

The decision making process of a state technical facilities based on rough set and vibroacoustic estimates

Rafal Boniecki^{1,1} and Mirosław Miciak¹

¹University of Technology and Life Sciences in Bydgoszcz, Institute of Telecommunications and Computer Science, Al. prof. S. Kaliskiego 7, 85-796 Bydgoszcz, Poland

Abstract. The article presents idea and develops of an algorithm to support decision-making process for assessment a technical object state, based on vibroacoustic signals and rough set theory. The article discusses how to implement the main elements of the proposed system using Java technology, with Swing graphics library and JFreeChart component. FirebirdSQL engine was used to be a data base for the objects of diagnosis. JDBC technology will be used for generating queries and communication between the application and database.

1 Introduction

Assessment of a technical object state poses the biggest problems connected with diagnostics of technical devices. Methodology of diagnostic tests involves performing classification tests, which include the following diagnostic steps [3]:

- diagnosis - a process involving assessment of a machine state in time t_b ,
- genesis – a process involving reconstruction of the machine previous states, eg. through defining the cause of the machine detected state, in the time of its damage,
- prognosis – a process involving assessment of the machine future state, eg. through scheduling the date of the machine next servicing Θ_a .

The main problems that occur during identification of a technical object state are:

- formulation of the machine diagnosis, prognosis and genesis purpose,
- change of the machine state during its operation,
- description of the machine state by means of its characteristics and dependence on diagnostic parameters,
- solution of the machine state diagnosing,
- solution of the machine state prognosis,
- solution of the machine state genesis.

An analysis of a technical object state is complicated due to operational - measurement data which are often characterized by large dimensionality. In defining the notion of high dimensionality, we mean that data sets (eg. measurement data corresponding to operational states) have a big number of attributes (variables). In practice, it often happens that many of these attributes are strongly correlated with each other and it is enough to consider only a small subset to provide a full picture of the described machine state or identify some regularities (or irregularities) in the object. Although, there are several methods for analysis of technical object operational states, it is usually impossible or hardly possible to identify the object usability only by analysing all the symptoms. Thus, among many tasks, which are involved in machine operation and maintenance, those which enable to reduce dimensionality of the set of symptoms are very important.

¹ Corresponding author: raboni@utp.edu.pl

The major problems connected with assessment of a machine state include: formulation of the state diagnosing, prognosis and genesis goal, change in a machine state during its operation, description of the state by means of its characteristics and parameters of the examined state signal and the dependence between the state characteristics and diagnostic parameters. Thus, the main problems that emerge while solving so formulated problems are:

- choice of ‘the best’ diagnostic parameters describing the machine current state and its changes during operation,
- determination of the diagnostic parameter boundary value,
- determination of the diagnostic parameter generated value and estimation of the technical object state, value performed by it in the past.

The main goal of the study was to design and create an application which will be useful in the process of making decisions on the technical object state assessment. Additionally, the application to be developed should enable communication between computers, regardless of available equipment and the system platform. The software will be written in Java language with the use of graphic library Swing and JFreeChart. FirebirdSQL engine was chosen to be a database storing the objects, whereas, communication and generation of queries, between the application and the database, will be carried out by JDBC technology.

2 Algorithm supporting the process of making decisions on technical object state

Identification of a technical object state, on the basis of available qualitative information, indicates the class of assignment. In this study, a method for assessment of a technical object state is proposed, based on statistic information obtained from vibroacoustic signals. Figure 1 shows an algorithm supporting the process of making decision on the technical object state.

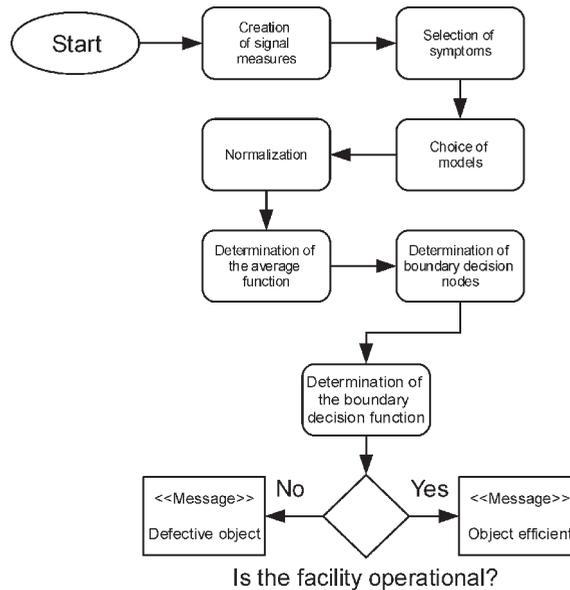


Fig. 1. The proposed algorithm supporting the process of making decision on the state of a technical object, on the basis of vibroacoustic signals and rough set.

2.1 Acquisition of measurement data and determination of signal measures

According to the developed methodology [4], the first stage of the presented algorithm (fig.1.) implementation is acquisition of measurement data. Measurement data is stored in

text files UNV, and for the needs of the created application, the data must be processed to provide the form of a unified dynamic structure. Measurement data, due to a random number of channels, is stored in a two-dimensional dynamic structure whose dimensions depend on the number of the signal samples and on the number of channels.

In the second stage, measures of vibroacoustic signals are created to form selected estimates of the vibration. In the second stage, measures of vibroacoustic signals are created to form selected estimates of the vibration. These measures are most frequently determined from signals in the domain of time, from their transforms in the domain of frequency or time and frequency.

Table 1 shows a list of denotations for selected estimators (symptoms), used in the developed application.

Table 1. List of denotations of selected estimators (symptoms).

	Denotation of estimator	Description of signal
1.	Media	Mean value of signal
2.	RMS(t)	Global value of effective value calculated for a signal in the domain of time
3.	RMS(f)	Global value of effective value calculated for a signal in the domain of freq.
4.	RMSp	Global value of spectral power top value
5.	Max	signal top value
6.	Min	Minimal value of signal
7.	Kurtosis	Kurtosis of data series
8.	Sesgo	Measure of data series asymetry
9.	D. Estdandar	Standard deviation
10.	Factor Forma	Coefficient of shape
11.	Factor Cresta	Top coefficient
12.	Factor Impulso	Coefficient of impulse
13.	Factor Holgura	Coefficient of clearance
14.	Fu (v)	Rice's frequency for vibration speed
15.	F u(d)	Rice's frequency for vibration displacement
16.	H(v)	Coefficient of vibration speed harmoniousness

2.2 Standardization of symptoms

A set of attributes undergoes (in the form of symptom columns) initial data processing to enable their comparison and further analysis. Due to the type of processed data, standardization was used, in result of which, the variable mean expected value assumes zero and the mean variance is equal to one. The most commonly used standardization method is, the so called, standardization z, which can be expressed by the below given formula:

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

where: x – value of symptom, μ - mean value of symptom, σ – standard deviation of symptom.

2.3 Selection of technical object symptoms

There can be many symptoms from one or more signals (measurement channels) but most of them are significantly interfered or independent of the monitored phenomenon. Thus, only those which are correlated with the phenomena should be accepted. Thus, only those which are correlated with this phenomenon, should be selected. The number of such signal measures must be big enough to compensate unavoidable random interferences in each of them. The proposed solution involves choosing symptoms with the use of rough set theory. This technique provides methods for determination of the most important attributes of a computer system without losing classification ability [9]. Rough set theory was developed by Zdzisław Pawlak [1,2,6,7], is based on the concept of upper and lower approximation of the set, approximating space and models of sets. Data is presented in the form of a table, whose columns correspond to the attributes (symptoms) and rows corresponds to particular objects (states of technical objects). Such a record is called an information system [9,10]. The basis of rough set theory is a reduct, which is the main part of the information system. Thanks to this, it is possible to distinguish all the distinguishable objects from the original set of attributes, hence, reducts are minimal subsets of attributes maintaining characteristics of the whole set of attributes. Literature provides many methods for determination of sets of information system reducts. For the needs of the application, the authors have taken into consideration the methods of: Johnson, global, local [5,8]. Significance of the symptoms is described based on the determined set of reducts, where the attributes are ordered according to the number of occurrence in the set. An exemplary set of reducts is presented in table 2.

Table 2. Set of reducts.

	Number	REDUCTS (no. of symptom)
1	9	{ 1, 3, 6, 7, 10, 12, 29, 31, 35 }
2	9	{ 1, 3, 6, 7, 10, 22, 29, 31, 35 }
3	8	{ 1, 3, 6, 10, 11, 12, 15, 31 }
4	9	{ 1, 3, 6, 10, 11, 12, 29, 31, 35 }
5	8	{ 1, 3, 6, 10, 11, 15, 22, 31 }
6	8	{ 1, 3, 6, 10, 11, 22, 29, 31 }
7	8	{ 1, 3, 6, 10, 11, 22, 31, 35 }

Symptoms ordered according to significance are presented in table 3.

Table 3. Symptoms ordered according to significance.

Symptom	1	3	6	10	31	11	12	29	35	22	15
Number	7	7	7	7	7	5	4	4	4	3	1

2.3 Determination of boundary decision function on the basis of *t* symptoms with the same weights

Undertaking a decision on the basis of a *t*-dimensional model can be formulated in the following way:

- In interval $\langle a; b \rangle$ of data there are $(t-1)(n+1)$ of different points $x_0, x_1, \dots, x_t(n+1)$ which can be called decision nodes, and the values of function $y(x)$ in these points are, respectively:

$$f(x_0) = y_0, f(x_1) = y_1, \dots, f(x_t(n+1)) = y_t(n+1)$$

$$\hat{S}_n = \frac{w_1 y_1(x_n) + \dots + w_{t-1} y_{t-1}(x_n)}{w_1 + \dots + w_{t-1}} \tag{2}$$

- Points determined by the symptom which is the first to achieve boundary value do not belong to decision nodes, for differentiation, these points are called extreme decision nodes and the function value in this point is:

$$y = f_s(x) \tag{3}$$

- When decision nodes are available, one can determine mean decision nodes in $\langle a;b \rangle$ interval, being a weighted mean for $t-1$ symptoms measures:

$$\dot{S}_n = \frac{w_1 y_1(x_n) + \dots + w_{t-1} y_{t-1}(x_n)}{w_1 + \dots + w_{t-1}} \tag{4}$$

where w_0, \dots, w_{t-1} are weights of particular symptoms.

$$\dot{S}_n = \frac{w_1 y_1(x_n) + \dots + w_{t-1} y_{t-1}(x_n)}{w_1 + \dots + w_{t-1}}$$

- When mean decision nodes are available, n -degree polynomial is used for approximation to achieve a continuous function, further referred to as the mean function, for $t-1$ symptoms, whereas the function value in this point is:

$$y = f_{t-1}(x) \tag{5}$$

- In order to determine a boundary decision function, the notion of prophylactic or preventive distance is introduced into $\langle a;b \rangle$ interval. Prophylactic distance dx_0 is a non-negative number, unequivocally assigned to a segment, meeting the following conditions:

$$|AB| = 0 \text{ then } \wedge \text{ only then, when } A = B \tag{6}$$

$$|AB| = |BA| \tag{7}$$

$$|AC| \leq |AB| + |BC| \tag{8}$$

where, A, B and C are random points of a plane.

Number $|AB|$ is a prophylactic distance of point A to point B which is expressed by:

$$|AB| = dx_0 = \sqrt{\left(\max_{\sin(t-1)} f_n(x_0) - f_{n-1}(x_0) \right)^2} \tag{9}$$

$$d_{x_0} = \sqrt{\left(\max_{0 \leq n \leq t-1} f_n(x_0) - f_{t-1}(x_0) \right)^2} \text{ when}$$

$$\text{QUOTE gdy } A = \left(x_0, \max_{0 \leq n \leq t-1} f_n(x_0) \right) \text{ i } B = \left(x_0, f_{t-1}(x_0) \right) A = \left(x_0, \max_{0 \leq n \leq t-1} f \right) \tag{10}$$

and

$$\text{QUOTE gdy } A = \left(x_0, \max_{0 \leq n \leq t-1} f_n(x_0) \right) \text{ i } B = \left(x_0, f_{t-1}(x_0) \right) B = \left(x_0, f_{t-1}(x_0) \right) \tag{11}$$

- Knowing prophylactic distance dx_0 , in point x_0 , boundary decision nodes are determined in point x_0 which belongs to interval $\langle a;b \rangle$. Boundary decision node in point x_0 is defined as:

$$\text{QUOTE gdy } A = \left(x_0, \max_{0 \leq n \leq t-1} f_n(x_0) \right) \text{ i } B = \left(x_0, f_{t-1}(x_0) \right) B w_0 = \left(x_0, dx_0, x_0 \right) \tag{12}$$

- Having available decision nodes in interval $\langle a, b \rangle$ we approximate them with n -th degree polynomial to obtain a continuous function, further known as boundary decision function for t symptoms, and the value of the function in this point is:

$$QUOTE \text{ gdy } A = \left(x_0, \max_{0 \leq n \leq t-1} f_n(x_0) \right) \text{ i } B = \left(x_0, f_{t-1}(x_0) \right) y = f_D(x) \quad (13)$$

- Using boundary decision function, decision on the technical object state can be made. with the use of symptoms.

2.5 Selection of models describing damage evolution

After implementation of the process of selection of symptoms, best describing a given object, it is possible to start building a model describing damage evolution of a given object. It enables comprehensive identification of a machine state using deck or stationary diagnostic systems, operating on the basis of the described algorithm. Power and exponential models of symptoms Asr , Xsz , Asz , Ask , were determined for the selected symptoms, by means of the proposed algorithm (section 2.4) for object $Sil24_d2$ (values of parameters of particular models are presented in table 4).

Table 4. Power and exponential models of selected symptoms for $Sil24_d2$ object.

Symptom	Power model	Exponential model	Boundary values
Asr_{Sil24_d2}	$4.89491 * x.^{0.146727168}$	$18.0803 * \exp(1.9551414E-6 * x)$	50
Xsz_{Sil24_d2}	$0.14341 * x.^{0.071569542}$	$0.26399 * \exp(1.7536301E-6 * x)$	5
Asz_{Sil24_d2}	$24.05980 * x.^{0.132450920}$	$77.03300 * \exp(1.8367322E-6 * x)$	200
Ask_{Sil24_d2}	$7.02050 * x.^{0.147649869}$	$25.7803 * \exp(2.0315695E-6 * x)$	69

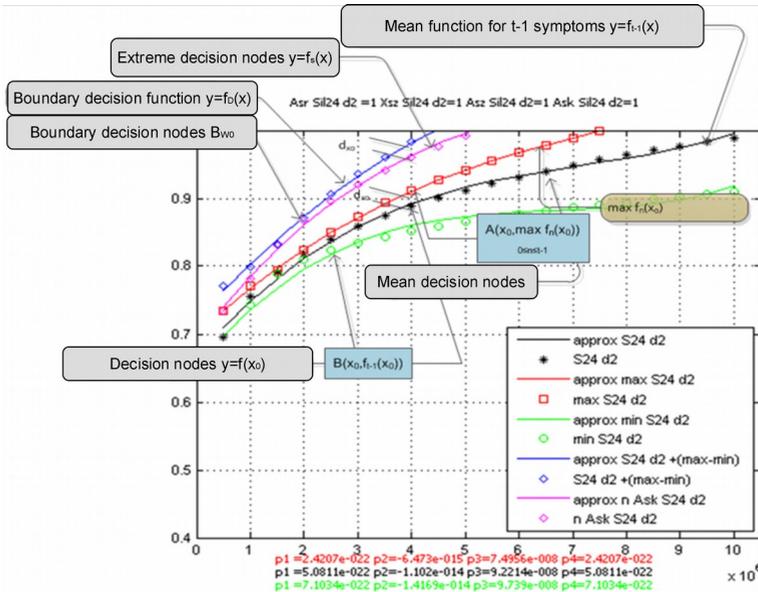


Fig. 2. Determination of the boundary decision function.

On the basis of the determined model $Sil24_d2$, it is possible to define a place where the boundary value is achieved.

3 Conclusions

The article presents implementation of an algorithm to support making decision on the state of a technical object, based on vibroacoustic data. Implementation of the application key

elements has been presented. Application of Java technology provides the application with mobility (independence of the operational system). Utilization of open-source technology significantly decreases the implementation costs. The application uses interface JDBC (Java DataBase Connectivity) for connection with SQL Firebird base. Base sql includes 12 tables, 108 procedures using UDF libraries (User Defined Function).

References

1. J. Bazan, M. Szczuka, A. Wojna, M. Wojnarski, *On Evolution of Rough Set Exploration System*, Lecture Notes in Artificial Intelligence 3366, 592–601, Springer-Verlag (2004)
2. J. Bazan, M. Szczuka, J. Wroblewski, *A New Version of Rough Set Exploration System*, Lecture Notes in Artificial Intelligence 2475, 397–404, Springer-Verlag (2002)
3. L. Bedkowski, *Elementy diagnostyki technicznej*, WAT, Warszawa (1991)
4. R. Boniecki, M. Miciak, R. Wiatr, *Modele nienadzorowanego grupowania obiektów technicznych z wykorzystaniem metod inteligencji obliczeniowej*, Studia i materiały Polskiego Stowarzyszenia Zarządzania Wiedza, 49, 35-48. Bydgoszcz (2011)
5. A. Ohrn, *Discernibility and Rough Sets in Medicine*, Tools and Applications. PhD Thesis, Dep. of Comp. and Inf. Science, Norwegian Univ. of Science and Technology. (1999)
6. Z. Pawlak, *Information Systems - theoretical foundations*. WNT, Warszawa (1983)
7. Z. Pawlak, *Rough sets*, Int. Journal of Parallel Programming 11. 341–356. (1982)
8. M. Salamo, E. Golobardes, *Global, Local and Mixed Rough Sets Case Base Maintenance Techniques*, Recent Advances in Artificial Intelligence Research and Development. 217-234, (2004)
9. E. Szydłowska, *Algorytmy selekcji atrybutów w zadaniach eksploracji danych*, XIII Konferencja użytkowników i deweloperów ORACLE, 262-267. Zakopane (2007)
10. B. Walczak, D.L. Massart, *Rough sets theory*, Gemoetrics and Intelligent Laboratory Systems 47, 1-16 (1999)