XVII IMEKO World Congress
Metrology in the 3rd Millennium
June 22–27, 2003, Dubrovnik, Croatia

NEURAL-LEXICAL CLASSIFICATION OF POWER QUALITY INSTRUMENTATION MULTIPLICITY

Maurizio Caciotta, Michela Cristiani

Electronic Engineering Dept. of "Roma Tre" University - Rome - Italy

Abstract – The problem of perceived electrical power quality by users of the Little Medium Enterprises (LME) is settled and matched with the difficulty to individuate the correct instrumentation that is able make their satisfaction. The exorbitant quantity of instruments that are on the market, that, the builders assure, measure power quality, cannot be completely managed by operators. The paper present a system based on automatic lexical procedure that read the manuals and extract the information that can be properly managed by a Kohonen non supervised neural network to obtain an optimal inner classification in view to cross data with an inquiry on effective perceived quality by users.

Keywords: Power Quality, Kohonen Neural Network, Lexical Analysis

1. INTRODUCTION

Properties and characteristics of both products and services, able to satisfy implicit or expressed demands of consumers, define quality (ISO 9001). This concept is subjective and varies by consumers. In the case of electrical energy supply, one must consider that service is also depending on consumers interaction between both the suppliers and the other consumers, and this suggests the necessity of deep knowledge of problems about electrical energy supply quality, giving support to minimum characteristics definition. The voltage supply anomalous behavior causes technical and economic consequences on the users. The technical ones are linked to noise sensibility of apparata that cause failure stoppage of work and damaging of devices. The economic consequences depend on kind of productive process and by the amount of losses and the maintenance operations linked to stoppage of work. Technical troubles interesting consumers [1] of the electric power are of many kinds and consequences are very different. A low level of supply quality, make difficult the industrialization of depressed geographic areas [2] obliging operators to acquire expensive continuity group (UPS-Uninterruptible Power Service) with a great energetical capability especially in sensible processes in factories of semiconductors, pharmaceutical, plastic, etc. [3-4]. The technical rules about the European distribution network are the CEI-EN 50160. The sections in which the power quality problem, seen by the Little Medium Enterprises (LME), has been subdivided are the section of perceived quality and the section of the instrumentation analysis disposable on the market. From the compared analysis of these two aspects we verify the instrumental covering of perceived demands by electric energy users. The aim of this work is to match the perceived quality of the electrical power with the suitable instrumentation on the market or, if not, to indicate the typology of the instruments needs to resolve determined problems. This work is pointed out to analyze the second of the problems.

2. INSTRUMENTS FOR THE POWER QUALITY

The instrumentation on the market connected with the measurement of all the quality parameters series, is exorbitant. It is impossible to try a classification, for the difference in the interpretation of the quality we want to research. We are convinced that the cross between the perceived quality and the instrumentation classification can help us to clarify what is happening on the market and which instruments classes are the best to identify the real parameters of perceived quality. An automatic analysis of instrument catalogues and manuals can allow winning an informative "impasse" caused by the complexity and numerosity. The first act has been to collect a certain number of instrumentation catalogues and manuals on the Power Quality on Internet, but not only, and distributed on certain number of industrial producers, to settle the Neural Network Parameters, extending, after, to the biggest number possible of manuals of the great as possible of producers. After, one has been decided to proceed in two phases: automatic lexical analysis of the text at first and than the neural classification through non-supervised Kohonen network.

3. LEXICAL ANALYSIS

To utilize the classification power of a non-supervised neural network the more specialized and frequent terms need to be known. The frequency of these words will be the numerical input of the neural network. Also the research of the repeated expressions is an analysis type useful, because many specialized characteristics are construct by phrases. To separate the expression and to put in evidence the most repeated, they can give important key words to the researchers. With this automatic system, often are put in evidence terms that we have not thought be so significant and that can allowed the success on resolve the classification problem. For this a lexical analyzer [5] has been chosen to utilize a database developed in ACCESS compass, able to:

- Scan word by word each access text
- Put a text word in a table
- Remove all the symbols different from the alphabetical ones (punctuation, mathematical symbols, etc.)
- Eliminate the terms of grammatical type unuseful for us (articles, conjunctions, prepositions, etc.), determine how many times each term appears in the text and to calculate its frequency.

Terms	ACR	AEMC	AMETEK	ARNETT
1mv	0	0	0	0
500v	0	0	0	0
ac	0	0.001447	0	0
accuracy	0.015707	0.004342	0	0
accurate	0	0	0.00346	0.003745
active	0	0	0.00346	0
advanced	0	0	0	0.015707

a)

Terms	Recurrence	Frequencies
power	13	2.17E-02
dip	11	1.84E-02
current	7	1.17E-02
analysis	6	1.00E-02
inputs	6	1.00E-02
transducers	6	1.00E-02
pulses	5	8.35E-03

b)

Fig. 1: a) a partial examples of word extracted by manuals of instrumentation; b) an example of a schedule of the lexical database with words, absolute occurrence and relative frequency on the text

The next step has been to put all the technical terms (about four hundreds terms on fifty two documents), arbitrarily chosen having frequency, in each document, greater than 1%, in a database

Successively we consider, inside this fusion, one hundreds and nineteen terms, the most important in our arbitrary judgment, being the input vector of Kohonen neural network.

4. KOHONEN NEURAL NETWORK

The Kohonen Neural Network (K.N.N.) [6-7-8-9] is able to classify different input according to "nearness" criterion that discovers and extract features from the input data in order to create bi-dimensional output of neurons as shown in Fig. 2, defining regions that responding to a spread of values around the trying input

The categories classifications will be limited to a nine for a better intelligibility of human brain. Instrumentations that can be applied to electrical power quality parameters determination would require a loop to obtain convergence of neural network so large that the computing time could be very length. More, if a new set of few instruments appears on market, all the convergence procedure must be repeated.

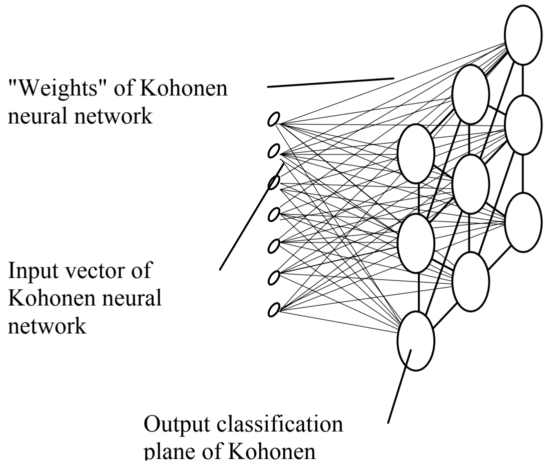

Fig. 2: general structure of a Kohonen neural network where all inputs are connect to every node network

Saving time we consider the possibility to choose set of limited number of instrumentations that will be treated as standard set. This procedure can be defined “Standard Settling” (S.S.).

Successively the great part of the instrument set, will be classified by “sterilized” K.N.N. eliminating the updating foreseen by the classical procedure.

All ones that, in working cycle of K.N.N. shall not be clearly identified can integrate the standard set of instruments, initially arbitrary.

In such an eventuality, an updating at the S.S. is run and a new classification tool is defined.

So three steps describe the procedure:

1. Kohonen Neural Network procedure on “Standard Set” (Standard Settling)
2. Minimizing of input number and possible declaration significance of firing a neuron.
3. Working cycle of “sterilized” K.N.N.

4.1 *Standard Settling*

The convergence algorithm organizes the nodes in the grid into local neighborhood that acts as feature classifiers on the input data. The output bi-dimensional map is self-organized by a loop process that "compares" all values of input vector with the "weights" attributed to each connection with the plan of output neurons. It is calculated the value of the minimum Euclidean distance and the "winner" neuron is considered. A certain neurons number around it, is revised with a range of action that decreases with the number of considered iterations. The start-up foresees, the network weights are settled to random values. The updating rules are:

$$d_{i,j} = \sum_{1}^{n} k\, [X_k(t) - W_{i,j,k}(t)]^2 \quad (1)$$

where n is the input vector length and t is the updating state.

$$W_{i,j,k}(t+1) := W_{i,j,k}(t) + \alpha(t)\,[X_k(t) - W_{i,j,k}(t)] \tag{2}$$

$$\alpha(t) = A * e^{-t/T_1} \tag{3}$$

$$N_C(t) = A_1 + A_2\, e^{-t/T_2} \tag{4}$$

This is the competitive learning. For each presented input to the network the algorithm compute the Euclidean distance $d_{i,j}$ between the input $X_k(t)$ and all the output node (i, j) and selects, as winner neuron, node that shows the minimum Euclidean distance. Weights of this neuron $W_{i,j,k}(t)$ and its neighbors is update. The algorithm is looped with a new input.

The deep of neighbors updating, will be decreased elapsing convergence procedure by α(t) in which A, A_1, A_2, T_1 and T_2 are the characteristics parameter, depending on application field of K.N.N. The extension definition of the neighborhood $N_C(t)$ is function of two parameters:

$$A_1 = \sqrt{I \cdot J} - 1 \tag{5}$$

and A_2. In this rule I e J are characteristic parameters limits, respectively, of i and j, that force $A_1 = 2$. The other parameters has been settled in experimental way utilizing fifty two instrument manuals that indicates:

$$0 < A < 1 \quad ; \quad 0 < A_2 < 1 \tag{6}$$

With a number of iterations between

$$50 < T_1 < 2100 \quad ; \quad 50 < T_2 < 2100 \tag{7}$$

the convergence is assured. The procedure has been optimized with the choosing $T_1 = T_2$. The preliminary experimental results has been obtained using

$$t > 200 \quad @\,A = .8; \quad @\,A_2 = .5 \tag{8}$$

These parameters settling will allows the utilization of the neural network, increasing the number of instruments considered for the classification. The word's frequency level arbitrarily established to 1%, expands input vector length, to 199 elements, as before told. It is possible that this great number of inputs make confused the classification capability of the neural network. To establish the real useful words necessary to obtain an optimal classification, inputs will be eliminated on the basis of the mean frequency calculated on all instrumentation considered compared with the weights settled by the neural network in convergence situation. Will be eliminated all inputs that will give the greater mean distances because such inputs will not win the competition with other neurons and will be not operative.

.

4.2 *Identification of the useful categories*

After the update of neural weights removing, to realize the "sterilized" algorithm from K.N.N, problem of neural network stressing, would be faced, consisting on trying to

2	0	1	2
0	3	3	3
1	3	3	2
2	3	2	1

7	0	1	2
0	4	2	2
1	3	2	1
2	3	1	1

12	0	1	2
0	3	6	8
1	7	7	7
2	8	8	8

51	0	1	2
0	2	4	4
1	3	4	4
2	4	4	4

Fig 3. Example of four instruments in which is evident the unequivocal belonging of instruments 2, 7,12 and 51 and the contemporary belonging of instrument

remove inputs concerning the greater mean distances calculated in front of mean input solicitation. In fact, we used 199 terms as input pattern, taking into account consideration of last paragraph. This extension can be perhaps reduced in accordance with the experience because in whole group of words some terms are informatively unnecessary and eventually will be replaced by others.

This problem is temporary set aside, concentrating our attention on categories declaration. Problems arise about the extraction of a lexical definition of a categories linked to neurons fire. In fact, procedure forecasts the isolation of instrumentation that is able to richly fire a single neuron. But it is evident that this limited set of instruments does not utilize all inputs of the K.N.N.. Such not excited inputs, give, however, a contribution to fire neuron that can be considered noise in output.

Now we take into consideration only the neuron indicated with indexes i=3 and j=3.

It is interesting only nine instruments fire neuron at its full value, as shown in Fig. 4.

AEMC-PQL

3	3	3
3	2	2
3	2	1

AEMC-3910

3	2	2
3	2	2
3	1	1

Ametek-ACE

4	2	2
3	2	1
3	1	1

Ametek-RIS4

3	2	2
3	2	2
4	2	1

AvPowerPA

3	2	2
3	2	2
3	2	1

LEM-Topas

4	2	2
2	1	1
2	1	1

Unipower8000

3	1	1
2	1	1
3	1	1

Unipower900F

4	1	1
4	1	1
5	2	1

Unipower812

3	1	1
3	1	1
4	2	1

Fig. 4: output layer about the nine instruments that fire a neuron i=3 and j=3 at a full value.

Examining the inputs of such instruments, it is possible to verify only ones involved. The free inputs give always a contribution to classification considering an Euclidean distance not properly significant that can be considered as an input noise.

So it is necessary to try to give a declaration of the category only by the involved input of such chosen instruments, as it is possible to extract from table of Fig. 5.

INSTRUMENT	Word1	Word2	Word3	Word4	.
AEMC-PQL	ac	rms	ieee	product	
AEMC-3910	case	meter	distorted	user	
AMETEK-ACE	lbs	specs	screen	word	
AMETEK-RIS4	specs	word	firmware	probe	
AvPowerPA2300	air	av	demon	limited	
LEM-Topas	option	electric	level	dc	
Unipower-8000	option	current	harmonic	standard	
Unipower-900F	pst	monitor	parallel	level	
Unipower-812	screen	word	pst	monitor	

Fig. 5: the terms more involved about each instrument, firing the neuron i=3 and j=3. The words are put in decreasing order respect at the Euclidean distance.

Clarifying the procedure considering the more near operative inputs, in Fig. 6 are reported the distances effectively remarked.

INSTRUMENT	Word 1	Word2	Word3	Word4	
AEMC-PQL	1.9E-5	1.8E-7	7.7E-6	1.9E-5	
AEMC-3910	2.4E-7	2.4E-7	3.4E-5	7.3E-5	
AMETEK-ACE	4.2E-8	1.5E-6	1.7E-6	5.3E-6	
AMETEK-RIS4	1.5E-6	5.3E-6	7.3E-6	3.5E-5	
AvPowerPA2300	1.6E-6	3.7E-5	3.7E-5	3.7E-5	
LEM-Topas	2.8E-8	5.7E-7	1.1E-6	1.8E-6	
Unipower-8000	7.6E-6	1.4E-5	3.8E-5	1.8E-4	
Unipower-900F	8.2E-5	1.0E-4	1.6E-4	2.4E-4	
Unipower-812	1.7E-6	5.2E-6	8.2E-5	1.0E-4	

Fig. 6: mean Euclidean distances calculated from the output "weights" on the plan of output neurons.

At least three ways can be considered to obtain the searched lexical category. The first is considering information can be extracted from Fig. 5 and Fig. 6, trying to construct a phrase with words selected. The second consists on isolating the instrument alone firing the neuron independently by the level of excited as shown in Fig 7.

Clearly, in such an eventuality, the contribution to the declaration of categories is linked to the firing level of neurons considered. The third procedure starts from the same Fig. 6, trying to individuate the entities stretched out on i and j axes. In this case we have the simplest, but the less classifying, capability. In fact if we consider the instrument n° 51 definition, characteristic of i =1 and j=1 neuron and the correlated definitions of the instruments n° 2, 8, 17, characteristics of i =3 and j=3 neuron, the channel number and the distortion individuation capability are the searched quantities.

channels

(51) 2	(1) 8	
(46) 9	(26) 5	(39) 5
(46) 10		(2,8,17) 1

distortion

Fig. 6: the instrument numbers (in the top of each cell) and their respectively levels of excited (in the bottom of the cell) are showed. The position of each cell corresponds to the localization of the neurons in the output grid. The direction of the axes indicate the increasing course of entities chosen to realize a double classification.

By such only examination it is no possible to correctly assign the axe to each entity. By the help of instrument n° 46 typical of i =3 and j=1 neuron, it is possible to affirm the i axe concerns channels while j axe concerns the distortion individuation capability.

4.3 *Use of the "sterilized" Kohonen neural network*

After the declaration categories determination, linked to the neurons firing, the "sterilized" K.N.N. can be utilized for any other input pattern outside the standard set, with only the step calculating the distance between the access pattern and each output node, resulting the standard settling to a precedent and definitive training. The devices that appear no classified are instead an indication of the instrumentation to put in the set of standard recycling whole the procedure. The whole of the values of the output layer, resulting by this singular operation, is an estimator of the instrumentation class.

5. PERSPECTIVES

After the individuation of critical parameters of the Kohonen Neural Network, related to a chosen "standard set", the instruments more properly involved on single neuron firing are used to give the less significance attribution to classification categories. A better classification can be obtained by words informatively involved in the process, but such a procedure, being more complex, is in progress. Also the better definition of input set of the neural network must be still faced.

REFERENCES

[1] G. Bucci, C. De Capua, C.Landi: *Power Measurements on high distorted signals : the experimental comparison between two alternative developed devices solutions*; Measurement, vol 23, n. 4, 1998

[2] R.C. Dugan, M.F. McGranaghan: *Electrical power system quality*; Mc Graw Hill, 1996

[3] G. Bucci, C.Landi: *On line digitally measurement for quality analysis of power systems under non sinusoidal conditions*; IEEE Trans on Instrumentation and Measurements, vol 48, n 6, Aug 1999

[4] J.Q. Zhang, Y. Shen: *Power measurement under non sinusoidal and noisy conditions*; Proc. of IMTC 2001, pp 1021-5 Budapest, May 2001

[5] http://www.funsci.com

[6] J. Taylor*: The historical background*; in Handbook of Neural Computing, IOP Publishing Ltd and Oxford University Press, 1997

[7] P.J. Werbos: *Why Neural Network*; in Handbook of Neural Computing, IOP Publishing Ltd and Oxford University Press, 1997

[8] P.J. Werbos: *A traditional roadmapping of artificial neural network:* in Handbook of Neural Computing, IOP Publishing Ltd and Oxford University Press, 1997

[9] J.L. Noyes: *Learning rules*; in Handbook of Neural Computing, IOP Publishing Ltd and Oxford University Press, 1997

[10] R. Beale, T. Jackon: *Neural Computing: An introduction*, Institute of Physics Publishing, Department of Computer Science, University of York, IOP Publishing Ltd 1990

Authors:

Maurizio Caciotta, Electronic Engineering Dept. of Roma Tre University; Via della Vasca Navale, 84, 00146 - Rome-Italy; Phone: +39-06-55177029; Fax: +39-06-5579078
e-mail: caciotta@misure.ele.uniroma3.it

Michela Cristiani, Electronic Engineering Dept. of Roma Tre University; Via della Vasca Navale, 84, 00146 - Rome-Italy; Phone: +39-06-55177085; Fax: +39-06-5579078
e-mail: mkcristiani@inwind.it