Application of ELM Algorithm in Characterization of Nonlinear Loads

Miona Andrejević Stošović, *Member, IEEE*, Marko Dimitrijević, *Member, IEEE* and Dejan Stevanović

Abstract—**In this paper artificial neural networks are used for identification of nonlinear loads on the network, as one of the methods for non-intrusive load monitoring. Specific parameters of a group of the nonlinear loads are first measured, and then a neural network is used in order to identify which load is active on the network. Performances of ELM algorithm are tested, in the sense of the activation function and number of hidden neurons. It is shown that the task of loads identification can be done successfully.**

Index Terms— **ELM algorithm, artificial neural networks, nonlinear loads.**

I. Introduction

WE are living in the energy efficiency era, and many different researches have been led recently in order to explore the ways to save the energy. The process of saving energy consists of numerous procedures that need to interact in order to give results. One of these aspects should provide information to the user about the consumption of individual electrical devices, so the user can manage the device operation.

There are many methods that are being used in monitoring energy consumption (Appliance Load Monitoring-ALM). Depending on the method of data collection, these are divided to two categories: Intrusive Load Monitoring – ILM and Non-Intrusive Load Monitoring – NILM.

ILM method is implemented with a number of sensors that are placed on each electrical device whose consumption is monitored. However, NILM method requires the existence of only one measuring device for the entire household. Although ILM methods are more accurate, NILM methods are much more common in practice because the realization costs are significantly lower.

Most of the known monitoring methods for identification devices use neural networks. We will show in this paper application of one of the popular algorithms for neural network training, here applied to characterization of nonlinear loads.

Realization of such a monitoring system consists of three phases: data acquisition, parameter extraction and device classification. The first step, the data acquisition, is done in the power meter. At this stage it is necessary to provide sufficient speed of sampling data, and it primarily depends on which parameter is analyzed. After collecting data, it is necessary to extract corresponding parameters. We can monitor the current and average values that can provide some information about the connected devices, but it is also common to monitor the time-dependent values, or to detect events on the network. When parameters are extracted, data is analyzed in order to identify the consumer.

This paper is structured as follows: in the second chapter we will shortly explain ELM algorithm for neural network training. Then in the third chapter we will describe measurement of the properties of nonlinear loads using virtual instrument. Characterization of the loads using neural networks is presented in the fourth chapter, and after that conclusions are given.

II. ELM Algorithm

"ELM" is short for "Extreme learning machine". This algorithm is named like this because neural networks are trained very fast when using this algorithm, even thousands of times faster than networks trained using backpropagation algorithm. The key points that differ ELM from traditional popular gradient-based learning algorithms for neural networks are [1]:

- The learning speed of ELM is extremely fast.
- In many cases ELM has better generalization properties than gradient-based learning.
- ELM algorithm is much simpler than most learning algorithms for feedforward neural networks.
- ELM algorithm can be used for training networks with many non-differentiable activation functions.

It is important to stress that this algorithm is used only for training single-hidden layer feedforward neural networks (SLFNs), thus it can be used for limited number of applications. Moreover, using ELM, SLFNs with N hidden nodes can exactly learn N distinct observations [1].

In other algorithms, all the parameters of the network need to be tuned, so there is dependency between different layers of parameters. So, these gradient-based algorithms can be very slow and may easily converge to local minima. These are the most usual problems occurring in neural networks training.

In order to avoid these usual problems, the authors tried to bypass them, so the main difference between these "classical" algorithms and ELM is that input weights and hidden layer

Miona Andrejević Stošović and Marko Dimitrijević are with the University of Niš, Faculty of Electronic Engineering, 14 Aleksandra Medvedeva, 18000 Niš, Serbia (e-mail: miona.andrejevic@elfak.ni.ac.rs, marko.dimitrijevic@elfak.ni.ac.rs).

Dejan Stevanović is with Innovation Centre of Advanced Technologies (ICAT), Bulevar Nikole Tesle 61, 18000 Niš, Serbia. E-mail: dejan.stevanovic@icnt.rs.

thresholds are randomly chosen, so they are called random hidden nodes. By introducing this new concept, the learning process is much simplified since these parameters, which are very numerous, do not need to be tuned. Some simulation results showed that this method makes learning extremely fast, as well as it produces good generalization performance [2], [3].

This algorithm has on more limitation. Namely, activation function used in the hidden layer must be infinitely differentiable in any interval, for N arbitrary distinct samples. Such activation functions include the sigmoidal functions as well as the radial basis, sine, cosine, exponential, and many others.

III. Parameters measurement

In linear circuits, the electric quantities – currents and voltages – are trigonometric functions of time. In this case, the power factor is defined as cosine of the phase difference between voltage and current sinusoidal waveform. When circuit consists of nonlinear loads, one should define new parameters related to higher harmonics and redefine power factor and notions of apparent and non-active power. Therefore, power factor should be generalized to a total power factor, where the apparent power, used in its calculations, includes higher harmonic as well. This is crucial in description and design of present day power systems that contain non-linear loads such as voltage rectifiers, and switched-mode power supplies. Phase difference between current and voltage waveforms and existence of higher harmonics has negative impact on power distribution system. The second phenomenon is often referred as *distortion*.

Having in mind that the problem of distortion becomes global, it can be either observed globally at the distribution system level, or locally by performing measurement of the properties of nonlinear loads.

The power factor, harmonics and distortion power measurements, requires special instruments. For example, use of a classical instrument will provide incorrect results when attempting to measure the AC current drawn by a non-linear load and then determine power, or the power factor. A true RMS ammeter and voltmeter must be used to measure the correct RMS electrical quantities and apparent power. To determine the real power and reactive power, a true RMS wattmeter designed to properly work with arbitrary waveform currents must be used.

In recent papers [4], [5], [6] a new approach to nonlinear polyphase load analysis is given: system with advanced options for nonlinear load characterization.

The system combines the paradigm of virtual the instrumentation with core advantages of classical instruments – determinism in measurement. The hardware of the system is realized using field programming gate array (FPGA), which controls data acquisition. The software is implemented in two parts: real-time application for parameter calculation that executes on real-time operating system (RTOS) and application for user control, data analysis and visualization that runs on general purpose operating system (GPOS). Described system allows measuring and calculating a number of parameters that characterize nonlinear loads, unobtainable with classical instruments. Main advantages of the system are scalability, openness and flexibility: it can be extended in number of calculated parameters, in number of independent measurement channels, as well in functionality. It can be employed to be a part of harmonic compensation circuitry [5] or aimed for hardware-in-the-loop and power-in-the-loop applications. Flexibility is reflected trough different implementations on different platforms for different purposes: laboratory equipment for real-time measurements (PXI controller with PXI-7813R FPGA card and external expansion chassis), compact industrial computer for real-time operation (installed on programmable automation controller) or simple portable instrument with USB interface. It consists of three components: acquisition subsystem (hardware implemented), real time application for parameter calculations (RTOS) and virtual instrument for user control, additional analysis, visualization and data manipulation (GPOS).

Acquisition subsystem comprises of acquisition modules for A/D conversion, FPGA circuit and interface for programmable automation controller. A/D resolution is 24-bit, with 50000 samples per second sampling rate and dynamic range ±300 V for voltage measurements and ±5 A for current measurements. FPGA circuit performs acquisition control and basic harmonic analysis.

Real time application calculates apparent, active and reactive power and various power quality parameters deterministically. Calculated quantities and parameters are stored on local storage. The application runs on real time operating system.

Virtual instrument for user control, additional analysis, visualization and data manipulation is user interface of described system. It can run on any personal computer, separately from rest of the system. Communication between two software components (real time application and virtual instrument) is achieved by TCP/IP. Electrical quantities and parameters can be presented numerically (RMS values, power, power factor, etc.) and graphically (waveforms, spectra, phasor diagrams).

For the purpose of this paper, we measured following quantities:

1) I_{RMS} – RMS value of current;
2) THD_{I} – Total Harmonic Distortion for current;
3) P – Active power;
4) I_{H} – Current RMS related to higher harmonics: $I_{\mathrm{H}} = \sqrt{\sum_{k=2}^{N} I_k^2} = \sqrt{I_{\mathrm{RMS}}^2 - I_1^2}$. I_{k} stands for RMS value of k-th harmonic and N is order of highest harmonic taken into calculation.

IV. Characterization of the Loads

The measured parameters will be now used in the process of loads characterization. We measured these parameters for three different devices: computer, monitor and bulb. We also measured parameters for the combinations of these devices, so in the Table I we presented these combinations, and also coded them. Each combination has to have its code, so that it can be recognized by the artificial neural network.

TABLE I
CODES FOR DEVICES

Code	Device
1	Computer and monitor
2	Computer and bulb
3	Computer, bulb and monitor
4	Computer
5	Monitor
6	Bulb and monitor
7	Bulb

In the Table II parameters presented in the previous paragraph for the specific devices are given. THD_I and P are given in the normalized form, since it is the form to be presented to the neural net. I_{RMS} and I_H are in the range [0, 1], so they do not have to be normalized.

TABLE II
CHARACTERISTIC PARAMETER VALUES FOR THE SPECIFIC DEVICES

Code	I_{RMS} (A)	THD_I	P	I_H (A)
1	0.415468	0.928061	0.347265	0.319782
2	0.604193	0.255288	0.765179	0.190188
3	0.732451	0.327231	0.9085	0.286547
4	0.244795	0.951062	0.200015	0.190278
5	0.17384	0.82316	0.152303	0.126972
6	0.545821	0.180137	0.705366	0.124369
7	0.424189	0.01986	0.568683	0.01094

When training the neural network, it is necessary to specify input file containing measured data and decided response (code from the Table), but we also need to specify ELM-method, classification or regression. Since we want to classify our loads, we specify classification. We proved that regression is not good method in Fig.1, where it is obvious that trained neural network does not work properly, though we used great number of neurons in the hidden layer (16).

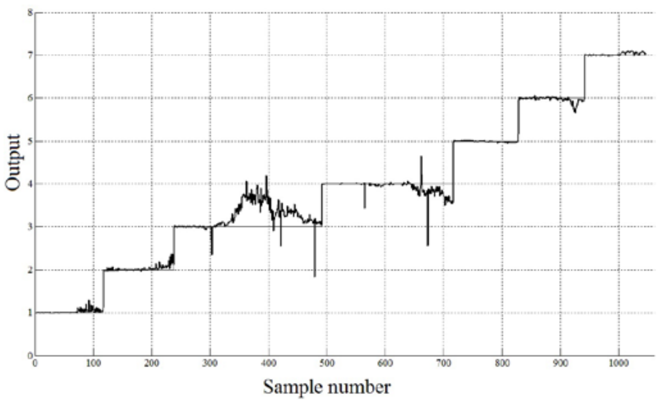

Fig. 1. Response of the artificial neural network with sigmoid activation function and 16 neurons in the hidden layer- Regression method

Data to be specified before the training process are also activation function in the hidden layer and number of neurons in the hidden layer. It is necessary to be careful when choosing number of neurons in the hidden layer, since too many neurons can cause overfitting problem [7], and when number of neurons is too small, the network is not trained properly.

In the Table III we gave examples for different number of hidden neurons, and we can see how this affects the training error. In the Fig. 2 we presented response of the artificial neural network with sigmoid activation function and 7 neurons in the hidden layer. From the Fig. 3 we can see that response has several incorrect values, but this is expected because this number of neurons is not enough. From the Fig. 3 we can see that we obtained satisfying response with 16 neurons in the hidden layer. Namely, for the specified combination of input parameters, the neural network recognizes what devices are active on the network by outputting the code of this specific combination.

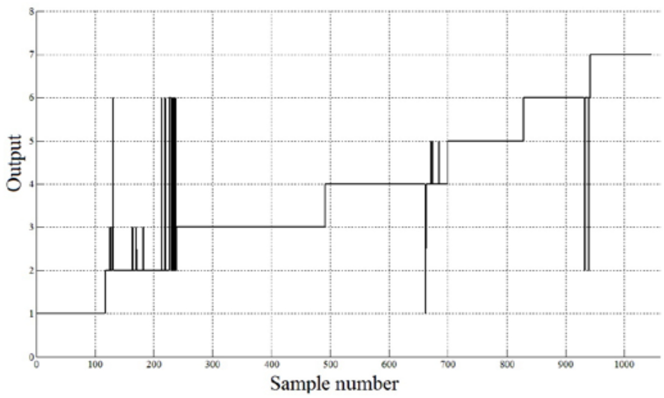

Fig. 2. Response of the artificial neural network with sigmoid activation function and 7 neurons in the hidden layer- Classification method

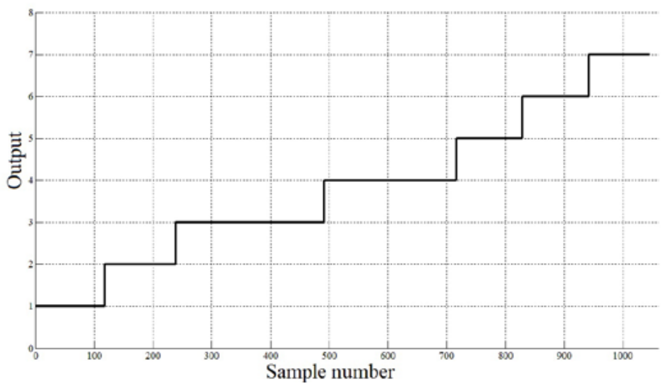

Fig. 3. Response of the artificial neural network with sigmoid activation function and 16 neurons in the hidden layer- Classification method

Since previous figures give only the general idea about training accuracy and number of hidden neurons, detailed data are systemized in the Table III. We have to emphasize that network training is performed with one set of data, and training results are given in the first column of the Table III. Examples are for different number of hidden neurons and different activation functions. Sigmoidal, sinusoidal and hard-limit (hardlim) activation functions are used. Generalization property of the network is confirmed by applying completely new set of data to its inputs. The results of these tests are given in the second column of the table. We can notice that when we use sigmoid and sinusoidal activation function, excellent results are obtained with 9 neurons, which is not the case for other activation functions. The last three rows in the table show that sometimes number of hidden neurons can be oversized, so more neurons in the hidden layer do not always mean better accuracy.

TABLE III
CHARACTERISTIC PARAMETER VALUES FOR THE SPECIFIC DEVICES

Training accuracy	**Testing accuracy**	**Number of hidden neurons**	**Activation function**
98.88%	96.37%	7	sig
100%	95.60%	7	sin
83.99%	77.72%	7	hardlim
100%	97.71%	8	sig
100%	93.88%	8	sin
71.63%	77.15%	8	hardlim
100%	99.24%	9	sig
100%	99.81%	9	sin
84.83%	86.90%	9	hardlim
100%	100%	10	sig
100%	100%	10	sin
85.67%	88.43%	10	hardlim
100%	100%	11	sig
100%	100%	11	sin
85.67%	89.29%	11	hardlim
100%	100%	15	sig
100%	100%	15	sin
93.82%	85.85%	15	hardlim
100%	100%	16	sig
100%	100%	16	sin
96.63%	94.17%	16	hardlim
100%	100%	25	hardlim
99.9%	99.9%	14	sin
100%	99.81%	30	sig
100%	99.81%	30	sin
95.22%	94.36%	30	hardlim

.

V. CONCLUSION

We examined in this paper characteristics of the ELM algorithm that was not used before for nonlinear loads identification. Since we obtained excellent results, we conclude that this algorithm can be used for much larger set of loads, in order to recognize which group of loads is active on the network.

ACKNOWLEDGMENT

Results described in this paper are obtained within the project TR32004 founded by Serbian Ministry of Science and Technology Development.

REFERENCES

[1] Guang-Bin Huang, Qin-Yu Zhu, Chee-Kheong Siew, Extreme learning machine: Theory and applications, Neurocomputing, Volume 70, Issues 1–3, December 2006, Pages 489-501, ISSN 0925-2312, http://doi.org/10.1016/j.neucom.2005.12.126.

[2] G.-B. Huang, Learning capability and storage capacity of two-hidden-layer feedforward networks, IEEE Trans. Neural Networks, 14 (2) (2003), pp. 274–281.

[3] S. Tamura, M. Tateishi, Capabilities of a four-layered feedforward neural network: four layers versus three, IEEE Trans. Neural Networks, 8 (2) (1997), pp. 251–255.

[4] M. Dimitrijević, “Elektronski sistem za analizu polifaznih opterećenja baziran na FPGA”, Doktorska disertacija, Univerzitet u Nišu, 2012.

[5] M. Dimitrijević, D. Stevanović, M. Andrejević Stošović, M. Petronijević, P. Petković, *Compensation of Harmonics in Smart Grid Using Photovoltaic Energy Source*, XI International Symposium on Industrial Electronics - INDEL 2016, Banja Luka, 03.11.-05.11. 2016, pp. 1-6.

[6] S. Djordjević, S. Bojanić, M. Dimitrijević, D. Stevanović, Monitoring potrošnje električne energije zasnovan na kontinualnom praćenju aktivne i snage izobličenja, Zbornik LIX konferencije ETRAN, Srebrno Jezero, 8.6.-11.6., 2015, EL2.3.

[7] T. Masters, “Practical Neural Network Recipes in C++”, Academic Press, San Diego, 1993.