n}\sum_{i=1}^{n}\exp\left(-\frac{\|\mathbf{x}-\mathbf{x}_i\|^2}{2\sigma^2}\right) \qquad (1)$$

PNN uses the simplified form of the function above.

$$g(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^{n} \exp\left[-\left(\frac{d(\mathbf{x}, \mathbf{x}_i)}{p\sigma}\right)^2\right] \qquad (2)$$

where:

d is Euclidean distance between unknown pattern $\mathbf{x}$ and learning vector $\mathbf{x}_i$,

n is a number of learning vectors belonging to given class,

p is size of input space,

σ describes shape of every radial function.

The decision about assigning vector $\mathbf{x}$ to i-th class is based on the following relation:

$$h_i c_i g_i(\mathbf{x}) > h_j c_j g_j(\mathbf{x}) \underset{\text{class } j \neq i}{\forall} \qquad (3)$$

where

h_i denotes a priori probability that pattern $\mathbf{x}$ belongs to i-th class,

c_i is misclassification cost.

PNN consists of three layers: input, radial and output layer. The task of the first layer is to provide data to the next layer of neurons. The second layer contains radial neurons – as many neurons as the learning set possess patterns. Each neuron spreads a radial function over single learning pattern. The output layer sums signals coming from the previous layer. Every class has exactly one representative in this layer. Signals generated by radial neurons that represent the same class as output neuron are summed with weight 1. Weights of the remaining connections are 0. Additional elements of PNN architecture can be next layers corresponded to parameters h and c.
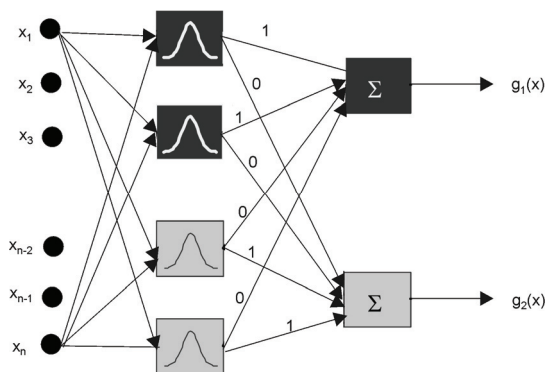

Fig. 2. Architecture of PNN

PNN possess all advantages and disadvantages connected with using learning data during reasoning. First of all, the learning process consists merely in memorizing learning patterns in the second, radial layer of neurons what is almost immediate activity. It is necessary to spend additional time solely on tuning parameter σ of radial neurons. A memorizing nature of PNN, especially in the case of software implementation of this network also induces negative effects. Analysis of large set of the learning data during reasoning causes long reply time. To obtain a network learned of every detail connected with a problem solved we have to add to the learning set new patterns what in turn prolongs performance time. Therefore, application of large learning sets entails necessity of simultaneous use of some acceleration methods. The paper presents several PNN improvements that speed up the reasoning process without sufficient deteriorating of performance of the system.

III. PNN ACCELERATION METHODS

In the paper two PNN acceleration methods are presented. Kohonen network is a key element of each of them. The first approach divides all learning data into groups and makes identification solely based on portion of them. The method is somewhat similar to kNN method where decision is taken based on k patterns that are placed in the nearest vicinity of the pattern unknown. The second method also relies on data grouping but in this case average data, fixed for each group, constitute the basis for reasoning. An additional element of this approach is the modification of the decision rule of the classical PNN.

The identification system implementing the first method consists of two main elements – KohonenNN and PNN. The task of Kohonen network is to roughly identify radio station. Two different learning philosophies could be used in relation to this network. Supervised learning could be applied (LVQ) and in this case Kohonen network distinguishes various signal encoding methods. An alternative solution is learning without a teacher and application any of unsupervised learning algorithms – SCS, neuron gas algorithm, learning with Gauss neighborhood etc. In this case the task of the network is to find group of signals similar to each other.

The next element of the system is the set of PNNs. Every network is attached to one neuron of Kohonen network. The task of PNN is accurate identification of a signal. The architecture of the overall system is depicted on Fig. 3.

The identification is performed in three phases. Initially, registered signal is subjected to feature extraction. The most informative part of signal spectrum is extracted and passed on to the input layer of Kohonen network. Kohonen network decides which PNN should be run and subsequently the network chosen performs ultimate identification of radio station.

The solution with two layers of networks accelerates functioning of the system. In the case of using classical
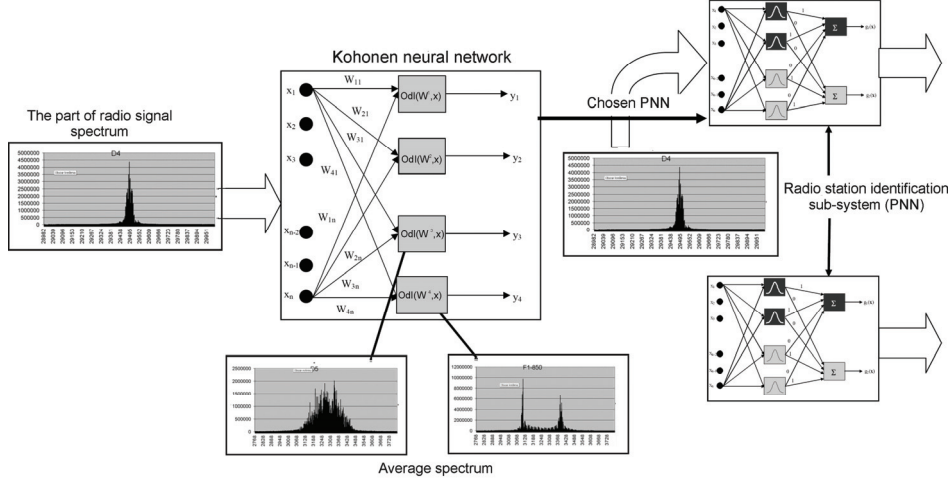
Fig. 3. The architecture of the system with the layer of Kohonen network

PNN without the layer of Kohonen network the system during identification is forced to take all learning data into consideration. Kohonen network causes that the number of data that the system has to analyze is limited to data located in the same cluster plus average patterns memorized in neurons of Kohonen network. Given that the number of neurons in Kohonen network is considerably less then the number of all learning data (one neuron of Kohonen network indicates one group of data) it is apparent that speed of the system proposed is higher then speed of classical PNN (except several unlikely situations).

The problem in the case of using unsupervised learning with respect to Kohonen network is to properly determine its number of neurons. Little number of neurons induces large clusters with large number of patterns and consequently not big acceleration of the system. In turn, large number of neurons increases speed of processing, however at the same time it causes that clusters are smaller and smaller and data located on the borderline between clusters could be misclassified. Simply, PNN designed based on sparse set of learning data, from one not large cluster, is in possession of insufficient knowledge to correctly identify radio station. Particularly, this involves areas of the cluster close to other clusters where knowledge of PNN could be mostly incomplete. For that reason we decided on the solution that consists in constructing PNN not only by means of data from one cluster (whose neuron won during competition) but also based on patterns from other adjoining clusters if the input signal lies in their proximity. In the case when the input pattern is located close to the center of the cluster, PNN solely contains data from this cluster. Thus, the input pattern always will be located in the center of some cluster (bigger or smaller) incorporating learning data for PNN. The decision whether to expand the learning set of PNN at data from the cluster corresponding to *i*-th

neuron of Kohonen network (neuron that lost during competition) depends on value of neighborhood rate:

$$\eta_i = \frac{d(\mathbf{x}, \mathbf{x}^W)}{d(\mathbf{x}, \mathbf{x}^i)} \qquad (4)$$

where $d(\mathbf{x}, \mathbf{x}^W)$ and $d(\mathbf{x}, \mathbf{x}^i)$ denote respectively distance between the input vector and the average vector memorized in the winner neuron and distance between the input vector and the mean vector from *i*-th neuron. If η_i is greater then some fixed threshold then the learning set is supplemented at data from the cluster indictaed by *i*-th neuron. Otherwise, PNN is constructed merely based on patterns from the cluster of winner neuron.

Another method of accelerating the system is use, as the learning data for PNN, average patterns that represent similar data grouped within one cluster. Initially, the whole set of the learning data is portioned into groups by means of Kohonen network. Next, each group is represented in the resultant PNN by the average vector. Usage of average patterns is well known method that speeds up memorizing solutions like PNN. However, making use of this method alone, without any change of PNN, can lead to deterioration in performance of the network. For that reason the modification of the same PNN is necessary. Function (2) is supplemented with the average vector "strength" parameter (*N*) that makes output value of every radial neuron dependant on the number of learning patterns located in the cluster of the average pattern:

$$g(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^{n_A} N_i \exp\left[-\left(\frac{d(\mathbf{x}, \mathbf{x}_i^A)}{p\sigma} \right)^2 \right] \qquad (5)$$

where $\mathbf{x}_i^A$ denotes the average pattern memorized in *i*-th radial neuron and n_A is the number of average learning vec-

tors belonging to given class. Lack of N parameter in the Eq. (5) would cause that places with single, alone learning vector as well as crowded areas of input space would be characterized by the same value of the decision function.

IV. DATASET

Radio signals used in the experiments come from warship radio stations. Frequency of signals is from the range 1,6-30 MHz. Military nature of radio stations causes that we deal with encoded signals. Crude radio signals are subjected to the feature extraction process. Initially, spectrum of every signal is fixed. Subsequently, the center of the most informative part of the spectrum is determined according to the following procedure. First, spectrum resolution is found,

$$R = \frac{f_s}{N_s} \qquad (6)$$

where f_s is sampling frequency and N_s denotes the number of samples of radio signal spectrum. In our case f_s is from the range 4 000 Hz-200 kHz and the number of samples is always 32 768. Further, the center sample is fixed,

$$S_c = \frac{f_i}{R} \qquad (7)$$

where f_i is an intermediate frequency. Ultimately, every radio signal is represented by 1000 samples – 500 to the left and 500 to the right from the central sample.
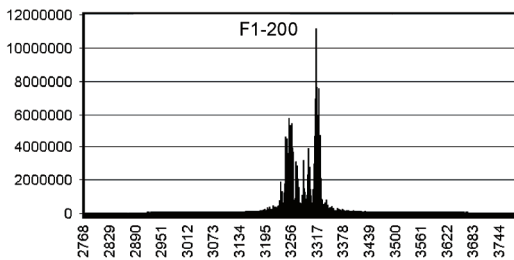

Fig. 4. Spectrum of the example radio signal

V. EXPERIMENTAL RESULTS

In the experiments two previously presented methods were compared with the classical PNN. Identification capability and reasoning acceleration in relation to the classical solution were examined. For this purpose data set consisting of 1200 radio signals was used. Signals, encoded in the same way, represented 6 different radio stations – 200 signals per radio station. The whole data set was portioned into two subsets of equal size– the learning (600 signals) and the testing part (also 600 signals). Generalization error constituted a measure of identification capability. This error in the case of use of the classical PNN amounted to 7% incorrectly identified signals from the test set. In turn, the total number of learning signals in relation to the average number of data analyzed by given solution, during identification of signals from the test set, determined acceleration ratio. During identification, the classical PNN always required analysis of all 600 learning signals. The system with the layer of Kohonen network needed data memorized in Kohonen network and data located in the cluster pointed out by the Kohonen network to classify single signal. In turn, the solution with use of average signals required all of them to identify a test signal. In order to prepare the layer of Kohonen network, unsupervised neural gas algorithm was used. Kohonen network and neural gas algorithm was applied to fix average radio signals. σ parameter of PNN radial neurons was fixed by means of the portion of the learning set.

The graphs (Fig. 5) show that the acceleration methods proposed are able to identify radio stations on the same level or even better then the classical PNN. The suggested approaches identify with similar result, simultaneously accelerating calculations 10 and almost 13 times (for 8% generalization error). Comparing both methods it is necessary to state that generalization error rises faster for the solution with use of average patterns. Reduction of the number of average learning patterns used to construct PNN speeds up identification but at the same time causes that the resultant decision function less and less accurately reflects distribution of individual data categories. In turn, the method of acceleration of inference used in the solution
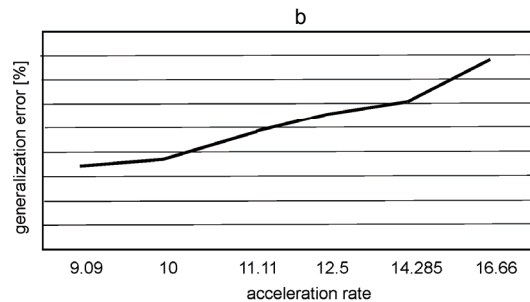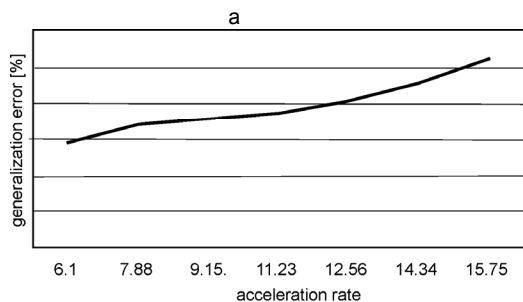

Fig. 5. Generalization error versus acceleration rate (a) – the solution with the layer of Kohonen network, (b) – the solution with use of average radio signals

with the layer of Kohonen network (i.e. adding new clusters) causes that knowledge of a network taking a decision is more and more local. The winner PNN often and often neglects information from further neighborhood that could be essential for identification process.

VI. CONCLUSIONS

The paper presents use of two acceleration methods of the classical PNN to identification of radio stations. The first method divides all learning data into groups and makes identification solely based on portion of them. The second solution also relies on grouping data but in this case average data, fixed for each group, constitute the basis of reasoning. An additional element of this approach is the modification of the decision rule of the classical PNN. During the tests conducted, both methods have proven that are able to speed up significantly inference and what is more to retain classification reliability demonstrated by the original solution. The experiments also exhibited that the method with the layer of Kohonen network somewhat outperforms the competitive solution with use of average patterns.

References

[1] M. Butz, *Rule-based Evolutionary Online Learning Systems: Learning Bounds,* Classification and Prediction, Scientific Literature Digital Library, http://citeseer.ist.psu.edu

[2] S. Chalup, F. Maire, *A Study On Hill Climbing Algorithms For Neural Network Training*, Scientific Literature Digital Library, http://citeseer.ist.psu.edu

[3] P. Cichosz, *Learning Systems*, WNT Warsaw (2000) (In Polish).

[4] D. Curran, C. O'Riordan, *Applying Evolutionary Computation to Designing Neural Networks: A Study of the State of the Art*, Scientific Literature Digital Library, http://citeseer.ist.psu.edu

[5] K. Doherty, R. Adams, N. Davey, *Hierarchical Growing Neural Gas*, Scientific Literature Digital Library, http://citeseer.ist.psu.edu

[6] Y. Freund, H. Seung, E. Shamir, N. Tihsby, *Selective sampling using the query by committee algorithm*, Machine Learning, 28, pp. 133-168, (1997).

[7] B. Fritzke, *A growing neural gas network learns topologies*, Advances in Neural Information Processing Systems, MIT Press, Cambridge, (1995).

[8] L. Gajek, M. Kałuszka, *Statistical Reasoning*, WNT, Warsaw (2000) (In Polish).

[9] F. Gruau, *Neural Networks Synthesis using Cellular Encoding and the Genetic Algorithm*, Scientific Literature Digital Library, http://citeseer.ist.psu.edu

[10] B. Guinand, A. Topochy, K. Page, M. Burnham-Curtis, W. Punch, K. Scribner, *Comparisons of Likelihood and Machine Learning Methods of Individual Classification*, Scientific Literature Digital Library - http://citeseer.ist.psu.edu

[11] J. Korbicz, A. Obuchowicz, D. Uciński, *Artificial neural networks*, AOW PLJ, Warsaw (1994) (in Polish).

[12] J. Kornacki, J. Ćwik, *Statistical learning systems*, WNT, Warsaw (2005) (In Polish).

[13] G. Kuchariew, *Processing and analysis of digital images*, Szczecin Univ. of Technology, Szczecin (1999) (in Polish).

[14] M. Mandischer, *Representation and Evolution of Neural Networks*, Scientific Literature Digital Library, http://citeseer.ist.psu.edu

[15] D. Michie, D. Spiegelhalter, C. Taylor, *Machine Learning, Neural And Statistical Classification*, Scientific Literature Digital Library, http://citeseer.ist.psu.edu

[16] S. Osowski, *Neural networks in algorithmic sense*, WNT, Warsaw (1996) (In Polish).

[17] R. Parekh, J. Yang, V. Honavar, *Constructive Neural Network Learning Algorithms for Pattern Classification*, Scientific Literature Digital Library, http://citeseer.ist.psu.edu

[18] R. Polikar, L. Udpa, S. Udpa, *Learn++: An Incremental Learning Algorithm for Supervised Neural Networks*, Scientific Literature Digital Library, http://citeseer.ist.psu.edu

[19] M. Raymer, L. Kuhn, W. Punch, *Knowledge Discovery in Biological Datasets Using a Hybrid Bayes Classifier/Evolutionary Algorithm*, Scientific Literature Digital Library, http://citeseer.ist.psu.edu

[20] R. Tadeusiewicz, M. Flesiński, I*mage recognition*, PWN, Warsaw (1991) (in Polish).

[21] D. Wilson, T. Martinez, *Reduction Techniques for Instance-Based Learning Algorithms*, Scientific Literature Digital Library, http://citeseer.ist.psu.edu

TOMASZ PRACZYK. *Education:* Military University of Technology, Warsaw – MSc (1996); Maritime University, Szczecin – PhD (2001). *Activities:* intelligent navigational systems, neural networks, genetic algorithms, neuroevolution, evolutionary reinforcement learning.