Jarosław KUREK[3], Stanisław OSOWSKI[1,2]

Politechnika Warszawska (1), Wojskowa Akademia Techniczna (2), Warsaw University of Life Sciences (3)

Diagnostic feature selection for efficient recognition of different faults of rotor bars in the induction machine

Streszczenie. *Praca przedstawia dwuetapową skuteczną metodę selekcji cech diagnostycznych dla potrzeb lokalizacji uszkodzonych prętów w maszynie indukcyjnej. Badania diagnostyczne prętów maszyny odbywają się na podstawie zarejestrowanych przebiegów czasowych prądu i napięcia stojana oraz strumienia wałowego rozproszenia. Wielkości te poddane są transformacji dyskretnej Fouriera (FFT) a uzyskane spektrum częstotliwościowe poddane jest analizie ukierunkowanej na wyłowienie cech diagnostycznych (harmonicznych) najbardziej różnicujących różne klasy uszkodzeń prętów maszyny. Zaproponowano dwustopniowy algorytm selekcji cech stosujący wielokrokową selekcje regresyjną eliminującą cechy skorelowane ze sobą i pozostawiającej w zbiorze jedynie cechy najlepiej skorelowane z rozpoznawaną klasą. Ten system selekcji cech został sprawdzony w praktycznym rozwiązaniu komputerowym on-line wykorzystującym jako klasyfikator sieć typu Support Vector Machine (SVM) i przetestowany na silniku indukcyjnym o zmodyfikowanej konstrukcji umożliwiającej symulację fizyczną uszkodzeń różnych prętów maszyny.(**Selekcja cech diagnostycznych w zastosowaniu do rozpoznania różnych uszkodzeń prętów maszyny indukcyjnej**).*

Abstract. *The paper presents two-step selection method of the most important diagnostic features in application to the localization of the faulty bars of the squirrel cage induction machine. The registered waveforms of the stator current, voltage and shaft flux are first transformed to the frequency spectrum using FFT. These harmonics are the potential candidates to be the features on the basis of which the SVM classifier is able to localize the faulty bars. The selection of the optimal set of harmonics is done in two-step approach. In the first step the wide set of harmonics is identified on the basis of comparison their values for two classes of data. In the second step the forward and backward regressive selection is applied to eliminate the strongly correlated features and to leave only the features well correlated with the recognized classes. The results of numerical experiments are presented and discussed in the paper.*

Słowa kluczowe: Selekcja cech diagnostycznych, diagnostyka prętów klatki, maszyna indukcyjna, SVM.
Keywords: Selection of the diagnostic features, fault localization of the faulty bars, induction machine, Support Vector Machine.

Introduction

The problem of automatic localization of the faulty bars in squirrel cage induction motor belongs to the task of pattern recognition [2,3]. On the basis of the measurements of the instantaneous values of the current, voltage or shaft flux the diagnostic features are generated, on the basis of which the chosen classifier will recognize the fault. To get the best results of the recognition we have to determine the optimal set of these features. This paper will present the two-stage selection method of the features.

The registered waveforms of the stator current, voltage and shaft flux are first transformed to the frequency spectrum using FFT. These harmonics are the potential candidates to be the features on the basis of which the SVM classifier is able to localize the faulty bars. The selection of the optimal set of harmonics is done in two-steps. In the first step the wide set of harmonics is identified on the basis of comparing their values for two classes of data. In the second step the forward and backward regressive selection is applied to eliminate the strongly correlated features and to leave only the features well correlated with the recognized classes. The results of numerical experiments concerning recognition of 10 types of bar faults will be presented and discussed in the paper

Problem statement

The aim of our research is to develop the automatic diagnostic system able to discover not only the case of occurrence of the fault of the rotor bars, but also to recognize the type of the fault. In our considerations we have focused on several essential cases of the bar damage most often met in practice. Each type of the bar fault will be called the class. In our experiments we have dealt with the induction machine of 33 bars in the squirrel cage. For this machine the considered cases included [3]:

class 1 – no fault
class 2 – one bar broken
class 3 – two subsequent bars broken
class 4 – three subsequent bars broken
class 5 – the 1st, 2nd, 3rd and 16th bars broken
class 6 – the 1st, 2nd, 3rd and 16th, 17th bars broken
class 7 – the 1st, 2nd, 3rd and 16th , 17th, 18th bars broken
class 8 – the 1st, 2nd and 16th bars broken
class 9 – the 1st, 2nd, 16th and 17th bars broken
class 10 – the 1st and 16th bars broken.

To apply the automatic system of fault recognition we have to gather the learning data set covering different cases of all considered faults. To generate such data we have used the constructively modified induction machine Sg132M-6B-S of 5.5kW [4]. The construction change has enabled us to simulate the broken bars in the induction motor in a physical way. In the experiments we have registered the instantaneous values of the following variables: the phase current, the phase voltage and shaft flux. The number of registrations was equal 36 for every cases so it means together 360 measurements. All registrations have been done by using the data acquisition card USB-6251 of the sampling frequency equal 10kHz.

The registered variables were normalized and established the basis for the diagnosis of the squirrel cage bars. The output of the system should decide if the cage of the motor is in the normal state or is damaged. In the case of damage the system should qualify it to one of 9 classes of damages mentioned above.

Feature selection

To solve this classification problem we have to generate the diagnostic features, on the basis of which the SVM neural network will be able to recognize the proper class. The excessive number of features, some of which may represent the noise, is harmful for the recognition accuracy of the patterns [1,5]. Thus the important problem in classification and machine learning is to decide on the efficient methods of selection of features according to their importance for the problem solution. The elimination of not significant features leads to the reduction of the dimensionality of the feature space and improvement of performance of the classifier in the testing mode on the data not taking part in learning.

Our features will rely on the Fourier representation of the measured data. Fig. 1 presents the current spectra corresponding to different states of the bars: the non-faulty state (a), one bar broken (b), 3 subsequent bars broken (c) and the case of 3 subsequent bars placed on 2 opposite sides of the rotor (6 bars broken). From the point of view of the diagnosis the extremely important is to discover which frequencies carry the most important diagnostic information, that will be used as the input signals for the SVM classifier.

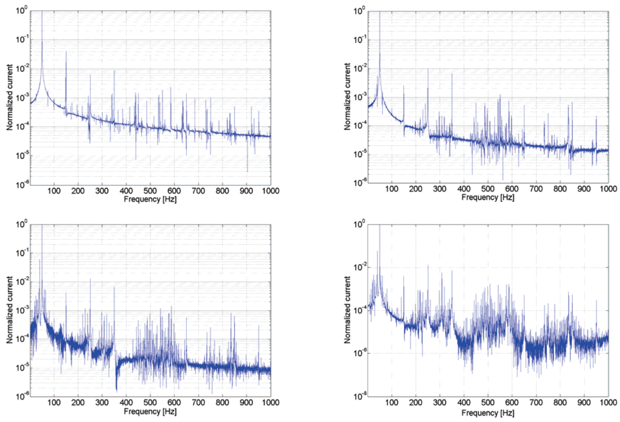

Fig. 1 The frequency spectra of the stator current of the induction machine at different state of the bars: a) non-faulty bars, b) one bar broken, c) fault of three neighboring bars, 4) fault of 6 bars

Step 1 - differential selection

The first step in designing the automatic system, able to recognize different faults is to find, which of the harmonics differentiate the faulty classes in the most unambiguous way. It is evident that visual approach to this problem is inappropriate and we have to develop some automatic method to discover the harmonics differing classes in the most distinctive way.

The feature suitable for the recognition between two classes should be characterized by high value of the summed differences over all considered observed cases. Small value of this summed difference means high diversity of these differences for the subsequent cases and lack of correlation of the feature with a class. Such analysis was performed for each feature at all combinations of classes.

Let us denote by **x** and **y** the vectors of harmonic distribution of the measured variables, (independently for the current, voltage and flux) belonging to two different classes considered as the potential features. Vector **x** represents one class and **y** the opposite one. We denote by *n* and *m* the number of samples of the vectors **x** and **y**, respectively. The resultant ranking measure $d_1(f)$ is calculated for each harmonic of the frequency *f* using the expression

(1) $$d_1(f) = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i(f) - y_j(f))$$

As a result we get the set of values of $d_1(f_i)$ characterizing the importance of each frequency f_i for two considered classes. High value of $d_1(f_i)$ means high influence of this particular frequency on the recognition between classes represented by **x** and **y**. On the basis of the value of $d_1(f_i)$ for the particular pair of classes we can sort the frequencies from the most important (the highest value of $d_1(f)$ measure) to the least important (the smallest one). Such sets of frequencies are determined independently for each pair of classes.

Moreover we have tried also its modified forms $d_2(f)$ defined on the basis of the sign of differences

(2) $$d_2(f) = \sum_{i=1}^{n}\sum_{j=1}^{m} sign(x_i(f) - y_j(f))$$

In this modification (eq. 2) we rely on the signs of the differences between the samples belonging to two different classes for the same frequency *f.* In the case when in all samples representing the classes we observe the same tendency (the feature of one class is always higher than in the second class) this feature is valuable and gets higher position in ranking. In the opposite case, when the samples of both classes are of different tendencies, the signs of differences are changing and compensate in (2). In such case the discriminative measure $d_2(f)$ will assume small value.

As a result of such analysis we get the set of harmonic frequencies arranged in the decreasing order of their discrimination abilities, independently for the current, flux and voltage for all considered pairs of classes. Among them we search for the best harmonics present simultaneously in the sets corresponding to all combinations of two classes. However among them there are some harmonics strongly correlated with each other. The correlated features may dominate over the set of other features and as a result reduce the potential recognition ability of the whole set. To reduce this effect we have performed the correlation analysis of the features selected in the first stage. This step eliminates the features correlated with each other and leaves only the features of the highest correlation with the recognized classes.

Step 2 - multistep regressive selection

To solve the problem of optimal feature suit we have applied the multistep regressive selection [8]. This method combines together two simple steps, repeated many times. The first step is a forward selection and the second step – the backward elimination. In the forward selection we include the features into the set under fulfillment of strictly defined entrance conditions. In the next step we eliminate the features, which fulfill precisely specified conditions for leaving the set. After changing the composition of features (as a result of the second step) the procedure is repeated by trying to include back the features which may fulfill now the entrance conditions at the changed composition of features.

In the multistep selection we assume the linear model of the system. It means that the actual output variable $y(\mathbf{x}_t)$ of the system is described by the linear relation

(3) $$y(\mathbf{x}_t) = a_1 x_{t1} + a_2 x_{t2} + \ldots + a_k x_{tk}$$

In this relation x_{tj} represents the *j*th feature at the time *t*, $\mathbf{x}_t$ is the vector of the actual set of features selected by the algorithm and $y(\mathbf{x}_t)$ is the actual output of the system. The set of features in the forward selection is adjusted in such a way, that the total error $e_t = |y(\mathbf{x}_t) - y_e(\mathbf{x}_t)|$ at a time *t* is minimized. This error represents the difference between the output signal $y(\mathbf{x}_t)$ of the actual model of the process and the empirical data $y_e(\mathbf{x}_t)$. To get the optimal results we investigate the statistical significance of each feature using Student test. The actual value of the t-test index associated with the *j*th feature is calculated according to the relation

(4) $$t(a_j) = \frac{a_j}{S(a_j)}$$

where $S(a_j)$ is a mean value of the error of approximation of the empirical data by our model. This value is determined by the equation

(5) $$S(a_j) = \sqrt{q_j \frac{\sum_{t=1}^{n} e_t^2}{n-k}}$$

In this equation *n* is the number of observations, *k* – number of features used in the regression model, q_j represents the value of *j*th diagonal element of the inverse of the correlation matrix $(\mathbf{X}^T\mathbf{X})^{-1}$, **X** is the matrix of the empirical data of the number of rows equal to the number of observations and the number of columns equal to the number of classes.

The procedure of feature selection starts from single feature. Then for each next potential feature (j=2, 3, …, k) we check whether it should enter the set of features. It means calculation of the actual value of $t(a_j)$ and comparing it with the nominal value of t_α at the assumed significance level of the Student test and (n-k-1) degrees of freedom. For each feature we have to find the $t(a_j)$ value and choose the maximal one. If the estimated value of $t_{max}(a_j)$ fulfills the inequality $t_{\max}(a_j) \geq t_\alpha$ we add jth feature x_j to the set of chosen features. If this relation does not hold we omit this particular feature. At any step the procedure is tried for all features not already selected. In practice we have chosen the significance level for entering the feature into the chosen set equal 0.05.

In the backward elimination of the feature we try to eliminate the particular feature from the set built in the forward selection if the actual value of the t-test index for this particular feature falls below the minimal significance level t_α (chosen a'priori for leaving the set). If the actually calculated t-test fulfills the condition $t_{\min}(a_j) \leq t_\alpha$ the feature x_j is eliminated from the set. The significance value t_α for leaving the feature set was adjusted on a higher level than for entering the set and was equal 0.1 in all experiments. The procedure of backward elimination is repeated for all features forming the actual suit. As a result of such procedure we select the features which are best correlated with the target and at the same time avoid the features correlated with each other.

The results of feature selection

The discussed above procedure of feature selection was applied to find these harmonics of current, voltage and flux of the induction machine, which carry the most useful information concerned with diagnosis of the rotor bars. We have carried independent numerical experiments concerning the current, voltage and shaft flux of the machine at nominal speed assuming the recognition of 10 defined above classes. At 10 classes we need to select the sets of features suitable for recognition of all 45 combinations of 2-class recognition problems. Five most important frequencies selected for recognition of 2 classes have formed the common set for 45 combinations. The calculations have been carried out for both measures (d_1 and d_2) applied in the first step. The quantity of harmonics selected in this way using both ranking measures are presented in Table I.

Table I The quantity of diagnostic harmonics for current, voltage and flux selected in the first step at application of two ranking measures

Ranking measure	Current	Voltage	Flux
$d_1(f)$	11	12	20
$d_2(f)$	33	35	34

Application of the second step of selection (multistep regression) has been applied to these selected sets of frequencies. As a result of it we have got the reduced set of diagnostic features, presented in Table II.

Table II The quantity of diagnostic harmonics for current, voltage and flux selected after application of the second step for both ranking measures

Ranking measure	Current	Voltage	Flux
$d_1(f)$	4	4	8
$d_2(f)$	18	17	16

These features contained the harmonics best correlated with the decision of the faulty bar diagnosis and at the same time of the least correlation among themselves.

The results of class recognition using SVM

To check the effectiveness of our procedure we have applied the selected features at the recognition of these 10 classes of fault. To get the most reliable results of the tests we have followed the cross-validation approach [3,6]. In this approach we have performed 200 experiments of learning and testing the Gaussian kernel SVM system [6,7] using 90% of data for learning and the remaining 10% for testing. The data used in both sets of all trials have been chosen randomly. The final error of the recognition of classes is defined as the average error of testing the system in all trials. All experiments have been done using Matlab [5].

Table III presents the testing results of the experiments in the form of the mean relative error of class recognition in 200 trials. Only the results of testing data not taking part in learning are presented here. The experiments have been performed using independently the features drawn only from the measured current (the second column), the voltage (the third column) and the shaft flux (the fourth column).

Table III The relative errors of class recognition

Ranking measure	Current features	Voltage features	Flux features
$d_1(f)$	0.26%	9.37%	0.87%
$d_2(f)$	1.35%	3.5%	1.78%

It is evident that the voltage features are least efficient and the current features belong to the best. This is due to the fact that the voltage is rather sensitive to the environmental conditions, able to introduce some noise or other artifacts into the voltage waveform coming from the electrical devices working nearby. The current and flux of the machine are filtered out of these noisy components thanks to the inductance of the machine. Hence they are better suited to generate good diagnostic features.

Conclusions

On the basis of presented results it is evident, that we can rely the rotor bar diagnosis on the observation of the stator current only. This is very fortunate from the practical point of view, since the registration of phase current of the machine is very easy in the on-line mode operation of the machine and does not require any special arrangement connected with stopping the machine at the measurements.

RFERENCES

[1] Guyon I., Elisseeff A., An introduction to variable and feature selection, *Journal of Machine Learning Research*, 2003, vol. 3, pp. 1158 – 1182
[2] Fang R., Induction machine rotor diagnosis using support vector machines and rough set, *Lecture Notes in Computer Science*, 2006, 4114:631-636
[3] Kurek J., Diagnostyka uszkodzeń prętów klatki maszyny indukcyjnej z zastosowaniem sieci neuronowych, rozpr. dokt. PW, 2008
[4] Kurek J., Biernat A., Osowski S., Markiewicz T.,. Diagnosis of induction motor using Support Vector Machine, VII Conf. CPEE, Odessa, Ukraine, 2006, 98-101
[5] Matlab user manual, MathWorks, 2007
[6] Osowski S., Sieci neuronowe do przetwarzania informacji, *Oficyna Wydawnicza* PW, 2006
[7] Schölkopf B., Smola A., Learning with kernels, Cambridge, *MIT Press, MA.* 2002
[8] Weisberg S., Applied linear regression, *Wiley,* 2005,

Authors: dr inż. Jarosław Kurek, *Szkoła Główna Gospodarstwa Wiejskiego, Email: kurek@iem.pw.edu.pl*
prof. dr hab. Stanisław Osowski, Politechnika Warszawska, *Wojskowa Akademia Techniczna, Email: sto@iem.pw.edu.pl*