XVIII IMEKO WORLD CONGRESS
Metrology for a Sustainable Development
September, 17 – 22, 2006, Rio de Janeiro, Brazil

PARAMETRIC FAULT LOCATION OF ELECTRICAL CIRCUIT USING SUPPORT VECTOR MACHINE

S. Osowski [1,2], T. Markiewicz [1], R. Salat [3]

[1] Warsaw University of Technology, Warsaw, Poland, sto@iem.pw.edu.pl
[2] Military University of Technology,Warsaw, Poland, markiewt@iem.pw.edu.pl
[3] Agriculture University, Warsaw, Poland, salat@sggw.edu.pl

Abstract: The paper is concerned with the application of the Support Vector Machine to the discovering of the parametric fault in analog electrical circuits. The recognition of fault is based on the measurements of the accessible terminal voltage and current of the circuit at the set of frequencies. The SVM network fulfills the role of the recognizing system and of the classifier. The numerical results of recognition of faulty elements in the LC filter of the ladder network structure are presented and discussed in the paper.

Keywords: SVM classifiers, parametric fault detection, diagnosis of circuits.

1. INTRODUCTION

The paper presents the neural approach to the detection of the parametric fault and its location in analog electrical circuits on the basis of the external measurements of the voltages and currents. It is assumed that the general structure of the circuit under investigation is known. The fault of element is understood as the parametric change of its value beyond the assumed tolerance limit [1,2,6]. Only fault of one element at a time is considered. The detection of the faulty element is performed on the basis of the measurements of the terminal signals, i.e., the terminal voltages and currents of the circuit, operating in the normal and faulty conditions. The measured signals should undergo the stage of preprocessing to extract the diagnostic features of the obtained information. These features form the input signals to the neural network, performing the role of classifier.

The neural network of the Support Vector Machine (SVM) type is trained on the examples of the patterns belonging to the normal operation of the circuit and to the representatives of the typical faults of the particular element. The important advantage of the neural method is its ability to treat efficiently the parametric changes of the values of the faulty elements. The method is effective and offers a great speed and acceptable accuracy for the fault detection in the large range of parameter changes. In comparison to the standard neural networks, like the multilayer perceptron or self-organizing Kohonen network [1,2,8] SVM offers much better generalization ability at limited number of training samples. The results concerning the recognition of single element faults in the resistively terminated LC prototype ladder filter are presented and discussed. They confirm great efficiency of the developed diagnostic system.

2. PRINCIPLE OF FAULT RECOGNITION AND LOCATION

In the recognition and localization of fault in analog electrical circuit it is assumed, that the general structure of the circuit under investigation is known [8]. The user is able to measure the external voltages and currents of the circuit, operating in the normal and faulty conditions in the steady state under sinusoidal excitation. We assume the fault of element as the parametric change of its value above the assumed tolerance limit.

The basic observation is that each state of the circuit, either normal or any single fault, is associated with the specific frequency characteristics of the magnitude and phase of the measured variables. These characteristics differ to some degree at various faults. The differences between the nominal and faulty states are used by us to make the recognition of the particular state of the circuit. This will be done by the neural classifier network. The classifier is trained on the data representing different examples of the nominal and faulty states of the circuit. After training, the parameters of the classifier are frozen and the system is ready for the on-line operation of the diagnostic task.

Let us assume that there is a sufficient number of independent signal measurements in the circuit, greater than the number of elements in the analog circuit of known topology and nominal values of its elements [8,11]. The measurements are concerned with the external accessible points at different frequencies. These frequencies should be chosen in a way to enhance the differences between different states of the circuit. Hence special procedure of frequency selection should be applied. The important point is to provide the highest sensitivity of the system to any changes in the parameters of the circuit element. Hence the natural way to determine the optimal set of frequencies is the application of the sensitivity analysis of the circuit [1]. The sensitivity curves of the magnitude and phase of the external measured variables with respect to the circuit elements are generated. The frequencies corresponding to the maxima of the absolute values of these curves are the candidates for testing frequencies.

The next step is to convert the measured variables into the diagnostic features. They should be normalized and at the same time as sensitive as possible to the changes of the parameters of the elements. To differentiate the feature values corresponding to individual faults we consider here the relative differences between the faulty and non-faulty modes of the circuit operation. Applying the general notation x for either measured voltage V or current I (magnitude or phase) we define the feature as the relative difference of this variable at nominal and actual (presumably faulty) state of the circuit at the frequency f_i,

$$x_r(f_i) = \frac{x(f_i) - x_n(f_i)}{x_n(f_i)} \tag{1}$$

The variable $x_n(f_i)$ means the measured quantity under non-faulty (nominal) operation of the circuit at the frequency f_i. The generated features $x_r(f_i)$ are the candidates for the input vector **x** defining the input signals for the classifier.

The next step is the validation and selection of the candidate set of features. Let us assume that the measurements have been done for one voltage and current of the circuit at the set of frequencies, for which the diagnostic features have been generated for the magnitude and phase according to eq. (1). As a result we get four possible candidate vectors: $\mathbf{V}_m = abs(\mathbf{V}_r)$, $\mathbf{V}_p = \arg(\mathbf{V}_r)$, $\mathbf{I}_m = abs(\mathbf{I}_r)$ and $\mathbf{I}_p = \arg(\mathbf{I}_r)$, where the vector of voltages $\mathbf{V}_r = [V_r(f_1), V_r(f_2), ..., V_r(f_{m1})]$ and currents $\mathbf{I}_r = [I_r(f_1), I_r(f_2), ..., I_r(f_{m2})]$ represent relative voltages and currents of the terminals at different frequencies f_i, normalized according to the relation (1), *abs* stands for the magnitude and *arg* for the phase of the corresponding complex values. Therefore the maximum size candidate feature vector **x** that may be used in learning, is given by $\mathbf{x} = [\mathbf{V}_m, \mathbf{V}_p, \mathbf{I}_m, \mathbf{I}_p]$. However different candidate vectors formed as the combinations od $\mathbf{V}_m$, $\mathbf{I}_m$, $\mathbf{V}_p$ and $\mathbf{I}_p$ may be also considered, for example $\mathbf{x} = [\mathbf{V}_m, \mathbf{V}_p, \mathbf{I}_m]$, $\mathbf{x} = [\mathbf{V}_m, \mathbf{V}_p, \mathbf{I}_p)]$, $\mathbf{x} = [\mathbf{V}_m, \mathbf{I}_m]$ or even in extreme cases $\mathbf{x} = [\mathbf{V}_m]$, $\mathbf{x} = [\mathbf{I}_m]$, etc.

Many different feature assessment methods are known and applied in practice [3]. To the most popular belong principal component analysis (PCA), correlation existing among features, correlation between the features and the classes, statistical analysis of mean and variance of the features or even application of SVM feature ranking. In this work we have applied the PCA based assessment of the features quality.

The PCA [4] is described as the linear transformation **y**=**Wx**, mapping the N-dimensional original vector **x** into L-dimensional output vector **y**, where $L<N$. The vector **y** preserves the most important elements of the original information. The transformation matrix **W** is composed of the eigenvectors associated with L largest eigenvalues of the correlation matrix $\mathbf{R}_{\mathbf{xx}}$ defined for the set of input vectors $\mathbf{x}_i$. Taking the value of L equal two or at most three we can map the original N-dimensional vectors $\mathbf{x}_i$ into the two or three–dimensional PCA space, that can be easily represented in a two- or three-dimensional coordinate system. Thanks to this the visual inspection of the trajectories of points representing the faults of elements can be performed. Good feature set corresponds to the trajectories of different faults separated from each other as well as possible. The set of features providing the best separation of trajectories of different faults should be considered as the potential input vector **x** applied to the neural classifier performing the recognition of the patterns associated with either fault of the element or with the nominal state of the circuit. The proposed diagnostic system structure is presented in Fig. 1.

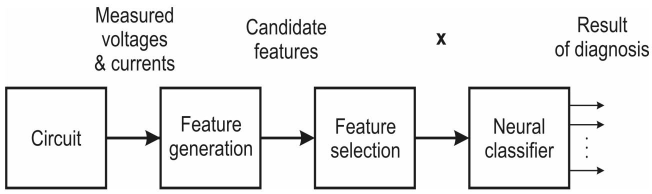

Fig. 1 The general scheme of the circuit diagnostic system

3. THE NEURAL CLASSIFIER

The important role in our diagnostic system fulfills the neural classifier, performing the final recognition of the circuit state on the basis of the applied input feature vector **x**. In our solution we have applied the Support Vector Machine (SVM) classifier, regarded now as the most effective classification tool [10,12]. The distinct advantage of the SVM network solution is its good generalization ability. Trained on the limited number of representative examples of each fault, the network is able to recognize the non-ideal (parametric) fault in the wide range of the faulty parameter values associated with the assumed tolerance of the non-faulty elements.

SVM is a linear system working in the highly dimensional feature space formed by the nonlinear mapping of the N-dimensional input vector **x** into a K-dimensional feature space ($K>N$) through the use of a mapping function $\varphi(\mathbf{x})$. The SVM network recognizes between two classes, represented by d_i=1 and d_i=-1. In the classification mode the equation of the separating hyperplane is given by the relation $y(\mathbf{x}) = \mathbf{w}^T \boldsymbol{\varphi}(\mathbf{x}) + b = \sum_{j=1}^{K} w_j \varphi_j(\mathbf{x}) + b = 0$, where $\boldsymbol{\varphi}(\mathbf{x}) = [\varphi_1(\mathbf{x}), ..., \varphi_K(\mathbf{x})]^T$ is a vector of mapping functions of hidden units and $\mathbf{w} = [w_1, ..., w_K]^T$ is the weight vector. The parameters of the hyperplane $y(\mathbf{x})$ are adjusted in a way to maximize the distance between the closest representatives of both classes. The primary learning problem [10,12] is formulated as the minimization of the objective function $\phi(\mathbf{w},\xi)$

$$\phi(\mathbf{w},\xi) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{p}\xi_i \qquad (2)$$

at the linear constraints defined for each learning data sample (i = 1, 2, ..., p) with ξ_i - the slack variable

$$\begin{aligned} &d_i\left(\mathbf{w}^T\varphi(\mathbf{x}) + b\right) \geq 1 - \xi_i \\ &\xi_i \geq 0 \end{aligned} \qquad (3)$$

The first term in equation (2) corresponds to the maximization of the margin of separation. The constant C is the regularization parameter responsible for the minimization of the learning errors. The higher is its value the bigger is the impact of this term on the final parameters of the hyperplane.

The most distinctive fact about SVM is that the learning task is reduced to the quadratic programming by introducing the so-called Lagrange multipliers α_i. All operations in learning and testing modes are done in SVM by using kernel functions satisfying Mercer conditions [12]. The kernel is defined as

$$K(\mathbf{x},\mathbf{x}_i) = \boldsymbol{\varphi}^T(\mathbf{x}_i)\boldsymbol{\varphi}(\mathbf{x}) \qquad (4)$$

The most often used kernels include radial Gaussian, polynomial, spline or linear functions [10]. The final problem of learning SVM, formulated as the task of separating learning vectors $\mathbf{x}_i$ into two classes of the destination values, either $d_i=1$ or $d_i=-1$, with maximal separation margin, is reduced to the dual maximization problem of the quadratic function [10]

$$\max \quad Q(\boldsymbol{\alpha}) = \sum_{i=1}^{p}\alpha_i - \frac{1}{2}\sum_{i=1}^{p}\sum_{j=1}^{p}\alpha_i\alpha_j d_i d_j K\left(\mathbf{x}_i,\mathbf{x}_j\right) \qquad (5)$$

with the linear constraints $\sum_{i=1}^{p}\alpha_i d_i = 0, \; 0 \leq \alpha_i \leq C$. The regularizing parameter C determines the balance between the complexity of the network, characterized by the weight vector $\mathbf{w}$ and the error of classification of data. For the normalized input signals the value of C is usually much higher than 1 and is adjusted by the cross validation procedure. The solution of (5) is expressed througth the Lagrange multipliers, on the basis of which the optimal weight vector $\mathbf{w}_{opt}$ is determined

$$\mathbf{w}_{opt} = \sum_{i=1}^{N_{sv}}\alpha_i d_i \boldsymbol{\varphi}(\mathbf{x}_i) \qquad (6)$$

In this equation N_{sv} means the number of support vectors, i.e. the learning vectors $\mathbf{x}_i$, for which the Lagrange multipliers are nonzero. The output signal y($\mathbf{x}$) of the SVM network is determined now as the function of kernels

$$y(\mathbf{x}) = \sum_{i=1}^{N_{sv}}\alpha_i d_i K(\mathbf{x}_i,\mathbf{x}) + b \qquad (7)$$

The signal $y(\mathbf{x})$ greater than 0 is associated with class 1 and the negative with the opposite one. Although SVM separates the data into two classes only, the recognition of more classes is straightforward by applying either one against one or one against all methods [5]. The more powerful is one against one approach in which many SVM networks are trained to recognize between all combinations of two classes of data. At M classes we have to train $M(M-1)/2$ individual SVM networks. In the retrieval mode the vector $\mathbf{x}$ belongs to the class of the highest number of winnings in all combinations of classes.

The important point in designing SVM classifier is the choice of the kernel function. The simplest linear kernel is usually inefficient due to the lack of linear separability of the data. The polynomial kernel may be also useless, if high degree of polynomial is needed, since in such case the system is becoming badly conditioned. The best results are usually obtained at application of Gaussian kernel and this kernel has been applied in all further experiments.

On the stage of designing the SVM classifier system the choice of the Gaussian spread σ and the regularization constant C is very essential. Especially important is the value of C, since it controls the tradeoff between the complexity of the machine and the number of non-separable data points used in learning. The small value of C results in the acceptation of more not separated learning points. At higher value of C we get the lower number of classification errors of the learning data points, but more complex network structure. The optimal value was determined for each pair of classes independently after additional series of learning experiments through the use of the validation test sets. The process of optimizing the values of C and σ was done together. Many different values of C and σ combined together have been used in the learning process and their optimal values are those for which the classification error on the validation data set was the smallest one. The SVM networks were trained using Platt algorithm [7].

To assess the performance of our diagnostic system properly we have compared it with the application of MultiLayer Perceptron (MLP) used as the classifier. MLP is the most known and typical multilayer neural network solution [4] applying the sigmoidal activation function. It performs the classification of data in one simple structure (no need for many classifiers, as in SVM case). The results of numerical experiments performed for the same data sets have confirmed the superiority of the proposed SVM solution.

4. THE CIRCUIT UNDER TEST

The theoretical considerations presented in the previous sections will be illustrated on the example of the prototype of the resistively terminated passive LC ladder filter of 9th order presented in Fig. 2

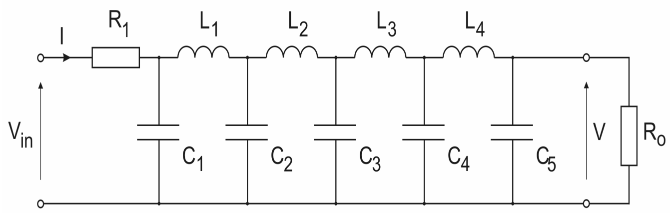

Fig. 2 The RLC ladder filter structure

It is the ninth order circuit containing 11 elements. The elements of the filter have been adjusted to realize the low-pass characteristics in the normalized range of frequencies. The normalized values of the elements used in experiments were as follows: $R_o=R_1=1$, $C_1=0.2$, $C_2=0.72$, $C_3=1.46$, $C_4=0.691$, $C_5=0.290$, $L_1=0.408$, $L_2=0.509$, $L_3=0.730$, $L_4=0.340$. Only the terminal points of the circuit are accessible for measurements.

The diagnostic task is to find out if the element value is different from its nominal one by more than the assumed tolerance limit (10%). Such case is regarded as a fault. We consider the single faults of elements (only one faulty element at the same time). It means that any kind of fault of the element is associated now with one class. In such case we have 12 types of circuit operations. One class represents the normal operation and 11 classes are associated with the fault of any of its eleven circuit elements.

Two terminal measurements are available in the network. At sinusoidal voltage excitation and resistive load (R_o) at the output terminal, the input current *I* and output voltage *V* at different frequencies can be determined. On the basis of these measured values we will generate the candidate features that may form the input vector **x** for the SVM network, using equation (1). We may rely here on the magnitude and phase frequency characteristics of these two measured variables. The frequency values used in the analysis of the circuit should be chosen first. They have been determined by the sensitivity analysis of the original circuit. The frequency set is composed of all frequencies for which we have observed the extremes of either magnitude or phase sensitivity characteristics of the output voltage and input current. Fig. 3 presents the exemplary four sensitivity curves for the magnitude and phase of the output voltage frequency characteristics obtained by using NAP program [9] with respect to two capacitances (C_3 and C_5) and two inductances (L_1 and L_3). The points corresponding to the extreme values of the sensitivity functions are chosen as the frequencies for the analysis of the filter. They were used for the generation of the learning and testing data.

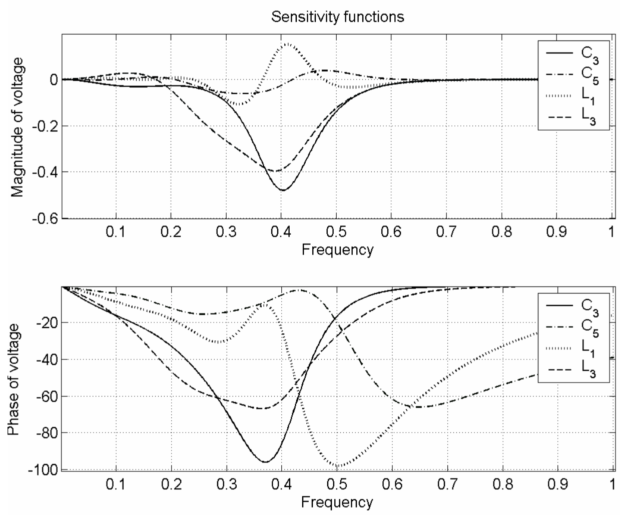

Fig. 3 The sensitivity of the magnitude and phase of the output voltage of the tested circuit

After performing the sensitivity analysis we have found 21 and 20 different frequency points corresponding to the extremes of the magnitude and phase of the output voltage, respectively, as well as 24 and 25 frequency points corresponding to the extremes of the magnitude and phase characteristics of the input current. At such number of frequency points the vector of magnitude voltage characteristics $\mathbf{V}_m$ is composed of 21 elements, the phase voltage characteristics $\mathbf{V}_p$ – 20 elements, the magnitude of the input current vector $\mathbf{I}_m$ – 24 elements and the phase of the input current vector $\mathbf{I}_p$ – 25 elements. Hence the maximal dimension of the feature vector **x** is 90.

We have considered here the parametric faults of the circuit elements. As the faults we understand all changes of the nominal values of resistances, inductances and capacitances beyond the assumed tolerance limit (10% in experiments). Each fault has been associated with the tolerance of the remaining non-faulty elements, changing randomly in the experiments from 0 to 5%.

5. THE RESULTS OF NUMERICAL EXPERIMENTS

The first important question is of the optimal feature representation of the measured data. All combinations of the magnitude and phase information contained in the normalized output voltage and input current (vectors $\mathbf{V}_m$, $\mathbf{V}_p$, $\mathbf{I}_m$, $\mathbf{I}_p$) may form the feature vector **x**. The same number of each case (either appropriate fault or normal operation) was used in experiments. The principal component analysis of the normalized data at different arrangement of the feature vectors has been performed. After assessing all results by visual inspection we have come to the conclusion that the full length vector **x** is not the best one since the combination of the magnitudes of the output voltage and input current, $\mathbf{x}=[\mathbf{V}_m\ \mathbf{I}_m]$ has provided a bit better separation of different fault trajectories.

Fig. 4 presents the PCA representation of the learning data corresponding to the best arrangement (the reduced dimension vector composed only of the magnitude information $\mathbf{x}=[\mathbf{V}_m\ \mathbf{I}_m]$).

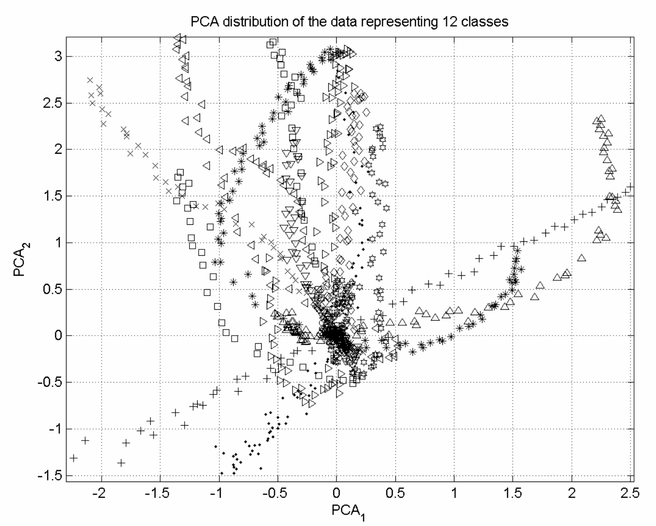

Fig. 4 The PCA representation of the measured data for the best feature vector $\mathbf{x}=[\mathbf{V}_m\ \mathbf{I}_m]$

Even in this figure the unique interpretation of the results is not easy since the distribution of the points belonging to 12 classes (11 faults of element plus the normal operation of the circuit) is extremely complex, especially very close to the center (normal operation of the circuit).

The next experiments have been directed to check the efficiency of different representations of the feature vector **x** for the fault recognition by the SVM network. We have used different possible combinations of magnitude and phase information to form feature vector **x**, beginning from the single individual vectors and ending on all of them combined together (**x**=[$\mathbf{V}_m$ $\mathbf{V}_p$ $\mathbf{I}_m$ $\mathbf{I}_p$]). The numerical experiments of classification by using Gaussian kernel SVM working in one against one mode, have been performed for the data samples evenly split for learning and testing sets (420 samples for each faulty element and for the normal state of the circuit). All different sets of features forming the vector **x** have been tried in experiments. The testing data have been applied only in the testing mode of the trained SVM network. Here we will show the results corresponding to 5% tolerance of the non-faulty elements. To deal with 12 classes we have applied one against one strategy [4], resulting in learning 66 individual SVM classifiers, recognizing two classes at a time.

Table 1 presents the cumulative comparison of the average testing misclassification rate in percentage, corresponding to different representations of the feature vector **x**. The results are given for the normal operation of the circuit (all element values within the tolerance limit) and for the faulty mode of operation. The faults of elements have been distributed in the wide considered region, for which the particular element value differs more than allowed by the tolerance (10%). The faulty elements changed from 0.05 to 0.9 and from 1.1 to 10 of their nominal values. In the case of faulty mode operation the mean values of the misclassification rate within all 11 classes have been calculated and presented in the table.

Table 1 Summary of the misclassification rate of the testing data at different representations of the feature vector x

Representation of feature vector	Total misclassification rate		
	Normal operation	Faulty mode	Mean
x=[$\mathbf{V}_m$ $\mathbf{I}_m$ $\mathbf{V}_p$ $\mathbf{I}_p$]	3.1%	0.41%	0.63%
x=[$\mathbf{V}_m$ $\mathbf{I}_m$ $\mathbf{V}_p$]	5.0%	0.41%	0.69%
x=[$\mathbf{V}_m$ $\mathbf{I}_m$ $\mathbf{I}_p$]	0.72%	0.58%	0.60%
x=[$\mathbf{V}_m$ $\mathbf{V}_p$ $\mathbf{I}_p$]	4.7%	0.39%	0.65%
x=[$\mathbf{I}_m$ $\mathbf{V}_p$ $\mathbf{I}_p$]	5.0%	0.45%	0.73%
x=[$\mathbf{V}_m$ $\mathbf{I}_m$]	**0.24%**	**0.54%**	**0.52%**
x=[$\mathbf{V}_m$ $\mathbf{I}_p$]	1.7%	0.76%	0.81%
x=[$\mathbf{V}_m$ $\mathbf{V}_p$]	5%	0.58%	0.85%
x=[$\mathbf{I}_m$ $\mathbf{I}_p$]	2.67%	1.19%	1.28%
x=[$\mathbf{I}_m$ $\mathbf{V}_p$]	4.7%	0.39%	0.66%
x=[$\mathbf{I}_p$ $\mathbf{V}_p$]	2.9%	0.87%	1.04%
x=$\mathbf{V}_m$	6.7%	1.71%	2.01%
x=$\mathbf{V}_p$	4.67%	1.04%	1.26%
x=$\mathbf{I}_m$	3%	1.34%	1.44%
x=$\mathbf{I}_p$	1.33%	1.60%	1.59%

The mean error (last column of the table) is computed as the ratio of the total number of misclassifications to the number of samples used in the experiments (at 11 faults and equal representation of classes the faults have predominant impact on the mean results).

The summary results presented in Table 1 point out that many different representations, providing similar level of misclassification ratio are possible. If we take into account the mean value of the total error, the best seems to be the reduced representation of the feature vector **x**=[$\mathbf{V}_m$ $\mathbf{I}_m$]. It is interesting that some partial representations of the features, for example **x**=[$\mathbf{V}_m$ $\mathbf{I}_m$ $\mathbf{I}_p$] or **x**=[$\mathbf{V}_m$ $\mathbf{V}_p$ $\mathbf{I}_p$] are also very good for recognition of the faults of elements and even better than full set of features.

Table 2 presents the detailed results of faulty element recognition at testing mode of the trained SVM network using testing set (not taking part in learning) for the best selected representation of the features, **x**=[$\mathbf{V}_m$ $\mathbf{I}_m$]. The first column shows which element is faulty and the second one – the average misclassification rate (in percentage) corresponding to the particular fault.

Table 2 The results of testing the SVM classifier on the data samples corresponding to different faults at the feature vector x=[$\mathbf{v}_m$ $\mathbf{i}_m$]

Faulty element	Misclassification rate
No fault	0.24%
R_1	0
C_1	0.71%
C_2	0
C_3	0.24%
C_4	0.71%
C_5	2.86%
L_1	0
L_2	0
L_3	0
L_4	0.95%
R_o	0.48%
Mean	0.52%

As it is seen the overall accuracy of the fault recognition at the uniform distribution of data in the whole region of the considered fault is satisfactory from the practical point of view. In the worst case (the fault of element C_5) the average misclassification rate was 2.86%. The misclassification ratio of the single fault of most elements has been reduced to the insignificant or even zero values. The total average error of the fault recognition, calculated as the ordinary mean of all errors, has been reduced to the value of 0.52%.

The detailed observation of error distribution has revealed that most errors were committed for the data placed very close to the border of the normal operation (the element values variation slightly below or above the tolerance limit of 5% and at the border of the faulty state (slight crossing of the assumed 10% tolerance limit regarded as a fault). To check in details how the classifier system is able to deal with this kind of data we have generated additional set of data corresponding to the nominal element values placed close to the tolerance limit ($5\% \pm 2\%$) and the faults placed on the border of the assumed tolerance $10\% \pm 2\%$). Once again the non-faulty elements have been disturbed randomly within the tolerance limit $\pm 3\%$. In this experiment half of the data generated in this way has been added to the already existing learning set and the other half was left for testing only. So in the experiment the testing data set contained only these most difficult cases. The testing has been performed for the SVM network retrained using the extended learning data set containing also the additional data representing the region very close to the tolerance limit. The performed test was extremely difficult since the testing set contained only the data most difficult for the recognition. However, even in this very demanding test the results are fairly acceptable. For most elements the fault recognition error was below 6%, although the recognition of some faults has been done with large error reaching in worst case even 40% (the capacitor C_5). The average misclassification rate calculated as the simple mean of all average errors was equal in this test 14.5%.

To check the impact of frequencies on the accuracy of the diagnostic system we have made additional experiments by applying the random choice of frequency values in the analysis of the filter. The number of frequency points of the analysis of the circuit was changing from 10 to 23. Their distribution was also changing. The SVM system has been trained and then tested using the same number of testing data. The total misclassification ratio on the testing data was this time higher and dependent on the number and distribution of the frequency analysis points. At 10 frequencies and their random distribution within the normalized range from 0 to 2 the average classification error was close to 1.5% (the actual values depend on the actually generated random values of elements). At 23 frequency points distributed randomly in the same range the average classification error has dropped to the value of 0.78%. The results of this test confirm the importance of the sensitivity analysis applied for choosing the number of frequency points used in the preparation of the data and the selection of their particular values.

In the last experiments we have compared the accuracy of our SVM based diagnostic system with the one applying the MLP classifier. The optimal MLP network containing 90 inputs, 20 hidden and 12 output sigmoidal neurons was learned using the same data set as in the previous experiments and then tested on the testing data. The obtained results are inferior in comparison to the SVM. The mean classification error for the testing data was equal 5.43%, which is much higher than 0.52% obtained by the SVM classifier. Besides this the training of the MLP classifier was extremely time consuming in comparison to SVM one. At the same number of data (420 samples of each of 12 classes) and application of LM algorithm it lasted at least 50 times longer.

6. CONCLUSION

The paper has presented the new approach to the fault detection and location in an analog circuit, based on the application of Support Vector Machine. The most difficult case of the parametric faults of circuit elements has been considered. The numerical experiments conducted for the 9^{th} order RLC ladder filter have confirmed that the developed diagnostic system works well and is able to locate the single faults with the acceptable accuracy for the whole range of parameter values from the tolerance limit to the short circuit or open circuit of the element. The important feature of the proposed solution is its high efficiency and great speed of operation. The training of the SVM network system, performed using Platt algorithm [7] was very quick. At 5040 data pairs and 12 classes it lasted no longer than 3 minutes on the PC computer of 2.4GHz and 512M RAM. Moreover, once the network was trained, the recognition of fault was achieved immediately, irrespective of the size of the circuit. Thus the solution is suited for the real time applications for fault detection and location in any linear circuit.

The distinct advantage of the SVM network solution over the standard MLP case is its good generalization ability. Trained on the limited number of representative examples of each fault, the network is able to recognize the non-ideal (parametric) fault in the wide range of changed parameter values and at some assumed tolerance of the non-faulty elements.

REFERENCES

[1] C. Alippi, M. Catelani, A. Fort, Automatic selection of test frequencies for the diagnosis of faults in analog circuits, Proc. IMTC., Anchorage, 2002, pp. 1503-1507

[2] A. Fanni, M. Giua, A. Marchesi, A. Montisci, A neural network diagnosis approach for analog circuits, Journal of Applied Intelligence, 1999, vol. 11, pp. 1-20

[3] I. Guyon, A. Elisseeff, An introduction to variable and feature selection. J. Machine Learn., 2003, 1158–1182

[4] S. Haykin, Neural networks, comprehensive foundation, Prentice Hall, 1999, New Jersey

[5] W. Hsu and C. J. Lin. A comparison of methods for multi-class support vector machines , IEEE Trans. NN, 13, 2002, pp. 415-425

[6] R. Liu, Testing and diagnosis of analog circuits and systems, Van Nostrand Reinhold, New York, 1991

[7] L. Platt, Fast training of SVM using sequential optimization, (in B. Scholkopf, et.al. Eds., Advances in kernel methods. MIT Press), 1998, pp. 185-208

[8] R. Sałat, S. Osowski, Analog filter diagnosis using SVM, Proc. ECCTD, Krakow, 2003, pp. 421-424

[9] T. Rubner-Petersen, NAP2 – a nonlinear analysis program for electronic circuits, University of Denmark, 1981

[10] B. Schölkopf, A. Smola, Learning with Kernels, MIT Press. 2002

[11] J. Starzyk, M. El-Gamal, An optimization approach to fault location in analog circuits, Proc. of ICAS, Montreal, 1984, pp. 456-459

[12] V. Vapnik, Statistical learning theory, Wiley, N.Y. 1998

[13] Matlab manual, MathWorks, Natick, 2001