



afiliowane przy

Wydziale Nauk
Technicznych
Polskiej Akademii Nauk

Diagnostyka

ISSN 641-6414



Nr 2(38)/2006

RADA PROGRAMOWA

PRZEWODNICZĄCY:

prof. dr hab. dr h.c. mult. **Czesław CEMPEL** *Politechnika Poznańska*

REDAKTOR NACZELNY:

prof. dr hab. inż. **Ryszard MICHAŁSKI** *UWM w Olsztynie*

CZŁONKOWIE:

prof. dr hab. inż. **Jan ADAMCZYK**
AGH w Krakowie

dr inż. **Roman BARCZEWSKI**
Politechnika Poznańska

prof. dr hab. inż. **Walter BARTELMUS**
Politechnika Wrocławska

prof. dr hab. inż. **Wojciech BATKO**
AGH w Krakowie

prof. dr hab. inż. **Lesław BĘDKOWSKI**
WAT Warszawa

prof. dr hab. inż. **Adam CHARCHALIS**
Akademia Morska w Gdyni

prof. dr hab. inż. **Wojciech CHOLEWA**
Politechnika Śląska

prof. dr hab. inż. **Zbigniew DĄBROWSKI**
Politechnika Warszawska

prof. dr hab. inż. **Marian DOBRY**
Politechnika Poznańska

prof. **Wiktor FRID**
Royal Institute of Technology in Stockholm – Szwecja

dr inż. **Tomasz GAŁKA**
Instytut Energetyki w Warszawie

prof. dr hab. inż. **Jan KICIŃSKI**
IMP w Gdańsku

prof. dr hab. inż. **Jerzy KISIŁOWSKI**
Politechnika Warszawska

prof. dr hab. inż. **Wojciech MOCZULSKI**
Politechnika Śląska

prof. dr hab. inż. **Stanisław NIZIŃSKI**
UWM w Olsztynie

prof. **Vasyl OSADCHUK**
Politechnika Lwowska – Ukraina

prof. dr hab. inż. **Stanisław RADKOWSKI**
Politechnika Warszawska

prof. **Bob RANDALL**
University of South Wales – Australia

prof. dr **Raj B.K.N. RAO**
president COMADEM International – Anglia

prof. **Vasyl S. SHEVCHENKO**
BSSR Academy of Sciences Mińsk – Białoruś

prof. **Menad SIDAHMED**
University of Technology Compiègne – Francja

prof. dr hab. inż. **Tadeusz UHL**
AGH w Krakowie

prof. **Vitalijus VOLKOVAS**
Kaunas University - Litwa

prof. dr hab. inż. **Andrzej WILK**
Politechnika Śląska

prof. **Alexandr YAVLENSKY**
Aerospace University Sankt Petersburg – Rosja

prof. dr hab. inż. **Bogdan ŻÓLTOWSKI**
ATR w Bydgoszcz

Wszystkie opublikowane prace uzyskały pozytywne recenzje wykonane przez niezależnych recenzentów.
Obszar zainteresowania czasopisma to problemy diagnostyki, identyfikacji stanu technicznego
i bezpieczeństwa maszyn, urządzeń, systemów i procesów w nich zachodzących.
Drukujemy oryginalne prace teoretyczne, aplikacyjne, przeglądowe z badań, innowacji
i kształcenia w tych zagadnieniach.

WYDAWCA:

Polskie Towarzystwo Diagnostyki Technicznej
02-981 Warszawa
ul. Augustówka 5

REDAKTOR NACZELNY:

prof. dr hab. inż. **Ryszard MICHAŁSKI**

SEKRETARZ REDAKCJI:

dr inż. **Sławomir WIERZBICKI**

CZŁONKOWIE KOMITETU REDAKCYJNEGO:

dr inż. **Krzysztof LIGIER**

dr inż. **Paweł MIKOŁAJCZAK**

ADRES REDAKCJI:

Uniwersytet Warmiński – Mazurski w Olsztynie
Katedra Budowy, Eksploatacji Pojazdów i Maszyn
10-736 Olsztyn, ul. Oczapowskiego 11
tel.: 089-523-48-11, fax: 089-523-34-63
www.uwm.edu.pl/wnt/diagnostyka
e-mail: diagnostyka@uwm.edu.pl

KONTO PTDT:

Bank Przemysłowo Handlowy S.A.
II O/ Warszawa
nr konta: 40 1060 0076 0000 3200 0046 1123

NAKLAD: 275 egzemplarzy

Spis Treści

Preface – Wojciech MOCZULSKI	6
Gracjan GŁOWACKI, Krzysztof PATAN – Uniwersytet Zielonogórski.....	7
<i>Fault Detection In Electrical Drive By Means Of Artificial Neural Networks</i> <i>detekcja uszkodzeń w silniku elektrycznym przy pomocy sztucznych sieci neuronowych</i>	
Rafał SZKLARCZYK – e-lan s.c.	11
<i>CLP Solution For New Variant Of VRP</i> <i>Rozwiązanie nowego wariantu problemu marszrutyzacji pojazdów metodą CLP</i>	
Józef RYBCZYŃSKI – IMP PAN w Gdańsku.....	17
<i>Acquisition Of Some Relations Concerning Bearing Misalignment Defects Applicable In Technical Diagnostics</i> <i>Pozyskiwanie pewnych relacji związanych z defektem rozosiowania łożysk mających zastosowanie w diagnostyce technicznej</i>	
Przemysław WESOŁEK, Marek KUBIAK – Politechnika Poznańska	21
<i>Multiobjective Fuzzy Approach To The Vehicle Routing Problem With Time Windows</i> <i>Wielokryterialny problem planowania tras z oknami czasowymi w wersji rozmytej</i>	
Wojciech SKARKA, Łukasz ZAJĄC – Politechnika Śląska.....	25
<i>Comparing Methods For Implementing VBA In Catia For Elaborating KBE Package</i> <i>Porównanie metod programowych z zastosowaniem VBA w systemie Catia na etapie tworzenie pakietu systemu KBE</i>	
Piotr PRZYSTAŁKA – Politechnika Śląska.....	31
<i>Heuristic Modeling Of Objects And Processes Using Dynamic Neural Networks</i> <i>Heurystyczne modelowanie obiektów i procesów przy pomocy dynamicznych sieci neuronowych</i>	
Wojciech SKARKA – Politechnika Śląska.....	37
<i>Constructing Model Of Knowledge Base According To MOKA Methodology</i> <i>Budowanie modelu bazy wiedzy zgodnego z metodologią MOKA</i>	
Dominik WACHLA – Politechnika Śląska.....	41
<i>Identification Of Submersible Pump Temperature Changes Model Using KDD Methods</i> <i>Identyfikacja modelu zmian temperatury pompy głębinowej z zastosowaniem metod odkrywania wiedzy w bazach danych</i>	
Krzysztof CIUPKE – Politechnika Śląska	45
<i>Cost-Sensitive Feature Selection</i> <i>Selekcja cech z uwzględnieniem kosztu ich pozyskania</i>	
Bogdan WYSOGLĄD , Ryszard WYCZÓŁKOWSKI – Politechnika Śląska	49
<i>An Optimization Of Heuristic Model Of Water Supply System With Genetic Algorithm</i> <i>Optymalizacja modelu heurystycznego sieci wodociągowej z zastosowaniem algorytmu genetycznego</i>	
Eugeniusz ZIENIUK, Andrzej KUŻELEWSKI, Wiesław GABREL – Uniwersytet w Białymstoku	53
<i>Artificial Intelligence Algorithms Combined With The Pies In Identification Of Polygonal Boundary Geometry</i> <i>Algorytmy sztucznej inteligencji połączone z PURC w identyfikacji kształtu wielokątnej geometrii brzegu</i>	
Grzegorz URBANEK – Politechnika Śląska	57
<i>Application Of Evolutionary Algorithms To Identification Of Diagnostic Inverse Models</i> <i>Zastosowanie algorytmów ewolucyjnych w identyfikacji odwrotnych modeli diagnostycznych</i>	

Agnieszka KIMACZYŃSKA, Sebastian KIMACZYŃSKI, Ryszard ROHATYŃSKI – Uniwersytet Zielonogórski61 <i>The Influence Of Some Controlling Parameters On The Efficiency Of An Immune Algorithm</i> <i>Wpływ parametrów sterujących na efektywność algorytmu immunologicznego</i>	61
Anna BZYMEK – Politechnika Śląska65 <i>Application Of Equivalent Source Method To Identification Of Models Of Noise Sources In Automotive Air Conditioning Systems</i> <i>Zastosowanie metody ekwiwalentnych źródeł w celu identyfikacji modeli źródeł hałasu w układzie klimatyzacji samochodowej</i>	65
Krzysztof MENDROK – AGH Kraków69 <i>Zastosowanie analizy modalnej do detekcji uszkodzenia wiaduktu</i> <i>Damage Detection Of The Viaduct With Use Of The Modal Analysis</i>	69
Jan KICIŃSKI – Instytut Maszyn Przepływowych PAN w Gdańsku.....75 <i>Wpływ wymiany ciepła w węzłach łożyskowych na symptomy diagnostyczne maszyn wirnikowych</i> <i>The Influence Of Heat Transfer In Bearings Knots On Diagnostics Symptoms Of Rotor Machines</i>	75
Stanisław RADKOWSKI – Politechnika Warszawska85 <i>Wykorzystanie informacji diagnostycznej w zorientowanej na bezpieczeństwo eksploatacji maszyn</i> <i>Presuming Of Diagnostic Information In Safety Oriented Use Of Machinery</i>	85
Jan Maciej KOŚCIELNY, Bolesław DZIEMBOWSKI – Politechnika Warszawska.....93 <i>Rozróżnialność uszkodzeń w układach liniowych</i> <i>Fault Distinguishability For Linear Systems</i>	93
Stanisław NIZIŃSKI, Sławomir WIERZBICKI – UWM Olsztyn.....101 <i>Inteligentny pojazd mechaniczny jako mechatroniczny obiekt diagnozowania</i> <i>Intelligent Motor Vehicle As A Mechatronic Object Of Diagnosis</i>	101
Lesław BĘDKOWSKI – WAT Warszawa.....109 <i>Diagnozowanie z dwupoziomą komparacją niepewnych symptomów i syndromu stanu obiektu</i> <i>Diagnosing With Double-Level Comparison Of Uncertain Symptoms And Syndrome Of An Object State</i>	109
Tadeusz DĄBROWSKI – WAT Warszawa.....115 <i>Badanie symulacyjne skuteczności diagnozowania komparacyjnego na przykładzie systemu alarmowego</i> <i>Efficiency Of Comparative Diagnosing For An Alarm System – Simulation Results</i>	115
Wojciech CHOLEWA – Politechnika Śląska121 <i>Sieci stwierdzeń w diagnostyce technicznej</i> <i>Statement Networks In Technical Diagnostics</i>	121
Eliza JARYSZ-KAMIŃSKA – Politechnika Szczecińska129 <i>Nadzór nad sprzętem pomiarowo-kontrolnym - unormowania i przepisy prawne</i> <i>Supervising Of Measuring-Control Equipment</i>	129
Krzysztof KOŁOWROCKI – Akademia Morska w Gdyni135 <i>Wielostanowe podejście do analizy bezpieczeństwa systemów</i> <i>Multi-State Approach To System Safety Analysis</i>	135
Adam ROSIŃSKI – Politechnika Warszawska.....143 <i>Metoda optymalizacji przeglądów okresowych transportowych systemów nadzoru</i> <i>Exploitation Strategies Of Monitoring Transport Systems</i>	143
Jerzy GIRTLEK – Politechnika Gdańska.....151 <i>Zastosowanie diagnostyki do decyzyjnego sterowania procesem eksploatacji urządzeń</i> <i>Application Of Diagnostics For Decision Control Over The Process Of Systems Operation</i>	151

Jerzy JAŻWIŃSKI, Sławomir KLIMASZEWSKI – Instytut Techniczny Wojsk Lotniczych Warszawa	159
<i>Identyfikacja stanu zdolności systemu</i> <i>Identification Of The Ability State Of The System</i>	
Wiktor OLCHOWIK – WAT Warszawa	167
<i>Metrologiczne aspekty diagnozowania</i> <i>Metrological Aspects Of Diagnosing</i>	
Janusz MINDYKOWSKI, Tomasz TARASIUK, Edward SZMIT, Daniel CZARKOWSKI – Akademia Morska w Gdyni	171
<i>Diagnostyka izolowanego systemu elektroenergetycznego na przykładzie jednostki pływającej</i> <i>Diagnostics Of Isolated Electrical Power System On Example Of The Ship</i>	
Mariusz BOGDAN, Józef BŁACHNIO – Politechnika Białostocka	177
<i>Analiza sygnału świetlnego odbitego od powierzchni w diagnostyce obiektów technicznych</i> <i>Analysis Of Light Signal Reflected Of A Surface In Technical System Diagnostics</i>	
Janusz GARDULSKI – Politechnika Śląska	187
<i>Badania diagnostyczne amortyzatorów</i> <i>Diagnostic Of Shock Absorbers</i>	
Piotr KRZYWORZEKA, Witold CIOCH – AGH Kraków	199
<i>PLD we wspomaganiu demodulacji drgań maszyn</i> <i>PLD Aided Demodulation Of Machine Vibrations</i>	
Henryk KAŻMIERCZAK – Przemysłowy Instytut Maszyn Rolniczych w Poznaniu	211
<i>Energetyczne charakterystyki degradacji systemu</i> <i>The Energetics Characteristics Of Machine's Degradation</i>	
Marek DZIDA – Politechnika Gdańska, Andrzej ADAMKIEWICZ – Akademia Morska w Szczecinie	217
<i>Zagadnienie bilansowania obiegu cieplnego z zastosowaniem rachunku wyrównawczego w diagnostyce okrętowych turbinowych silników spalinowych</i> <i>A Problem Of Balancing Thermal Cycle Using Reconciliation Calculations In The Diagnostics Of Marine Gas Turbine Engines</i>	

WARTO PRZECZYTAĆ

Zeszyty Naukowe – Katedry Podstaw Konstrukcji Maszyn, Politechniki Śląskiej	225
---	-----

PREFACE

Technical diagnostics is aimed at estimating technical condition of the object of diagnosis, given evidence collected about this object, and diagnostic knowledge and skill attributable to the subject of diagnosis. As an object of diagnosis one can understand either a machine or a piece of equipment, or a useful process the machine/equipment is intended to carry out. The subject of diagnosis has been since many years a human operator or diagnostician. Nowadays, automated or computerized systems are ever more and more often used for this purpose.

For many years, the main area of interests of scientists involved in the development of foundations of technical diagnostics (including the author itself) has been how to measure signals, and then how to select optimal signal features, and subsequently how to arrive at conclusions when countable set of the values of features was given. However, in case of modern technical diagnostics, where ever more and more attention is paid to model-based technical diagnostics of objects and processes, several scientific groups switched their interests to gathering knowledge that could be used in diagnostic reasoning. For this reason, methodology of identifying models of objects and processes attracted the diagnostic society. At the very beginning, physical and input-output models were most frequently used. They were easy to apply in case of quite uncomplicated objects and processes. Unfortunately, quite a few of objects and processes are too complex to allow us to create models efficient enough for applications in model-

based diagnostics. Hence, this approach became infeasible.

To overcome these difficulties, soft (or heuristic) models of objects and processes were introduced. Heuristic models take advantage of achievements of Artificial Intelligence, since they may be represented by such prominent methods and means, as Artificial Neural Networks, Fuzzy Logic Systems, Bayesian Networks, rulesets, qualitative models, and many others. Nowadays there are ever more and more applications of usage of these models.

Bearing this in mind, the editor invited groups active in the area of heuristic modeling to share results of their work with diagnostic society in the country. In the following the Reader can find several descriptions of work done in the framework of heuristic modeling of objects and processes. Individual papers represent all the stages of development of the methodology of heuristic modeling of objects and processes, starting from rough concepts and ideas, and finishing with short reports about verification of methods and software tools. It is worth stressing that the prevailing part of papers has been prepared by young Colleagues. It is our hope that the Readers of "Diagnostyka" will find here some interesting issues.

Finally, special thanks should be expressed to the Editor and the Secretary of Editorial Staff for their kind approval for including this Section "Heuristic Modeling of Objects and Processes" into the current issue, and for their patient assistance by editing.

Wojciech Moczulski
Guest Editor

FAULT DETECTION IN ELECTRICAL DRIVE BY MEANS OF ARTIFICIAL NEURAL NETWORKS*

Gracjan GŁOWACKI, Krzysztof PATAN

Institute of Control and Computation Engineering, University of Zielona Góra
ul. Podgórna 50, 65-246 Zielona Góra, Poland, e-mail: {g.glowacki, k.patan}@issi.uz.zgora.pl

Summary

The paper deals with problem model-based of fault detection electrical drive by using neural networks. The multilayer perceptron with tapped delay lines has been applied to model the diagnosed process at the nominal operation conditions. In turn, decision about faults has been performed using simple MultiLayer Feedforward Network (MFN). The electrical drive under consideration (AMIRA DR300) works in the closed loop and is controlled by PID controller. This laboratory electrical drive renders it positive to simulate a several faulty scenarios. In this way the proposed fault detection scheme can be tested on a number of faulty conditions.

Keywords: fault detection, neural network, neural classifier, modelling, electrical drive.

DETEKCJA USZKODZEŃ W SILNIKU ELEKTRYCZNYM PRZY POMOCY SZTUCZNYCH SIECI NEURONOWYCH

Streszczenie

Artykuł przedstawia problem detekcji uszkodzeń w silniku elektrycznym przy pomocy sieci neuronowych. Do zamodelowania diagnozowanego obiektu pracującego w warunkach normalnych użyto sieci jednokierunkowych z liniami opóźniającymi. Następnie, jako blok decyzyjny o wystąpieniu uszkodzeń zastosowano zwykłe jednokierunkowe sieci wielowarstwowe. Do przeprowadzenia badań wykorzystano silnik prądu stałego firmy AMIRA (DR300). Silnik pracuje w układzie zamkniętym z regulatorem PID i umożliwia symulację pewnych scenariuszy uszkodzeń. Dzięki temu możliwe jest przetestowanie zaproponowanego schematu detekcji uszkodzeń na przykładzie wadliwych warunków pracy obiektu.

Słowa kluczowe: detekcja uszkodzeń, sieci neuronowe, klasyfikator neuronowy, modelowanie, silnik elektryczny.

INTRODUCTION

Electrical Direct Current (DC) and Alternating Current (AC) engines are very often used in many industrial applications [7]. The changing conditions of operation and intensive exploitation result in systematic wearing off of individual parts of engines. This phenomenon can be interpreted as an incipient fault, which in the final phase changes to an abrupt fault and causes large damages in an engine. It is very important in this case to detect faults at an early stage and apply a special procedure to avoid them so that the worn off elements can be replaced. The faults considered manifest themselves at an early stage by a decreased efficiency, but finally, if the fault is not detected, some parts of the engine can be damaged. Thus, it is important to

develop a reliable fault detection algorithm which should detect even small changes in the system behaviour. The traditional methods of monitoring engines require a direct inspection, so the whole process must be stopped for the time of the inspection. Such approaches are usually time-consuming and causes financial losses for the company. Another methods use online identification of engine parameters. Fault detection in this case is done by monitoring the values of the parameters. Unfortunately, these methods require a detailed mathematical model of an engine and the expert who interprets the results should have a deep knowledge about an engine and experience in evaluating its behaviour.

An interesting alternative solution can be obtained using artificial intelligence. Artificial

* The work was supported by the Integrated Regional Operational Programme (Measure 2.6: Regional innovation strategies and the transfer of knowledge) co-financed from the European Social Fund.

neural networks seem to be particularly very attractive when designing fault detection diagnosis schemes. Artificial neural networks can be effectively applied to both the modelling of the plant operating conditions and decision making [1, 2, 3].

1. FAULT DETECTION SCHEME

In order to design fault diagnosis system for the considered-electrical engine, a neural network is used to model the process at normal operation conditions. First, the network has to be trained for this task. Process modelling in closed loop control and is controlled by PID controller. The model is designed applying the multilayer perceptron with tapped delay lines. Two neural models are designed model of tachometer $y^T = f_{NN1}(u)$, where y^T is tachometer measurement and u is control signal, and model of motor drive $y^D = f_{NN2}(u)$, where y^D is motor drive measurement.

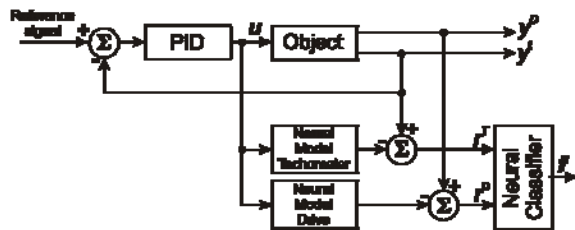


Fig. 1. The FD scheme based on neural network

2. ELECTRICAL ENGINE

The effectiveness of the fault detection scheme has been examined using laboratory stand of the Institute of Control and Computation Engineering of the University of Zielona Góra – AMIRA (DR300). The laboratory stand shown in Fig. 2 can be used to control the rotation speed of a DC engine with a changing load.



Fig. 2. Laboratory stand with a DC engine

Laboratory object considered consists of five main elements: DC engine M_1 , DC engine M_2 , two engine-speed indicators, clutch K .

The input signal for the engine M_1 is an armature current. Its value is determined by a cascade control loop. The servo-amplifier of the current controller works in four modes, thus the current direction can be chosen appropriately according to the demanded direction of rotation. The output of the object is the

rotational speed of the engine. The rotational speed can be measured using two sensors: a tachometer on optical sensor, which generates impulses that correspond to the rotations of the engine. The shaft of the engine M_1 is connected with the identical engine M_2 by the clutch K . The engine M_2 works in the generator mode and the generated current is controlled by another controller. The basic technical data concerning the laboratory system are shown in Table 1.

Table 1. Laboratory system technical data

Variable	Value
<i>Engine</i>	
rated voltage	24 V
rated current	2 A
rated power	30 W
rated speed	3000 rpm
rated moment	0.096 Nm
moment of inertia	$17.7 \cdot 10^{-6} \text{ Kg m}^2$
Resistance	3.13 Ω
<i>Tachometer</i>	
output voltage	5 mV/prm
moment of inertia	$10.6 \cdot 10^{-6} \text{ Kg m}^2$
<i>Clutch</i>	
Moment of inertia	$33 \cdot 10^{-6} \text{ Kg m}^2$

The engine M_1 is controlled using the servo-amplifier, where the control signal has the form of the voltage from range -10V to +10V with amplification 0.4 A/V. The engine M_2 is also controlled using the servo-amplifier, where the control signal has the form of the voltage from the range -10V to +10V with the amplification 0.237 A/V. The tachometer serves to measure indirectly the rotational speed of the engine M_1 . The output range of the tachometer is -10V to +10V [6].

A set of potential faulty scenarios defined for the engine considered. The faults were simulated artificially using the elements of the laboratory system. It was impossible to generate real faults in the laboratory environment. The faults are divided into two groups: tachometer faults and mechanical faults of the engine M_1 , which manifest themselves as a decreasing efficiency of the engine. Tachometer faults f^T were simulated by increasing tachometer measure by 5% (fault f_1^T), 10% (fault f_2^T), 15% (fault f_3^T) and decreasing tachometer measure by 5% (fault f_4^T), 10% (fault f_5^T), 15% (fault f_6^T). In order to generate the mechanical faults, the engine M_2 connected with the engine M_1 via the clutch K was used to simulate an additional faulty load. It was assumed that faults can be abrupt. Mechanical faults f^D were simulated by setting load values to 0.5A (fault f_1^D), 1A (fault f_2^D), 1.5A (fault f_3^D), -0.5A (fault f_4^D), -1A (fault f_5^D), -1.5A (fault f_6^D). The effectiveness of the designed fault detection

system was tested using data generated during fault simulations. Faulty data were prepared for all designed scenarios. Some results of fault detection are presented in the next section.

3. EXPERIMENTAL RESULTS

3.1. Process Modelling

Due to the complex nature of the fault detection problem, the electrical engine was modeled by the multilayer perceptron with tapped delay lines. The considered neural network belongs to the class $N^2_{1,10,1}$ (2 processing layer, 1 input layer, 1 hidden layer with 10 neurons and 1 linear output unit). Each neuron in hidden layer possesses hyperbolic tangent transfer function. Number of past outputs used for determining the prediction was equal to 3, number of past inputs was equal to 3.

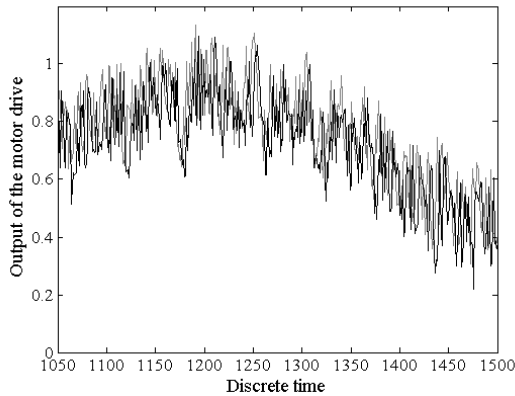


Fig. 3. Output of the motor drive and neural model

In Fig. 3. one can observe that the neural model of the motor drive approximate the modelled object precisely. Moreover, in the Fig. 4. the residual signal for normal condition of the object is presented.

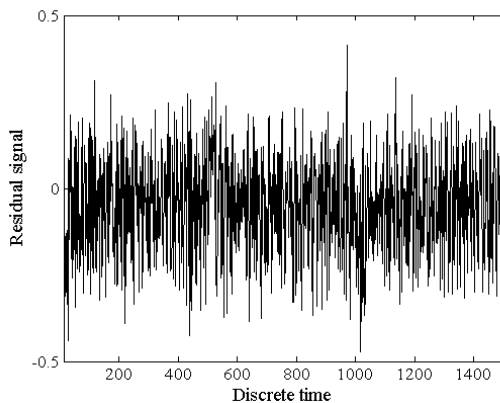


Fig. 4. Residuals for normal conditions

3.2. Tachometer faults

Decision about faults has been performed using simple MultiLayer Feedforward Network (MFN). The neural classifier structure of the neural network was $N^2_{2,5,1}$. Each neuron has sigmoid activation function. The parameters of the network were adjusted by using the well-known Levenberg-

Marquardt method. The learning set contains 300 samples for nominal condition, and for each faulty scenario a set of 100 test samples has been generated.

Analysis of the considered working conditions show that it is extremely difficult to differentiate between particular faulty scenarios. This situation is presented in Fig. 5. Even normal operation conditions are not clearly separated from faulty ones.

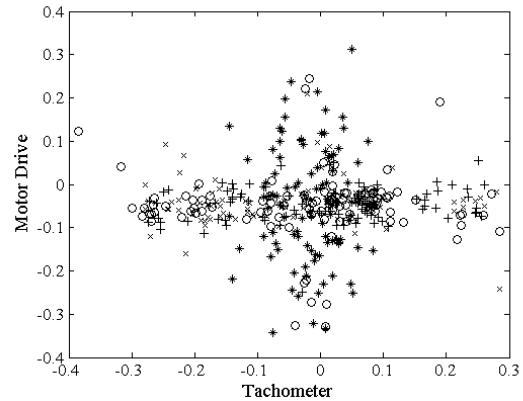


Fig. 5. Distribution of residuals patterns (normal conditions - *, f_1^T - o, f_2^T - x, f_3^T - +)

Taking into account that all tachometer faults are hardly distinguishable one from another, let us consider two classes of system behaviour. First class represented by 0 represents normal operation conditions, and second class represented by 1 represents generalized tachometer fault.

The sensitivity of the proposed fault diagnosis system is checked using the so-called false detection rate [4] defined as follows:

$$r_{fd} = \frac{\sum_i t_{fd}^i}{t_{from} - t_{on}} \quad (1)$$

where t_{fd}^i is the period of i th false fault detection, t_{on} is the benchmark start up time, t_{from} is the period of time from the begin of fault start-up.

The fault detection quality is evaluated using true detection rate [4] defined as follows:

$$r_{td} = \frac{\sum_i t_{td}^i}{t_{hor} - t_{from}} \quad (2)$$

where t_{td}^i is the period of i th true fault detection, t_{hor} is the benchmark time horizon.

In the case of tachometers fault patterns related to the nominal condition overlap patterns collected for faulty scenarios, (Fig.5) what results in high false detection rate at the nominal condition $r_{fd} = 0.2225$.

Table 2. True detection rate

f^T	f_1^T	f_2^T	f_3^T	f_4^T	f_5^T	f_6^T
r_{td}	0,83	0,82	0,80	0,87	0,88	0,87

3.3. Mechanical fault of the engine

In this case problem of distinction between nominal and faulty conditions less complicated. One can see in the Fig. 6 that patterns associated with the nominal state are located in the separated area from the faulty space. Thus, the false detection rate at the nominal condition is much better than in the previous case and $r_{fd} = 0.071$.

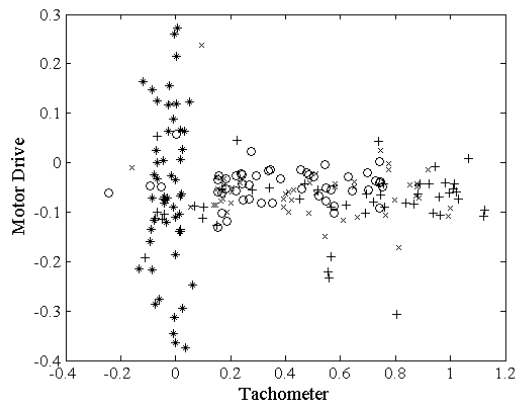


Fig. 6. Distribution of residuals patterns (normal conditions - *, f_1^D - o, f_2^D - x, f_3^D - +)

Table 3. True detection rate

f^D	f_1^D	f_2^D	f_3^D	f_4^D	f_5^D	f_6^D
r_{td}	0,93	0,94	0,78	0,91	0,94	0,95

4. SUMMARY

The paper deals with application of neural networks to model-based fault detection systems. On the basis of the experimental results presented in the paper, the multilayer perceptron with tapped delay lines has appeared to be applicable to the system modelling during designing the FD systems. All experimental simulations were carry out with real object, electrical engine.

On the basis of simulation experiments it to turned out that particular faults are not distinguish and an attempt of their isolation by means of simple MultiLayer Feedforward Network ended in failure.

In both cases faults can be isolated as a group of faults.

REFERENCES

- [1] Korbicz J., Kościelny J. M., Kowalczyk Z., Cholewa W. (Eds.): *Fault Diagnosis: Models, Artificial Intelligence, Applications*. Springer-Verlag, Berlin, 2004.
- [2] Duch W., Korbicz J., Rutkowski L., Tadeusiewicz R. (Eds.): *Neural Network* Vol. 6. Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2000. (in polish)
- [3] Gupta M. M., Jin L., Homma N.: *Static and Dynamic Neural Networks*. John Wiley & Sons, New Jersey, 2003.

- [4] Patan K., Parsini T.: *Identification of neural dynamic models for fault detection and isolation : the case of real sugar evaporation process*. Journal of Process Control, Vol. 15, pp. 67-79.
- [5] Korbicz J., Patan K., Obuchowicz A.: *Dynamic neural networks for process modelling in fault detection and isolation systems*. Journal of Applied Mathematics and Computer Science, Vol. 9, No. 3, pp. 519-546.
- [6] Kowal M.: *Optimization of Neuro-Fuzzy Structures in Technical Diagnostics Systems*. University of Zielona Góra Press, Zielona Góra, 2005.
- [7] Leonhard W.: *Control of Electrical Drives*. Springer, Berlin, 1996.
- [8] Witczak M.: *Advances in Model-Based Fault Diagnosis with Evolutionary Algorithms and Neural Network*. Journal of Applied Mathematics and Computer Science, Vol. 16, No. 1, pp. 85-99.
- [9] Basseville M., Nikiforov I. V.: *Detection of Abrupt Changes: Theory and Application*. Prentice-Hall, Englewood, 1993.
- [10] Janczak A.: *Identification of Nonlinear Systems Using Neural Networks and Polynomial Models*. Springer, Berlin, 2005.



Gracjan GŁOWACKI received the M.Sc. degree in software engineering from Technical University of Zielona Góra, Poland in 2005. He is first-year doctoral student in Computer Science as well as Electrical Engineering. Since 2004 he has been with the Institute of Control and Computation

Engineering. His current research include fault detection and diagnosis, neural networks and their applications.



Krzysztof PATAN received the M.Sc. degree in electrical engineering from Technical University of Zielona Góra, Poland in 1996 and Ph.D. degree in machine design and exploitation from Warsaw University of Technology Poland in 2000. Since 1996 he has been with the Institute of Control and Computation Engineering, University of Zielona Góra. His research interests include dynamic neural networks, fault diagnosis, optimisation methods.

CLP SOLUTION FOR NEW VARIANT OF VRP

Rafał SZKLARCZYK

e-lan s.c.

Piłsudskiego Street 42/2, 43-300 Bielsko-Biała, Poland, e-mail: rafal.szklarczyk@e-lan.pl

Summary

In the paper a mathematical model of MDMGVRP is presented. The two CLP programs solving above are described and discussed. The new variant of VRP consider certain number of commodities (goods) transported. In classic VRP there is only one commodity. The second main difference to classic problem is limited stock in depots. This two aspects cause the problem more complicated and prevent from direct applying of common known algorithms for solving the MDMGVRP. Two CLP solver tools were used to formulate CLP programs. The first was Cosytec CHIP the other was GNU-Prolog. Finally a demonstration problem is presented and solution is discussed.

Keywords: vehicle routing problem, VRP, MDMGVRP, CLP, CHIP, GNU-Prolog.

ROZWIĄZANIE NOWEGO WARIANTU PROBLEMU MARSZRUTYZACJI POJAZDÓW METODĄ CLP

Streszczenie

W artykule przedstawiono nowy wariant problemu marszrutyzacji pojazdów – MDMGVRP, czyli problem marszrutyzacji pojazdów z wieloma magazynami i wieloma asortymentami. Zagadnienie zostało opisane modelem matematycznym a następnie rozwiązane metodą CLP. Opisano dwa programy wykonane przy użyciu narzędzi CLP: Cosytec CHIP i GNU-Prolog. Na koniec przedstawiono przykładowe dane oraz rozwiązanie uzyskane przy pomocy opisanych programów.

Słowa kluczowe: problem marszrutyzacji pojazdów, VRP, MDMGVRP, CLP, CHIP, GNU-Prolog.

INTRODUCTION

In nowadays world of high level of competition there is extremely important to hold a strict regime of cost control. Especially parts of cost that do not bring any added value to a products shall be most reduced. A transportation cost is one of the kind.

Control of logistic processes is one of main matters of concern of operational research. The CLP is a method that help to perfect us dealing with the problem.

1. MDMGVRP VS VRP

You face Vehicle Routing Problem every time when you need to determine routes for certain fleet of vehicles which have limited capacity and are located in one or more depots. The vehicles must distribute goods to a clients. Every client has a demand for a certain number of goods. The distances between client locations and depot locations are given. Usually the aim is to minimize sum of total distance traveled by all cars.

About fifty years have elapsed since Dantzig and Ramser [1] introduced VRP but in the last decade VRP was a matter of interest for operational

research. Good overview of the problem was presented by Toth and Vigo [2].

A classification of VRP was introduced by Garn [3]. He also proposed quite wide set of bibliography [4] Sometimes TSP with multiple salesmen is classified as VRP [5], but in fact it is much less complicated and is particular instance of VRP. There is quite a lot of VRP variants, all focused on different problem restrictions.

Every day life comes with various situations, that is why we need to formulate new mathematical models and find a new solution methods. MDMGVRP is for Multiple Depots Multiple Goods Vehicle Routing Problem.

In this case demands are determined for certain set of commodities. The additional constraint is that a stores of goods and a vehicles can be located in a certain number of depots. Every depot has limited stock of goods. This are reasons that make new variant of VRP more complicated.

2. MATHEMATICAL MODEL

Let M be a number of cities, V a number of vehicles and A number of commodities (types of goods). We assume that commodities are packed with the same kind of wrapping, so their number is

determined in same units, which are packages of commodity.

The distances between cities are given with matrix O :

$$O = [o_{i,j}] \quad i, j = 1, \dots, M$$

where i means start city, j means end city.

There are depots in some cities. In depots a limited number of commodities is stored:

$$S = [s_{a,m}] \quad a = 1, \dots, A; \quad m = 1, \dots, M; \quad s_{a,m} \in \mathbb{N}$$

where a is means commodity, m the city where $s_{a,m}$ packages of commodity a is stored.

There are some clients located in the cities. Each client has certain demand:

$$D = [d_{a,m}] \quad a = 1, \dots, A; \quad m = 1, \dots, M; \quad d_{a,m} \in \mathbb{N}$$

where a is means commodity, m the city where $d_{a,m}$ packages of commodity a is required.

There can not be a depot and a client located in the same city:

$$\forall m \in \{1, \dots, M\}: \sum_{a=1}^A s_{a,m} = 0 \vee \sum_{a=1}^A d_{a,m} = 0$$

Each vehicle has certain capacity l_v that is determined with package of commodity unit:

$$l_v \in \mathbb{N} \quad v = 1, \dots, V$$

and each vehicle has exploitation cost c_v , which is determined as a cost of travel of one distance unit by the vehicle v :

$$c_v \in \mathbb{N} \quad v = 1, \dots, V$$

Each vehicle starts from specified depot:

$$p_v \quad v = 1, \dots, V; \quad p_v \in \{1, \dots, M\}$$

where p_v means the city of start of vehicle v .

Let us define:

$$X^v = [x_{m_1, m_2}^v] \quad v = 1, \dots, V; \quad m_1, m_2 = 1, \dots, M; \quad x_{m_1, m_2}^v \in \{0, 1\}$$

where X^v is the matrix of travel of vehicle v , that is determined in following way:

$$x_{m_1, m_2}^v = \begin{cases} 1 & \text{if vehicle } v \text{ travels from city } m_1 \text{ to } m_2 \\ 0 & \text{otherwise} \end{cases}$$

Particular vehicle can visit particular city at most once, so:

$$\sum_{m_1=1}^M x_{m_1, m_2}^v \leq 1 \quad \forall v \in \{1, \dots, V\}, \forall m_2 \in \{1, \dots, M\}$$

That is also known that particular vehicle leaves particular city at most once:

$$\sum_{m_2=1}^M x_{m_1, m_2}^v \leq 1 \quad \forall v \in \{1, \dots, V\}, \forall m_1 \in \{1, \dots, M\}$$

If particular vehicle takes part in fulfilling the demands, it would leave its start depot. Otherwise it stays at start depot and all its travel matrix elements are equal to 0:

$$\forall v \in \{1, \dots, V\}: \sum_{m_2=1}^M x_{p_v, m_2}^v = 0 \Rightarrow \sum_{m_1=1}^M \sum_{m_2=1}^M x_{m_1, m_2}^v = 0$$

Let Z^v be a matrix of load, defined in following way:

$$Z = [z_{a,m}^v] \quad v = 1, \dots, V; \quad a = 1, \dots, A; \quad m = 1, \dots, M; \quad z_{a,m}^v \in \mathbb{N}$$

where $z_{a,m}^v$ is the number of packages with commodity a loaded to vehicle v in city m .

The total number of units loaded in a particular city can not be greater than a stock in the city (depot):

$$\forall m \in \{1, \dots, M\}, \forall a \in \{1, \dots, A\}: \sum_{v=1}^V z_{a,m}^v \leq s_{a,m}$$

It also have to be less than a capacity of the vehicle:

$$\forall m \in \{1, \dots, M\}, \forall a \in \{1, \dots, A\}, \forall v \in \{1, \dots, V\}: z_{a,m}^v \leq l_v$$

The matrix of unload is defined in the same way:

$$R = [r_{a,m}^v] \quad v = 1, \dots, V; \quad a = 1, \dots, A; \quad m = 1, \dots, M; \quad r_{a,m}^v \in \mathbb{N}$$

where $r_{a,m}^v$ is the number of packages with commodity a unloaded from vehicle v in city m .

All demands shall be fulfilled. It means that number of commodities unloaded in a particular city is equal to demand in the city:

$$\forall m \in \{1, \dots, M\}, \forall a \in \{1, \dots, A\}: \sum_{v=1}^V r_{a,m}^v = d_{a,m}$$

In MDMGVRP (apart from SDVRP) one client need to be served by a single vehicle:

$$\forall m \in \{1, \dots, M\}, \exists v \in \{1, \dots, V\}: r_{a,m}^v = d_{a,m} \quad \forall a \in \{1, \dots, A\}$$

Each vehicle loads in total the same number of commodities as total number of unloaded commodities during its journey:

$$\forall v \in \{1, \dots, V\}, \forall a \in \{1, \dots, A\}: \sum_{m=1}^M z_{a,m}^v = \sum_{m=1}^M r_{a,m}^v$$

A vehicle must visit a city to unload or load commodities. The model must define such constraint, so vehicle must travel in or out the city if number of loaded or unloaded commodities is greater than 0. In this model vehicles do not have to return to their start depots. So there is:

$$\forall v \in \{1, \dots, V\}, \forall m \in \{1, \dots, M\}: z_{a,m}^v > 0 \vee r_{a,m}^v > 0, \quad a = 1, \dots, A \Rightarrow$$

$$\sum_{m_2=1}^M x_{m, m_2}^v = 1 \vee \sum_{m_1=1}^M x_{m_1, m}^v = 1$$

Each vehicle can not be overloaded and underloaded in any visited city. This means that a load of vehicle, which is equal to sum of load state in previous city and change of load (load or unload) in current city, must be less than or equal to vehicle capacity and greater than or equal to zero:

$$\forall v \in \{1, \dots, V\}, \forall m \in \{1, \dots, M\}: 0 \leq B(v, m) \leq l_v$$

where:

$$B(v, m) = \sum_{a=1}^A (z_{a,m}^v - r_{a,m}^v) + B(v, m_0) \quad x_{m_0, m}^v = 1$$

and in the start depot of vehicle v :

$$B(s, p_v) = \sum_{a=1}^A (z_{a, p_v}^v - r_{a, p_v}^v)$$

You need to remember that vehicle shall visit a city to perform load or unload there. You can describe this relation in following way:

$$\sum_{a=1}^A (z_{a, m_0}^{v_0} + r_{a, m_0}^{v_0}) \neq 0 \quad v_0 \in \{1, \dots, V\}, m_0 \in \{1, \dots, M\} \Rightarrow$$

$$(\exists m \in \{1, \dots, M\}: x_{m_0, m}^{v_0} = 1 \vee \exists m \in \{1, \dots, M\}: x_{m, m_0}^{v_0} = 1)$$

Total cost depends on total distance traveled by all vehicles and their exploitation costs, so:

$$F(X^1, \dots, X^V) = \sum_{v=1}^V \sum_{m_1=1}^M \sum_{m_2=1}^M x_{m_1, m_2}^v * o_{m_1, m_2} * c_v$$

To solve the problem means to find such X^v , Z^v and R^v , that keep all relations defined above. The goal of optimization is to find the solution with minimum value of objective function F .

3. CLP PROGRAMS

Constraint Logic Programing is still not very popular approach however it is flexible and effective.

Two CLP tools have been used to formulate the programs: Cosytec CHIP and GNU Prolog. Both are Prolog descendant with constraint over finite domains functions. They are similar but there are some differences. The syntaxes are slightly different but more important differences are in functionality. CHIP offers more build-in predicates [6] that help a lot to solve even very complicated problems. GNU Prolog has less predicates but offers a possibility of extending features with function libraries written in C. It also accepts nonlinear constraints [7] that was useful to solve MDMGVRP. Another advantage of GNU-Prolog is that the program is distributed freely under GPL license.

The MDMGVRP programs consist of 4 main sections:

1. Entry data and constraints,
2. Main (computing) part,
3. Additional predicates,
4. Distance matrix definition.

First vehicle parameters are defined:

$L1=15$, $C1=3$, $MS1=4$,

L means capacity, C exploitation cost, MS start depot. The digits following letters are indexes. The mean number of vehicle. When there are more than one index of variable in the program, the rule of order is that the first index means vehicle, second is for commodity kind, third is for city.

There are stocks determined forth:

$S=[S1, S2, S3, S4]$,
 $SA1=[SA11, SA12, SA13, SA14]$,
 $SA2=[SA21, SA22, SA23, SA24]$,
[...]
 $S1\# = SA11+SA21+SA31+SA41+SA51+SA61$,
 $S2\# = SA12+SA22+SA32+SA42+SA52+SA62$,
[...]
 $SA1=[0, 0, 0, 7]$,
 $SA2=[0, 0, 4, 2]$,

$S1, S2...$ mean total stocks of all commodities in the cities. $SA21$ means stock of commodities 1 in city 2.

The demands are defined in the same way:

$D=[D1, D2, D3, D4]$,
 $DA1=[DA11, DA12, DA13, DA14]$,
 $DA2=[DA21, DA22, DA23, DA24]$,

In the last section of program a distances between the cities are defined. It is done with predicate *odl*:

$odl(, 0, 0)$.
 $odl(1, 1, 0)$.
 $odl(1, 2, 1)$.
 $odl(1, 3, 3)$.
[...]

Predicate *odl* has three arguments: first is source city, second is destination city and third the distance

between them. The first row is used to calculate objective function value in case when vehicle do not travel from some city.

Next line means that the distance from city 1 to city one is 0 distance units. It is obvious. Third line means that distance from city 1 to city 2 is 1.

Back to the first row it means that distance from any city to dummy city number 0, which really means that a vehicle stops, is equal to 0 distance units. When a vehicle stops it do not affect with growth of objective function. It is more understandable after definition of travel matrix.

The definitions in first section of the program are mixed with constraints like:

$ZA123\# \leq SA23$,

This example means that loading of vehicle 1 with commodity 2 in city 3 shall be less or equal appropriate stock in the city.

One of most important constraints is:

$R12 * D2\# = R12 * R12$

It holds when total unload of vehicle 1 in city 2 is equal to demand in the city or when is equal to 0. It is useful to define relation which says that single customer shall be served by single vehicle. However it is possible to use above only in GNU Prolog. In CHIP, because there are no nonlinear constraints, you need to write it in a different way:

if ($R12\# \neq D2$) then ($R12\# = 0$),

The travel matrix X^v in program is replaced with vector T^v , which size is equal to number of cities. The index of element means a source city, the value of an element means destination city. If the value is equal to 0, it means that vehicle stops, it does not leave the city.

For example matrix X :

$$X = \begin{bmatrix} 0, & 1, & 0, & 0 \\ 0, & 0, & 0, & 1 \\ 1, & 0, & 0, & 0 \\ 0, & 0, & 1, & 0 \end{bmatrix}$$

is replaced with a vector T :

$T = [2, 4, 1, 3]$

After definitions and constraints the main section of program is located:

```
statistics(real_time, [Czas_s, _]),
    fd_minimize((
        fd_labeling(RA11),
        fd_labeling(RA12),
        fd_labeling(RA21),
        fd_labeling(RA22),
        fd_labeling(ZA11),
        fd_labeling(ZA12),
        fd_labeling(ZA21),
        fd_labeling(ZA22),

        niezaladowane_stoja(Z1, T1),
        niezaladowane_stoja(Z2, T2),
        jedz_gdzie_trzeba(Z1, R1, T1, MS1),
        jedz_gdzie_trzeba(Z2, R2, T2, MS2),

        fd_labeling(T1),
        fd_labeling(T2),
        sprawdz_przeladowanie(MS1, T1, Z1, R1, 0, L1),
        bilans(MS1, T1, ZA11, RA11, 0),
        bilans(MS1, T1, ZA12, RA12, 0),
        sprawdz_przeladowanie(MS2, T2, Z2, R2, 0, L2),
        bilans(MS2, T2, ZA21, RA21, 0),
        bilans(MS2, T2, ZA22, RA22, 0),
```

```

dlugosc_trasy(T1,1,DT1),
dlugosc_trasy(T2,1,DT2),
CEL is (DT1*C1)+(DT2*C2),

statistics(real_time,[Czas_l,_]),
    write('Cel'),nl,
    write(CEL),nl,
    write('EUREKA!'),nl,
    write('Zaladunki'),nl,
    write('Z1='),write(Z1),nl,
    write('ZA11='),write(ZA11),nl,
    write('ZA12='),write(ZA12),nl,
    write('Z2='),write(Z2),nl,
    write('ZA21='),write(ZA21),nl,
    write('ZA22='),write(ZA22),nl,
    write('Rozladunki'),nl,
    write('R1='),write(R1),nl,
    write('RA11='),write(RA11),nl,
    write('RA12='),write(RA12),nl,
    write('R2='),write(R2),nl,
    write('RA21='),write(RA21),nl,
    write('RA22='),write(RA22),nl,
    write('Trasy'),nl,
    write('T1='),write(T1),nl,
    write('T2='),write(T2),nl,
    write(''),nl,
    Czas is Czas_l-Czas_s,
    write('Czas liczenia='),
    write(Czas),write(' ms'),nl,
    write(''),nl
),
CEL
),

```

In GNP Prolog the main role plays predicate *fd_minimize*. It runs a code given as first parameter until the second parameter (in this case variable CEL) achieves minimum value.

The first part of code inside *fd_minimize* is labeling of loads and unloads. Then there are some additional constraints in predicates *niezaladowane_stoja* and *jedz_gdzie_trzeba*. They provide that every vehicle that do not fulfill any demand stays in start depot and that every city where loads or unloads are performed is visited by the appropriate vehicle.

Then there is labeling of vectors Tv and predicates *sprawdz_przeladowanie* and *bilans*. The predicates provide that all vehicle routes are continuous. Vector that contain two routes like $T=[2,1,0,5,4]$ are forbidden. The predicates also provide balance checking for routes. It means that any vehicle can not be over or underload all route long.

4. DEMONSTRATIONAL PROBLEM

Consider MDMGVRP with given 2 vehicles, 2 commodity kinds and 6 cities. There are two depots and four customers.

Below stock values are defined:
 $SA1=[0,0,0,7,2,0]$,
 $SA2=[0,0,0,4,0,0]$,

and demand values:
 $DA1=[2,4,1,0,0,0]$,
 $DA2=[1,2,0,0,0,0]$,

Vehicle parameters are following:
 $L1=10$, $C1=3$, $MS1=6$,
 $L2=6$, $C2=1$, $MS2=6$,

The distance matrix:

$$O = \begin{bmatrix} 0, & 1, & 3, & 2, & 4, & 5 \\ 1, & 0, & 1, & 2, & 3, & 4 \\ 3, & 1, & 0, & 1, & 3, & 3 \\ 2, & 2, & 1, & 0, & 1, & 2 \\ 4, & 3, & 3, & 1, & 0, & 1 \\ 5, & 4, & 3, & 2, & 1, & 0 \end{bmatrix}$$

GNU Prolog that is running on machine with Intel processor (Celeron 1.4GHz, 256MB RAM, OS MS Windows XP Home) solves the problem above in time of 32397 milliseconds and shows the solution:

```

Cel
15
EUREKA! Rozwiazanie ostateczne:
Zaladunki
Z1=[0,0,0,8,2,0]
ZA11=[0,0,0,5,2,0]
ZA12=[0,0,0,3,0,0]
Z2=[0,0,0,0,0,0]
ZA21=[0,0,0,0,0,0]
ZA22=[0,0,0,0,0,0]
Rozladunki
R1=[3,6,1,0,0,0]
RA11=[2,4,1,0,0,0]
RA12=[1,2,0,0,0,0]
R2=[0,0,0,0,0,0]
RA21=[0,0,0,0,0,0]
RA22=[0,0,0,0,0,0]
Trasy
T1=[0,1,2,3,4,5]
T2=[0,0,0,0,0,0]
Czas liczenia=32397 ms

```

It also signalize that there are more optimal solutions (with the same objective function value). It is possible to view them all.

CHIP needs time of 38505 milliseconds for completing the task and shows only the first solution found.

5. SUMMARY

CLP is useful method to solve MDMGVRP a new variant of VRP and quite complicated one.

MDMGVRP is still an issue of research as there is a need to determine maximal size of the problem that can be solved with CHIP and GNU-Prolog.

The CLP solution should be compared with some other methods and maybe performance should improved by using heuristic methods or other means.

REFERENCES

- [1] Dantzig G. B., Ramser R. H. *The Truck Dispatching Problem*. Management Science, 1959, Vol. 6, pp. 80–91.
- [2] Toth P., Vigo D. *The Vehicle Routing Problem*. SIAM Monographs on Discrete Mathematics and Applications, 2002, pp. 1–25.
- [3] Garn W.: http://osiris.tuwien.ac.at/~wgarn/VehicleRouting/vehicle_routing.html. Internet pucation, Wien, 2002.
- [4] Garn W.: <http://osiris.tuwien.ac.at/~wgarn/VehicleRouting/neo/biblio.html>. Internet pucation, Wien, 2002.

- [5] Jadczyk R., Trzaskalik T.: *Algorytmy genetyczne, ewolucyjne i metaheurystyki. Informatyka w badaniach operacyjnych*, Prace naukowe AE, Katowice, 2005, pp. 79-111.
- [6] COSYTEC SA.: *CHIP System documentation*. Orsay Cedex, 1997.
- [7] Diaz D.: *GNU Prolog documentation*, Paris, 2002.



Rafał SZKLARCZYK is a graduate of Academy of Computer Science and Management in Bielsko-Biała (computer science engineer) and Academy of Mining and Metallurgy in Cracow (master of management). Now he is running

a software engineering company with success for over five years. He is also a PhD student at Silesian University of Technology with special interest on operational research and CPL methods.

ACQUISITION OF SOME RELATIONS CONCERNING BEARING MISALIGNMENT DEFECTS APPLICABLE IN TECHNICAL DIAGNOSTICS

Józef RYBCZYŃSKI

Institute of Fluid-Flow Machinery, Polish Academy of Sciences
80-952 Gdańsk, ul. Fiszer 14, tel.: (+48) 058 699 52 73, fax: (+48) 058 341 61 44
E-Mail: ryb@imp.gda.pl

Summary

The investigations aimed at determining the tolerable misalignment areas of two neuralgic bearings in the turbine set, taking into account criteria of permissible load of the bearings. The areas have been evaluated by numerical analysis of the discrete model of the machine with the aid of a package of MESWIR computer codes. The ranges of tolerable misalignments of the bearings to any direction are presented in the graphs as the envelopes of maximum permissible bearing load. Basing on the method it is possible to correct the location of particular bearings to optimize tolerable bearing misalignment areas. The research will make possible to formulate a system of diagnostic relations connected with the bearing misalignment defects.

Keywords: technical diagnostics, rotating machine, slider bearing.

POZYSKIWANIE PEWNYCH RELACJI ZWIĄZANYCH Z DEFEKTEM ROZOSIOWANIA ŁOŻYSK MAJĄCYCH ZASTOSOWANIE W DIAGNOSTYCE TECHNICZNEJ

Streszczenie

Badania miały na celu wyznaczenie obszarów bezpiecznych rozosiowań dwóch neuralgicznych łożysk turbozespołu z punktu widzenia kryterium dopuszczalnego obciążenia łożysk. Obszary te zostały wyznaczone w drodze numerycznej analizy dyskretnego modelu maszyny przy użyciu pakietu oprogramowania MESWIR. Zakresy bezpiecznych rozosiowań łożysk w dowolnym kierunku są prezentowane na wykresach jako miejsce geometryczne maksymalnych dopuszczalnych obciążeń łożysk. Bazując na zastosowanej metodzie możliwa jest korekta położenia poszczególnych łożysk maszyny w celu optymalizacji obszarów dopuszczalnych rozosiowań łożysk. Praca umożliwi sformułowanie systemu relacji diagnostycznych związanych z defektem rozosiowania łożysk.

Słowa kluczowe: diagnostyka techniczna, maszyna wirnikowa, łożysko ślizgowe.

1. INTRODUCTION

Large-power turbine sets consist of a number of rotors linked together by couplings and supported in a number of slider bearings. For such a machine, the bearings have to be set in a precisely defined manner with respect to each other, namely shafts should create a catenary line [1, 2, 3]. Deviations of bearing locations with respect to their design location change operating conditions in this one bearing and also in other bearings. The static load of the shafts and bearings as well as the shaft line shape is changed [2, 4, 5]. Exceeding the tolerable bearing misalignments makes further operation of a machine at least dangerous [2, 5].

Relations between bearing misalignments and the machine bearing loads, make in fact a set of relations linking the defects and their symptoms. Therefore matching like bearing displacement – bearing load characteristics can be considered as diagnostic relations and included to the diagnostic knowledge basis for the examined machine. Such

a knowledge basis is necessary for automatic diagnostic systems.

The motivations for examining the effect of bearing misalignment defect on bearing load characteristics originated from the earlier research activities in the subject, which resulted in determining tolerable displacements of bearings in the 13K215 turbine set in two directions only: horizontally to the left and to the right and vertically upward and downward [6, 7].

2. APPLICATION OF THE DIAGNOSTIC RELATIONS IN DIAGNOSTIC SYSTEMS

For their operation, automatic diagnostic and advisory systems need the information, based on which they can formulate diagnoses, suggest, or even make decisions in certain situations of object's behaviour. This information should be delivered to the diagnostic system in such a form that the inference algorithms, implemented in the system in the form of rules bearing the names of diagnostic relations, could make use of it. Preferably, this

information should have a form of sets of already formulated rules, but advanced diagnostic systems can formulate these rules themselves. The diagnostic relations are the cause-effect relations linking the type of defect with its symptoms [3, 9]. If the set of diagnostic relations is based on discrete values assumed for defect simulations, these relations can be given a form of simple logical sentences of the following type: „if (cause) then (effect)”, ($P \rightarrow S$). Sets of relations of $P \rightarrow S$ type are usually obtained using of the simulation method. The method consists in assuming a defect in the machine model and finding a response to the defect simulated in this way. A direction in which diagnostic and expert systems are developed is „artificial intelligence”. This name is given to systems capable of simulating the way of inference characteristic of human beings. A similar method of formulating conclusions in computer science is referred to as heuristics. The heuristic inference allows finding an approximate solution to the examined problem, without guarantee of its correctness. However, this solution is frequently the optimal and only possible one in a certain situation, for instance when an exact solution algorithm or sufficient database are missing.

The investigations presented in this article aimed at creating sets of diagnostic relations describing bearing misalignment defect in a multi-support rotating machine, a large-power turboset. In the examined the causes were the locations and scales of rotor misalignment defect, described by numerical values of displacements of particular bearings. Each individual cause P_i can be defined by a set of three numerical values:

$$P_i = [\text{bearing_number}, \Delta_x, \Delta_y],$$

where: Δ_x – horizontal bearing displacement,

Δ_y – vertical bearing displacement.

A displacement of the bearing to any direction may be presented as a vector sum of the Δ_x and Δ_y . The effect of the defect, describing the resultant static and dynamic state of the machine, is expressed by a set of values of quantities selected taking into account certain requirements of the diagnostic system. Generally, the set of values describing the effect is larger, and its size depends on the number of parameters considered significant for diagnostic purposes. If the state of bearing loads is of some interest from the diagnostic point of view, this set may include average pressures on bush surfaces in bearings in horizontal and vertical projections. If of certain interest is the dynamic state of the machine, this set can be composed of values characterising the intensity of absolute and relative vibrations, horizontal and vertical, of all bearings.

From the point of view of practical applications in automatic diagnostic systems of high importance are the relations, which allow indicating a defect from the set of its symptoms, i.e.: $S \rightarrow P$. These are inverse relations to those directly obtained using simulation methods. A way to find the relations of $S \rightarrow P$ type is inverting the $P \rightarrow S$ type relations

obtained using simulation methods. An expected result of this operation is a set of mutually unique relations. This task can be done in an exact and unambiguous way only when certain assumptions concerning the sets of causes and effects are made. In practice it is not unambiguous due to the physical nature of the object of examination and the resultant nature of phenomena treated as symptoms, as well as the nature of the simulation procedure applied for obtaining these relations. The set of simulation results is limited in size and does not include all possible effects of the assumed defects. Moreover the results of simulations are burdened with errors, therefore obtaining sets of inverse diagnostic relations is only possible using the mentioned above heuristic methods.

These methods were successfully applied by a research team of the Silesian University of Technology for examining bearing misalignment defect in a turboset [8, 9]. They inversed the diagnostic relations, which had been obtained from simulation calculations in IF-FM, and then used these relations in the developed experimental diagnostic system making use of heuristic procedures [8]. This system is capable of indicating, with certain probability, the type of defect from given symptoms. The experimental diagnostic system was trained on sets of simulation-based diagnostic relations describing the bearing misalignment defect in the 13K215 turboset in horizontal and vertical direction [6, 7]. These data included relations between maximum permissible displacement of bearings 5 and 6 in four directions: right, left, up and down, and the vibration state of the turboset resulting from those displacements. The same material was used for verifying the operation of the system.

In this article presented are the results of investigations being the continuation and extension of the above investigations. Determined were ranges of permissible displacements of the two turboset bearings in arbitrary direction from the point of view of the permissible bearing load criterion and the permissible vibration criterion. Discrete maximum states of permissible bearing displacements were linked with the parameters characterising the machine state, such as loads and vibrations of all bearings. This way sets of diagnostic relations were created in the form:

$$P_i [\text{bearing number}, \Delta_x, \Delta_y] \rightarrow$$

$$\rightarrow S_i [\text{bearing number}, Q_x, Q_y],$$

$$P_i [\text{bearing number}, \Delta_x, \Delta_y] \rightarrow$$

$$\rightarrow S_i [\text{bearing number}, A_x, A_y, V_x, V_y].$$

All these sets of diagnostic relations can be used in the experimental diagnostic system developed at the Silesian University of Technology. The article presents part of this work, which is the results of investigations of permissible displacements of bearings 5 and 6 in the 13K215 turboset, and related diagnostic relations from the point of view of bearing load criterion. The diagnostic relations

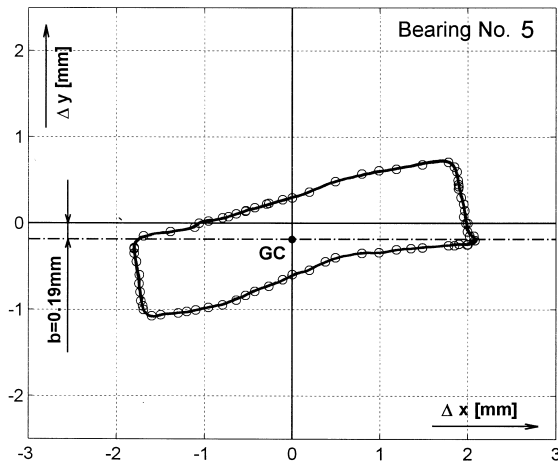


Fig. 1. The map of tolerable misalignment of the bearing No. 5 with regard to load criterion

making use of this criterion are better determined than those making use of the vibration criterion, and can be more useful as the starting data for training and developing the experimental diagnostic system.

3. METHODOLOGY OF INVESTIGATIONS

The investigations were carried out for a large power turbine set consisting of a 200 MW turbine and a generator [6, 7]. The object is a four-cylinder set, which rotors are supported on seven slider bearings. Four shaft sections are linked together by three semi-stiff couplings. Of special interest is the distribution of the pair of bearings between the LP turbine cylinder and the generator. The two bearings are located close to each other and share a common support, which is a reason of great mutual interdependence of the two bearings. The investigations were limited to misalignment of bearings No. 5 and 6 to any direction. These bearings are heavily loaded and are the source of relatively big maintenance troubles [6, 7].

The calculations were carried out using a series of the MESWIR system computer codes for non-linear analysis of rotors supported in slider bearings. The system was developed and is used in IFFM, Gdansk [6, 10]. In the MESWIR system bearings are regarded as hydrodynamic slider bearings of finite length. Oil film diathermic model is applied in the model. Bearing characteristics are calculated by solving the two-dimensional Reynolds equation applying Reynolds boundary conditions. Rotors are modeled by means of finite element method [10]. Scripts written in the MATLAB package were used to supervise the calculations. The calculations were carried on until a required convergence of iterative process was obtained i.e. until bearing misalignments attained the limiting values determined by the assumed criterion of permissible bearing load with sufficient accuracy.

An analysis concept of misalignment defect o bearings have been illustrated and explained in particulars in previous papers [6, 7], therefore here is

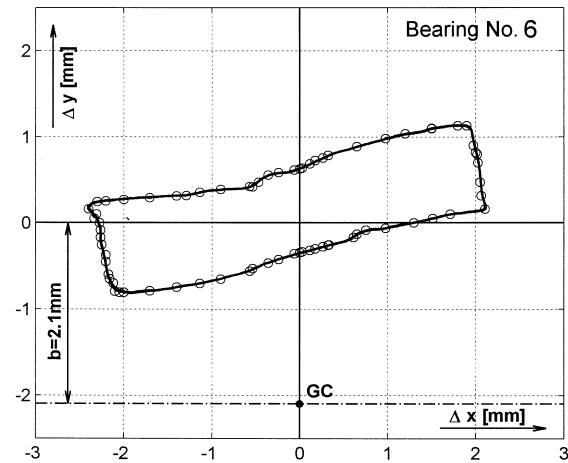


Fig. 2. The map of tolerable misalignment of the bearing No. 6 with regard to load criterion

shortly mentioned. The task was carried out by introducing a bearing misalignment defect into a basic model of the machine and observing effects. Dislocations of the bearings, representing the defects of the machine, were added to the basic bearing dislocation values resulting from rotor catenary in the basic case. The basic case was represented by a defect-free numerical model of the turbine set, in which all bearing centres were located along the design catenary line.

It was assumed that qualification criterion that allows considering the status of the turbine set acceptable is simultaneous fulfilment for all seven turbine set bearings of the bearing load condition. The average pressure on the bushing surface must be lower than 2 MPa. The limiting pressure on a bush surface was assumed after turbine industry for designing large slider bearings.

4. RESULTS

The Tab. 1 contains calculated values of maximum tolerable misalignments of analyzed bearings No. 5 and No. 6 purely in horizontal and vertical direction: to the right, to the left, upward and downward. The number of bearing in which the load was exceeded as well as direction of load being exceeded is pointed out in the last column. The table allows comparison of areas of purely vertical and purely horizontal tolerable misalignments of bearings No. 5 and No. 6. It is seen in the last column of the table that permissible values of bearing load can be surpassed in any bearing, not only this one on which a defect was imposed.

Ranges of tolerable displacements of the bearings No. 5 and No. 6 to any direction are graphically presented as maps of tolerable displacements in Fig. 1 and 2. The origin of the coordinate system represents the initial location of the centre of bearing bush, with respect to which the bearing was dislocated. A dotted-broken line represents an axis crossing the geodesic line (horizontal straight reference line) at the point marked GC. It results from Fig. 1 and 2 that tolerable

misalignment areas calculated with regard to the load criterion are confined by almost regular closed loops. The graphs for the two bearings theoretically should be similar, as the bearings have almost the same construction and the same direction of rotation. But they become similar in mathematical meaning when one of them is turned 180 deg around the coordinate system origin.

Tab. 1. Calculated values of maximum tolerable misalignments of bearings

No. of bearing	Direction of displac.	Displacement delta [mm]	Load exceed in bearing
5	right →	$\Delta x = 1.9811$	bear. 5, dir. x
	left ←	$\Delta x = -1.0605$	bear. 5, dir. y
	up ↑	$\Delta y = 0.2963$	bear. 5, dir. y
	down ↓	$\Delta y = -0.5949$	bear. 6, dir. y
6	right →	$\Delta x = 1.2994$	bear. 5, dir. y
	left ←	$\Delta x = -2.2720$	bear. 5, dir. x
	up ↑	$\Delta y = 0.6290$	bear. 6, dir. y
	down ↓	$\Delta y = -0.3435$	bear. 5, dir. y

The observed equality of ranges of allowable bearing displacements of the two neighbouring bearings to two opposite sides may lead to the conclusion that these bearings impose load to each other because bearing No. 5 is located too high with respect to the bearing No. 6. Moving bearing No. 5 down by about 0.2mm or moving bearing No. 6 to opposite direction by the same value loads one of the bearing and lightens the second bearing at the same time. This would result in higher vertical symmetry of the diagrams. Similarly, moving bearing No. 5 to the right by about 0.5 mm or moving bearing No. 6 to the left by the same value also results in higher horizontal symmetry of the diagrams. Symmetry of the diagrams cause wider tolerance of the machine to random, always possible, emergency displacements of these bearings.

Very important part of the results of the work, impossible for paper presentation is set of parameters expressing the static and dynamic state of the machine corresponding to particular bearing misalignment defects. The sets of parameters containing information about machine defect and its status can be included into the knowledge base of the turbine set diagnostic system, which is built-up.

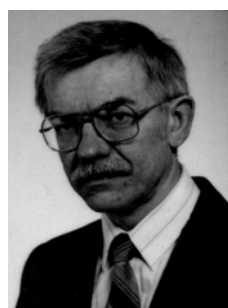
5. SUMMARY

1. The areas of tolerable misalignments of bearings delimited with regard to load criterion reveal approximate central symmetry.
2. For the pair of bearings located close to each other, one bearing reveals the range of tolerable misalignments in a given direction close to that revealed by other bearing in opposite direction.
3. The mirror reflected shapes of the area of tolerable misalignments of the two neighbouring bearings suggests, that the bearings impose the load reciprocally one to another.

4. The fact, that the bearings load themselves reciprocally suggests that the constructional rotor line is not optimal from viewpoint of machine's resistance to bearing misalignments. It can be corrected basing on results of this work.

REFERENCES

- [1] Vance J. M.: *Rotor dynamics of Turbomachinery*, 1985, A Wiley – Interscience Publ., New York.
- [2] Rao J. S.: *Vibratory Condition Monitoring of Machines*, 2000, Narosa Publishing House, New Delhi, Chennai, Mumboi, Calcuta, London.
- [3] Rybczynski J.: *Optimization of Rotor Critical Speeds by Change of Features of Machine's Bearings*, Proceedings of ASME Turbo Expo 2004, June 14–17, Vienna, No. GT2004-54291.
- [4] Rao J. S., Sreenivas R.: *Dynamic Analysis of Misaligned Rotor Systems*, Proceedings of bVETOMAC-1 Conferene, Oct. 25-27, 2000, Bangalore, India.
- [5] Sekhar A. S., Prabhu B. S.: *Effects of Coupling Misalignment on Vibrations of Rotating Machinery*, Journal of Sound and Vibration, Vol. 185, No. 4, 1995, pp. 655-671.
- [6] Rybczynski J.: *Analysis of Acceptable Nonalignment of Bearings of Large Power Turboset*, Diagnostyka, Vol. 30/2 (2004), p. 97.
- [7] Rybczynski J.: *Acceptable Dislocation of Bearings of the Turbine Set Considering Permissible Vibration and Load of the Bearings*, Key Engineering Materials, Vols. 293-294, Sept. 2005, pp. 433-440.
- [8] Cholewa W. White M. F.: *Inverse Modelling in Rotordynamics for Identification of Imbalance Distribution*, Machine Vibration, (2), 1993.
- [9] *Modelling and Diagnosing of mechanical, aerodynamic and magnetic interactions in power turbo-sets*. Multi-author work, ed. J. Kicinski, Wyd. IMP PAN, Gdansk, 2005.
- [10] Kicinski J., Drozdowski R., Materny P.: *The Non-linear Analysis of the Effect of Support Construction Properties on the Dynamic Properties of Multi Support Rotor Systems*, Journal of Sound and Vibration, 1997, 206(4).



Józef RYBCZYŃSKI, graduated from Gdansk University of Technology at Mechanical Engineering Faculty in 1971. PhD in mechanics. Employed since 1972 at Institute of Fluid Flow Machinery, PAScis in Gdansk in Mechanics of Machines Department. Area of interest: dynamics of

rotors supported in slider bearings, journal bearings, machine diagnostics, experimental investigations into the rotor and foundation vibration, measurement techniques. Member of Polish Assn. Diagnostic Engineering, Machine Bldg. Com. PAScis and ASME.

MULTIOBJECTIVE FUZZY APPROACH TO THE VEHICLE ROUTING PROBLEM WITH TIME WINDOWS

Przemysław WESOŁEK*, Marek KUBIAK*

* Poznań University of Technology, Institute of Computing Science
Piotrowo Street 2, 60-965 Poznań, Poland,
e-mail: {Przemyslaw.Wesolek, Marek.Kubiak}@cs.put.poznan.pl

Summary

The paper presents a model of the vehicle routing problem with flexible (fuzzy) constraints. This kind of model allows a decision maker to explore a set of alternatives with diverse cost and constraint satisfaction levels. The model is tested on well-known instances of the vehicle routing problem with time windows adjusted to the fuzzy case. They are solved by a multiobjective Pareto Memetic Algorithm. The obtained results indicate that the introduction of fuzzy constraints leads to exploration of new alternatives which may be interesting to a decision maker.

Keywords: fuzzy vehicle routing, flexible constraints, multiobjective optimization.

WIELOKRYTERIALNY PROBLEM PLANOWANIA TRAS Z OKNAMI CZASOWYMI W WERSJI ROZMYTEJ

Streszczenie

Artykuł prezentuje model problemu planowania tras z elastycznymi (rozmytymi) ograniczeniami. Model takiego rodzaju pozwala decydentowi na wybór rozwiązania spośród zbioru alternatyw ze zróżnicowanym kosztem i stopniem spełnienia ograniczeń. Ten model został przetestowany na klasycznym zestawie instancji problemu planowania tras z oknami czasowymi dostosowanych do przypadku rozmytych ograniczeń. Rozwiązania są uzyskiwane przez użycie wielokryterialnego algorytmu memetycznego. Uzyskane wyniki wskazują na to, że wprowadzenie elastycznych ograniczeń prowadzi do odkrycia rozwiązań, które mogą być interesujące z punktu widzenia decydenta

Słowa kluczowe: problem planowania tras, elastyczne ograniczenia, optymalizacja wielokryterialna.

1. INTRODUCTION

Models of the Vehicle Routing Problem (VRP) proposed in literature mainly define it as single objective one (with minimization of cost of a whole routing plan), with hard constraints (e.g. related to capacity of vehicles or time of delivery) and deterministic parameters (like travel time). However, decision makers (DMs) involved in the planning process (e.g. the owner or executive staff of a transportation company) usually know that there is more variability connected with parameters of such problems: travel times vary depending on circumstances, times of delivery may be renegotiated with customers. Therefore, even optimal solutions of the mentioned models may not be enough for DMs to make up their minds, because there are scarcely any options (variants) generated from these models which could be a basis for 'what-if' analysis. Moreover, the optimal or nearly optimal solution might be regarded by DMs as a 'magic' one. Even if there are multiple solutions generated, they are related only to the defined objective function, constraints and parameters, so they do not

reveal new possibilities to DMs, which could arise with slight changes to e.g. travel times and time windows.

Some models of the VRP try to address the issues of changes to parameters of the problem by introducing a level of constraint satisfaction. The model developed by Zimmermann [11] allows to define flexible constraints with fuzzy values. The original objective is converted into an additional constraint with its upper limit being a fuzzy number called aspiration level. The new objective is to maximize satisfaction of all constraints.

This model has major drawbacks: a decision maker may not understand the meaning of satisfaction value (it can be either a not fully satisfied constraint or not fully satisfied aspiration level). Also, the aspiration level has to be set in advance, forcing a DM to make a guess (potentially incorrect).

In [10] the authors introduced soft time windows, the violations of which result in a penalty. Value of the penalty depends on the type of customer and is a component of the objective function.

The aforementioned models do not provide a DM with a set of alternatives.

This drawback is not present in the model described in [8], which is a multiobjective one and allows constraints to be violated. One of the objectives is the weighted number of not fulfilled constraints. However, the extent to which they are violated is not measured. Due to this fact, the amount of violation may be either negligible or huge, thus unusable for a DM.

Another aspect of real-world situations which is not addressed by these models is the uncertainty of travel times. There are models which try to handle this aspect, but they are usually limited to stochastic approaches, rich in strong assumptions (e.g. independency of travel times or the type of their distribution) [3].

This paper describes a new model for the VRP with time windows (VRPTW) [9]. On one hand, the model allows for generation of multiple alternatives and, on the other hand, provides means of measuring the level of constraints satisfaction. Additionally, in contrast to the stochastic approach, the fuzzy model of travel times is proposed.

The purpose of this paper is to demonstrate that introduction of flexibility into constraints does not decrease the quality of results. Another goal is to show that the acceptance of small violations of constraints can lead to solutions with lower cost, which may be valuable to DMs.

2. FUZZY VEHICLE ROUTING PROBLEM WITH TIME WINDOWS (FVRPTW)

Let N be the set of customers. Each customer i has some demand (the amount of goods) and must be served within the time window $[e_i, l_i]$. Customers are visited by vehicles from the set K , each having the capacity q . Some time is required to travel between a pair of customers. All vehicles start and end their travel in the same depot. The objective is to find a set of routes (i.e. the sequences of customers) with minimal cost, such that:

- each customer is placed on exactly one route
- the time at which the customer i is visited is between e_i and l_i
- the sum of customers' demands on each route is not greater than q

The above model assumes hard time windows. To allow small violation of time constraints, the time windows must be presented in a different way. Previous works ([1], [4]) show that fuzzy intervals are a good model to achieve this requirement.

For the customer i his time window is fuzzified (see Fig. 1 for an example). Only the right part of the time window (i.e. the latest service time, l_i) is fuzzified, because the left part plays role only when the vehicle arrival takes place before the e_i and has nothing to do with the customer's satisfaction.

Formally, soft time window of customer i is represented by a fuzzy set of service times satisfying

the customer, with its membership function $\mu_i(t)$ equal to the level of customer's satisfaction when the service starts at time t_i .

The original time constraints are no longer present in the flexible model. Instead, the satisfaction levels assigned to time windows are part of solutions' quality. They constitute an additional objective: to minimize the global satisfaction level (i.e. the satisfaction level of least satisfied constraint). This makes the problem multiobjective.

As mentioned in previous sections, the assumption that travel times are fixed numbers known a priori is very strong. Therefore, fuzzy travel times are introduced. Each travel time T_i is modeled as a 3-point fuzzy number. This representation requires least assumptions and is easy to build with a human expert, e.g. by asking him questions in the form: "What is the expected, smallest and greatest travel time between X and Y ".

The arithmetic of fuzzy values is simple, compared to the one of stochastic variables, and the concept of satisfying flexible constraints with uncertain values is often used ([1]). An example of fuzzy time value satisfying flexible constraint is shown on Fig. 1. The concepts used here have roots in the possibility theory [2].

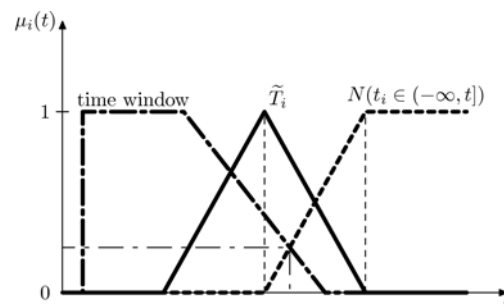


Fig. 1. Fulfillment of a flexible constraint by uncertain value

3. THE PROPOSED ALGORITHM

The model described above is solved by a multiobjective metaheuristic algorithm called Pareto Memetic Algorithm (PMA) [5]. Its goal is to find a set of approximately Pareto-optimal solutions. The algorithm uses a mechanism of scalarizing functions to drive the search in the direction of the Pareto-optimal set (Pareto set, PS). In this work the achievement scalarizing functions are used [5].

The main loop of the PMA, which starts right after generation of initial solutions by a randomized heuristic, is presented in Fig. .

The general idea of this algorithm is to generate new approximately Pareto-optimal solutions by recombination of already found solutions which are good with respect to a randomly chosen scalarizing function. The offspring resulting from this recombination is then improved by a local search heuristic optimizing the same scalarizing function. The rationale behind this procedure is based on the

assumption that recombination of good solutions is likely to produce good starting points for local search.

repeat

Draw at random a weight vector λ

From the current set CS of solutions draw with uniform probability a sample T of solutions

Recombine the best and the second best solution on $s(\mathbf{z}, \dots, \lambda)$ from T obtaining $\mathbf{x}_1$

Apply a local search heuristic optimizing $s(\mathbf{z}, \dots, \lambda)$ to $\mathbf{x}_1$ obtaining $\mathbf{x}_1'$

If $\mathbf{x}_1'$ is better on $s(\mathbf{z}, \dots, \lambda)$ then the second best solution in T **then**

Add $\mathbf{x}_1'$ to the current set CS of solutions

Update the set of approximately Pareto-optimal solutions with $\mathbf{x}_1'$

until the stopping criterion is met

Fig. 2. Main loop of the PMA. $s(\mathbf{z}, \dots, \lambda)$ stands for a scalarizing function value with weight vector λ for a solution with objective values $\mathbf{z}$

The Pareto Memetic Algorithm was adapted to the FVRPTW by means of problem-specific randomized heuristic procedure, local search and recombination operators. There is no explicit mutation operator defined.

The randomized heuristic procedure generates solutions using the insertion method I1 described in [9], with parameter values drawn randomly with uniform probability. This heuristic was chosen because it generates good quality solutions. During construction of routes I1 requires that all hard constraints are met, so it had to be adjusted to the case of flexible time windows. If time constraints were removed totally, it could result in all initial solutions having satisfaction equal to zero, especially in problems with very tight time windows. Thus, the modified I1 heuristic requires that constraints are not fully violated (they have non-zero satisfaction level).

The local search uses steepest approach to find the best neighbor of the current solution. The operators used to construct the neighborhood are moving a customer from one route to any position on another one or switching two customers between routes.

Two recombination operators were implemented: the route-based crossover (RBX) [7] and the common-edges-preserving crossover (CEPX) [6]. RBX generates an offspring by first copying a randomly chosen set of routes from one parent. Then, the offspring is completed with routes from the other parent, omitting the customers already inserted when necessary. Consequently, RBX contains a random component. CEPX is, on the contrary, a fully deterministic crossover. Routes of

its offspring are composed only of edges which are common to both parents.

The PMA was implemented in Java.

4. COMPUTATIONAL EXPERIMENT

In order to generate approximations of Pareto sets and to assess the efficiency of the proposed algorithm (especially of the two crossover operators) a computational experiment was designed. Well-known Solomon's instances of VRPTW [9] were used. From the whole set of Solomon's instances we selected 3 groups (called R, C, RC in short) with the largest number of customers (100); 56 instances in total. All these instances were augmented with information about fuzzy time windows and travel times. Each time window was enlarged by 10% of its original width, with customer's satisfaction decreasing linearly to zero. Travel times were fuzzyfied by 5% in either direction (compare Fig. 1).

Two versions of the PMA were run: one with RBX only, the other with CEPX. They were executed 15 times for each instance, 420 seconds each run, on PCs with Intel Pentium 4 processor 3.2 GHz and 1GB RAM, running Windows XP.

Evaluation of results of a multiobjective algorithm requires that specialized measures of quality of approximations of PS are employed. In this paper two such measures are used: coverage $C(A,B)$ of approximation A of PS by approximation B [12] and the average best value of the weighted Tchebycheff scalarizing function $R(A)$ for one approximation A [5].

$C(A,B)$ is a percentage of solutions in A which are dominated (covered) by some solutions from B. It is normalized and not symmetric. $C(A,B)=1$ indicates that A is completely dominated by B. If $C(A,B)=0$ then the approximations A and B of PS are incomparable.

$R(A)$, shows a quality of an approximation A with respect to a reference point, usually being an ideal point. This measure is nonnegative and not normalized. The closer it is to 0, the better the approximation A is (being generally closer to the ideal point).

Table 1 presents values of the coverage measure C for the two employed algorithms. Values of the R measure are shown in Fig. 3.

Table 1. Average values of the C measure between approximations generated by two algorithms in groups of instances

Group of instances	C(RBX,CEPX)	C(CEPX,RBX)
C	16,7%	63,2%
R	0,8%	52,9%
RC	0,8%	53,0%
All	5,6%	56,0%

As can be seen from the Table 1 and Fig. 4, the solutions obtained from the computations using

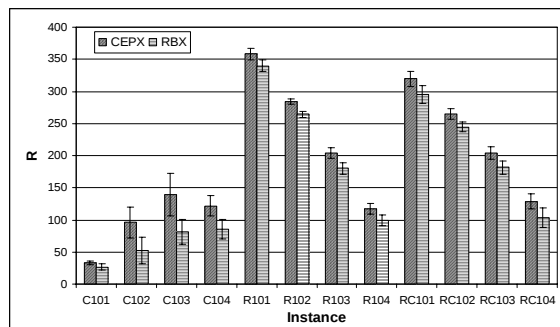


Fig. 3. Values of the R measure for approximations of PS generated by different algorithms

CEPX operator are usually dominated by the solutions resulting from computations using RBX. It may be due to the fact that in the case of CEPX the whole procedure of recombination and local search is deterministic. Thus, the only random component of the algorithm determining the structure of solutions is randomized initial heuristic, which might be not enough to achieve diversified solutions.

One approximation of a Pareto front which we obtained from the experiment is shown in Fig. 4. The horizontal line represents the cost of the optimal solution to the presented problem.

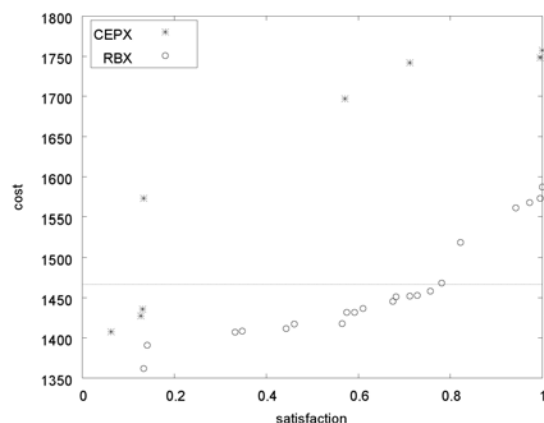


Fig. 4. An example of a Pareto front approximation for instance R101

In this figure it can be seen that the cost of a solution can be greatly lowered in exchange for a satisfaction decrease. This fact gives a DM a chance to choose between two interesting, but highly different solutions: one with reasonable cost and fully satisfied constraints, the other with smaller satisfaction, but less expensive. The knowledge about such solutions can be the basis for renegotiation of time windows with customers.

5. CONCLUSIONS AND FURTHER RESEARCH

The constructed FVRPTW is valuable if a DM should have the possibility to choose a solution from a set of good and diversified ones. Using the proposed approach, a DM is able to choose time

windows, which, when altered, can result in much less expensive routes.

The comparison of crossover operators shows that there is a need for randomness in the PMA: the use of randomized RBX operator leads to better results than the use of deterministic CEPX.

The main direction for further research is the improvement of heuristics and operators, which are suitable to the model of flexible constraints.

REFERENCES

- [1] Dubois D., Fargier H., Prade H.: *Fuzzy constraints in job-shop scheduling*, Journal of Intelligent Manufacturing, 1995, Vol. 6.
- [2] Dubois D., Prade H.: *Possibility Theory*. Plenum Press, New York, 1988.
- [3] Hadjiconstantinou E., Roberts D.: *Routing under uncertainty: an application in the scheduling of field service engineers*. In: Toth P., Vigo D. (eds.): *The Vehicle Routing Problem*, 2002, Philadelphia, SIAM.
- [4] Hapke M., Wesołek P.: *Handling imprecision and flexible constraints in vehicle routing problems: fuzzy approach*. Technical report RA-005/2003, Poznań University of Technology, 2003.
- [5] Jaskiewicz A.: *On the computational efficiency of multiple objective metaheuristics. The knapsack problem case study*. European Journal of Operational Research, 2004, Vol. 158, pp. 418-433.
- [6] Kubiak M.: *Systematic construction of recombination operators for the vehicle routing problem*. Foundations of Computing and Decision Sciences, 2004, Vol. 29, No. 3.
- [7] Potvin J.-Y., Bengio S.: *The vehicle routing problem with time windows part ii: genetic search*. INFORMS Journal of Computing, 1996, Vol. 8, No. 2.
- [8] Rahoual M., et al.: *Multicriteria genetic algorithm for the vehicle routing problem with time windows*. Metaheuristics International Conference, Porto, Portugal, 2001, pp. 527-532.
- [9] Solomon M. M.: *Algorithms for the vehicle routing and scheduling problems with time window constraints*. Operations Research, 1987, Vol. 35, No. 2, pp. 254-265.
- [10] Taillard É., Badeau P., Gendreau M., Guertin F., Potvin J.-Y.: *A tabu search heuristic for the vehicle routing problem with soft time windows*. Transportation Science, 1997, Vol. 31, No. 2, pp. 170-186.
- [11] Zimmermann H.-J.: *Description and optimization of fuzzy systems*. International Journal of General Systems, 1976, Vol. 2, pp. 202-216.
- [12] Zitzler E., Thiele L.: *Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach*. IEEE Transactions on Evolutionary Computation, 1999, Vol. 3, No. 4, pp. 257-271.

COMPARING METHODS FOR IMPLEMENTING VBA IN CATIA FOR ELABORATING KBE PACKAGE

Wojciech SKARKA*, Łukasz ZAJĄC**

* Silesian University of Technology, Department of Fundamentals of Machinery Design
Konarskiego 18a, 44-100 Gliwice, Poland, e-mail: wojciech.skarka@polsl.pl

** Silesian University of Technology, Department of Fundamentals of Machinery Design
Konarskiego 18a, 44-100 Gliwice, Poland, e-mail: ukasz_zajac@poczta.onet.pl

Summary

The paper present comparison of methods for implementing VBA for creation of KBE application in CATIA system. The usage of VBA is one of the methods of automation of Generative Model creation in CATIA system. Creation of KBE Package forms one of the last phases of the whole process of KBE creation which consists of capturing knowledge necessary for KBE system, formal representation of knowledge, and transferring and implementing knowledge in KBE package. While creating KBE application, Generative Model is being formed and for the purpose of formalization of creation process of that model CATIA Knowledgeware tools are used. In order to have automation of that process it is necessary to ensure programming techniques, including VBA. The suggested 4 methods of implementing VBA for automation of Generative Model creation guarantee greater stability and repeatability of the model.

Keywords: Knowledge Based Engineering, Generative Model, CATIA.

PORÓWNANIE METOD PROGRAMOWYCH Z ZASTOSOWANIEM VBA W SYSTEMIE CATIA NA ETAPIE TWORZENIE PAKIETU SYSTEMU KBE

Streszczenie

W artykule przedstawiono porównanie metod zastosowania VBA w końcowym etapie tworzenia aplikacji projektowej opartej na wiedzy (Knowledge Based Engineering - KBE) w systemie CATIA. Zastosowanie VBA jest jedną z metod automatyzacji tworzenia modelu autogenerującego w systemie CATIA, który jest głównym elementem aplikacji KBE. Tworzenie pakietu aplikacji KBE jest jednym z ostatnich etapów całego procesu tworzenia systemu KBE na który składają się akwizycja wiedzy koniecznej dla systemu KBE, formalna reprezentacja wiedzy i transfer i integracja wiedzy w aplikacji KBE. W fazie tworzenia aplikacji KBE tworzony jest model autogenerujący i dla formalizacji procesu tworzenia tego modelu wykorzystywane są narzędzia Knowledgeware systemu CATIA. Automatyzacja tego procesu wymaga zapewnienia technik programowych w tym zastosowania VBA. Zaproponowane 4 metody implementacji VBA do automatyzacji procesu tworzenia modelu autogenerującego pozwalają na zwiększenie stabilności i powtarzalności tego modelu.

Słowa kluczowe: Projektowanie oparte na wiedzy, model autogenerujący, CATIA.

INTRODUCTION

In designing processes of machines and devices computer aiding tools are commonly used and we can hardly imagine designing without using CAD, CAM or CAE tools. However, at the present stage of development it is not possible to talk about common usage of KBE (Knowledge Based Engineering) [2] systems in solving designing task. Two groups of such systems can be differentiated: the first one covers newly created systems whereas the latter one systems which are natural extension of already existing tools (e.g. CAD). Newly created KBE [11] systems, despite their positive effects of usage, are not commonly used. In the latter group

the existence of modules in the existing CAD tools can be noticed, which enable knowledge implementation in these systems [3]. For the purpose of creating newly formed KBE systems, ordering methodology is usually used which arranges the process e.g. CommonKADS [4], MOKA [10]. In the process of development of KBE system the following phases are commonly differentiated [8], [10]:

- Identify,
- Justify,
- Capture,
- Formalize,
- Package,
- Activate.

The most important phases from the point of view of the creation of the very KBE tool are: *Capture*, *Formalize* and *Package* phase. In these phases the approach similar to suggested in MOKA methodologies, which are oriented to a given CAD tool can be used.

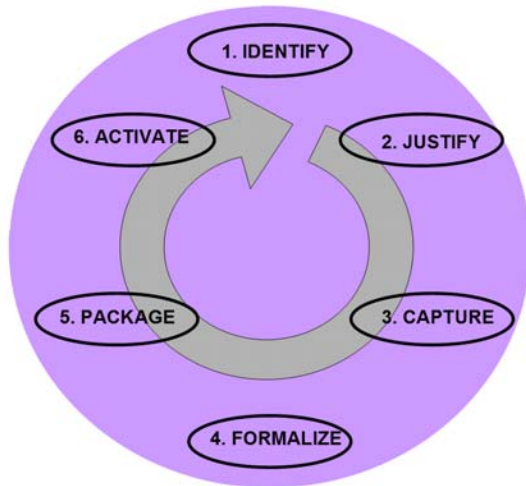


Fig. 1. General KBE lifecycle [8], [10]

For CATIA system such a methodology has been elaborated [8]. It is based on the model of MOKA knowledge base and uses Informal Model from that methodology in the process of knowledge capture. However, in that case MOKA metamodel has been adopted to use with CATIA system. It consists of two autonomous parts, with one being independent from the used CAD system and the other allows orienting to a given CAD system. In this way it is easy to adjust knowledge base which is being created to 'cooperation' with other CAD systems. During the process of Package, which is the creation of KBE tools, CATIA system tools are used for the formation of Generative Model [5], [6], [7]. This model can be created directly either by means of using Knowledgeware tools of CATIA system or by automation of the process by means of usage of programming tools e.g. VBA. It is one of the most time consuming tasks of that process and at the same time there is a lack of experience in that field. A big disadvantage is the choice of a given method of Generative Model creation by means of VBA, since there are at least a few hypothetical forms of that model to choose from.

Decisive factor which determines the choice is the level of time consumption as well as stability in using the model which results in proper generation of geometric model in the range of given sets of input data for designing process. In the further part of the paper four such methods have been suggested and compared which allow making *Package* process of KBE system creation easier with the aid of CATIA tools.

1. VB in VBScript

By means of using *Knowledge Advisor (KA)* modul in CATIA system it is possible to integrate *Visual Basic* code and CATIA *Part* file. *Visual Basic* code can be included in Macro as VBScript and in this case it can be used in *Power Copy* function. Script is activated by *Reaction* function that activates with the change of parameters (Fig. 2).

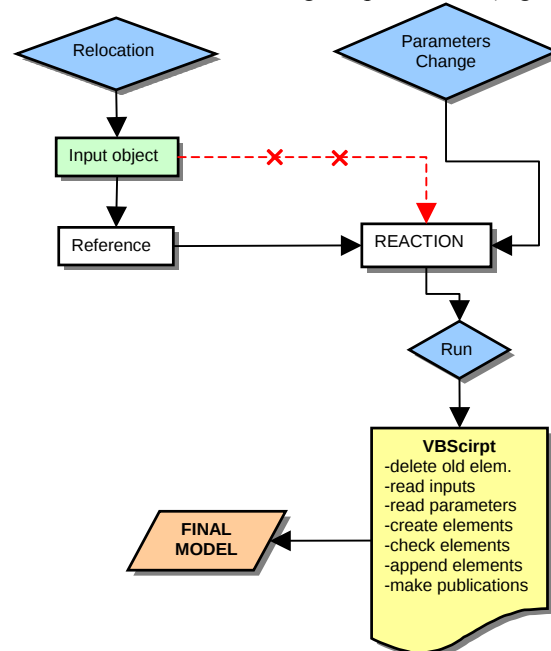


Fig. 2. VB in VBScript

Two CATIA *Open Bodies* have to be created before transformation. First *Open Body* should consist of references to input elements because it is necessary to have a set of unchanging names of references in VBScript. *Reactions* also have to be connected to those references. Connection of *reaction* directly with *inputs* cause errors while *inputs* are moved (red line crossed out in Fig. 2). Unfortunately, choosing such a kind of referencing makes it impossible to connect elements by *Formula* tool. Formulas are updated at the final stage as *Power Copy* is added and it generates errors. Table 4.1 shows examples how to create references to *inputs*. Second *Open Body* will be composed of products of VBScript.

Table 1: Example of references to *input* elements

Input Elements	References to input elements
Points	Points defined by coordinates
Directions	Lines defined by a point and a direction
Profiles	Parallel Curve with offset = 0

After transformation of basic elements (document, part, hybrid shapes etc.) VB code deletes all *Hybrid Shapes* in second *Open Body*.

Reading *input* elements and value of parameters should be made by searching method. The reason is that Power Copy adds some words to element names when it is activated. Next new elements are created with checking correctness of model. Only necessary features are appended to further work in CATIA tree. In conformity with the rules how to create KBE models, final features also have to be published.

Figure 3 shows example of profile created in this method. User through manipulating a series of parameters can get new model. VBScript checks that center point is in profile range. To check location of center point virtual line between this point and COG (Center Of Gravity) is created. Then intersection of line and profile is made. If intersection exists (error = 0) than point will be out of range and user is informed about that.

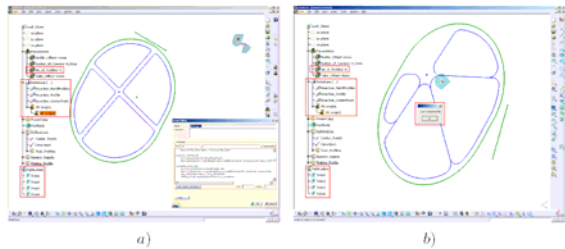


Fig. 3. Example of VB in VBScript usage

If one of the profiles is impracticable, user is informed about that and next profiles are created. To check correctness of profile *On Error Resume Next* function is used (see Fig. 3.b)

2. VBA as an environment to build complex model

This method is based on the usage of *SelectElemet2* function and it does not require the usage of *Power Copy*. Pointing at elements on CATIA tree allows creation of complete model in VBA. Interface is created in VBA, which allows user to point clearly at *input* elements and change value of parameters. Errors are identified by *check update error*. At first program checks if special *Open Body* exists, if it does not exist then *Open Body* will be created to include model, if it does exist then all *Hybrid Shapes* in it will be deleted. After that VBA program prepares to show window, values of parameters (as it was mentioned before) are read and a list of input elements is read. List of elements is created to enable showing what was pointed before and to enable *Cancel* option. Next user makes changes and decides whether to continue or to cancel work. If everything is satisfactory then a new model will be created or if *Cancel* option will be chosen then VBA will recall old parameters and create old model again. Figure 5 shows an example of this method. When user indicates tube profile

which is not of circular shape he must also indicate a point from which place in a range of profile one should start guide line. When it is impossible to create tube only guide line will be created to show user what is wrong.

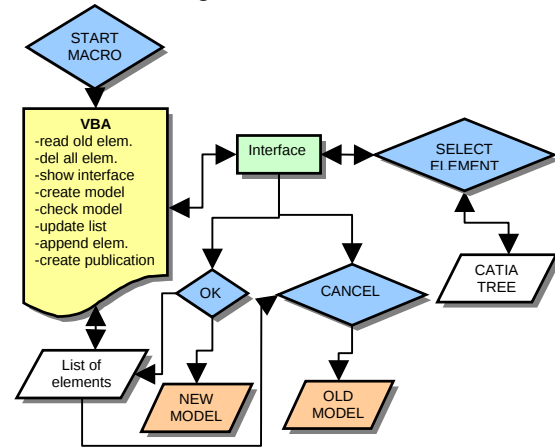


Fig. 4. VBA as an environment to build complex model

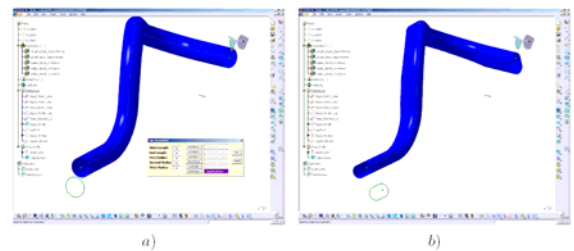


Fig. 5. Example of use of VBA environment to build complex model

3. VBA in conjunction with Power Copy

This method forms a combination of two previous methods. The interface is realized by VBA and management of model elements by *Power Copy* (Fig. 6). *Power Copy* has sets of correctly named parameters and *Open Bodies*. When *Power Copy* is inserted user points his input elements. In this case VBA does not have to read inputs when macro is started, and thanks to that interface becomes simpler. This method of transformation also includes creation of different design option in one VBA macro. Thanks to that a model has clear structure and is more stable as compared with model created in CATIA (see Fig. 7). At the beginning VBA macro checks if *Power Copy* exists, if it does not exist then macro will stop and user will have to insert *Power Copy*. Next macro reads all parameter values and creates copy of it to allow canceling work and recalling old version. After that old model is deleted and a user has to make a choice of design option. When selection is accepted then a proper set of parameters is showed. User can change values of parameters and next macro checks if it is possible to create model. Finally model is appended on a CATIA tree. Flange model (Fig. 7) is an example

of implementation of this method. Macro can generate different type of flanges (stamped flange, stamped flange with chamfer and deep drawn flange). Each flange can have form of two to four screw holes. Additionally there is a possibility to generate flanges for tubes of other cross-section than circle. To sum up, macro allows creating twelve different types of flange.

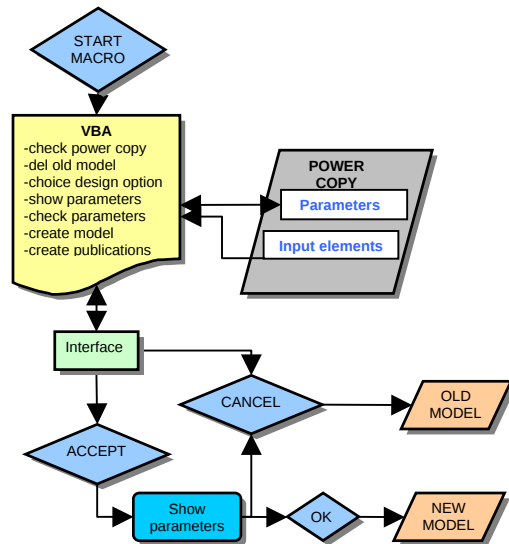


Fig. 6. VBA in conjunction with Power Copy

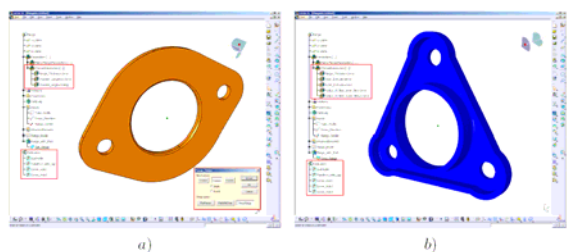


Fig. 7. Example of model build through Power Copy with VBA

4. PARSING ALGORITHM AND AUTOMATIC MODEL TRANSFORMATION

This method was elaborated to make automatic generation of VBA code from CATIA model possible. Basic idea was to use parsing algorithm to autotranslation models. First step of this idea is to read hierarchy of CATIA elements. This step is a tokenization process and it gives information about type of features of models. Every read element will be included in table with information about its parents. Reading the table by macro allows rebuilding a hierarchy of CATIA tree. Hierarchy must be read from bottom to the top then features that have no parents must be read first. While hierarchy is being read parser has information about

type of each element and its parents and generates VB code to create this element.

The problem is how to build grammar as a piece of VB code. In CATIA model, especially in surface model, we have a lot of features and each has many creation options. Unfortunately every option of feature has different parents and so it is impossible to create all grammar by hand. In order to solve this problem it is necessary to use specialized libraries of CATIA system and application development software oriented to CATIA [1]. The above mentioned problems resulted in the fact that the method has been abandoned despite potentially promising possibilities.

5. SUMMARY

The paper presents comparison of programming methods including the usage of VBA at the phase of package of KBE creation. Three methods have been presented and carefully checked and verified, based on the examples of subassemblies of automotive industry. Additionally, the concept of the fourth method and its preliminary tests has been discussed. The presented methods have significant importance as far as automation of *Generative Model* is concerned and this model is the main object of the last three phases of methodology of KBE creation with the aid of CAD tools. These are the following phases: *capture*, *formalize* and *package*. The use of the existing CAD/CAM tool in the process causes considerable problems because it is necessary to maintain universality of the paradigm of conduct and, at the same time, its possibilities of specialization for the purpose of a given CAD/CAM tool [9]. The presented programming methods ensure this type of specialization for the purposes of CATIA system. All of the presented methods allow automation of *Generative Model* creation from the sources of knowledge gathered and recorded in the process of *capture* and *formalize* of the methodology. They constitute extension and automation in *Knowledgeware* function operation of CATIA system. So far independent usage of *Knowledgeware* function in practice has made it impossible for system creation of *Generative Model* [5, 6] due to the lack of automation aiding of its creation.

Table 2: Comparison of methods for implementing VBA for creation of KBE application in CATIA

Method	VB i VBScript	VBA as environment to build complex model	VBA with collaboration with Power Copy	Parsing algorithm and automatic model transformation
Pros	• Fast transformation • Control of model stability • Autogenerating of code by indicating element in CATIA tree (insert object resolution) • VBScript constitute an integral part of Power Copy	• Whole model is created in VBA • Stability control in VBA • It doesn't need Power Copy • User indicate input elements in CATIA	• Whole model is created in VBA • Stability control in VBA • Simply interface • Clear structure of model • Different design option in one VB macro	• Clear concept of the method • Universality • Easy possibility of development • Possibility of modularization
Cons	• Reaction causing errors when Power Copy is put and must be deactivated in Power Copy when it is created • Making references to input elements • Script Editor poorly supports writing code	• Interface complicated in use • Creation of interface takes much more time than translating model form CATIA to VBA	• Large VB code to include all design options • VBA must cooperate with Power Copy	• Problem to build grammar of CATIA model • Difficult to convert CATIA model to VBA • It has not been fully tested yet

REFERENCES

- [1] CAA RADE. IBM Software. Application Development (CAA RADE) Discipline. www-306.ibm.com/software/applications/plm/caa/disciplines/caa/
- [2] Hopgood A. A: *Intelligent Systems for Engineers and Scientists*. CRC Press LLC 2001
- [3] Sandberg M.: *Knowledge Based Engineering – In Product Development. Technical Report*. Division of Computer Aided Design. Lulea University of Technology. Sweden, 2003
- [4] Schreiber G. et al.: *Knowledge Engineering and Management. The CommonKADS Methodology*. MIT Press 2000
- [5] Skarka W.: *Automation of designing process using generative models*. (in polish) Mechanik. 11/2005 s.781-782, Agencja Wydawnicza SIMP, Warszawa, 2005
- [6] Skarka W.: *Developing Generative Models through Informal Knowledge Base*. Proceedings of the Symposium on Methods of Artificial Intelligence AI-METH 2005 and the Workshop on Knowledge Acquisition in Mechanical Engineering. p. 115-116 November 16-18, 2005, Gliwice, Poland, 2005
- [7] Skarka W.: *Integrated Models as Knowledge Representattion in Designing Process*. (in polish) Mechanik. 1/2006 s.958-959, Agencja Wydawnicza SIMP, Warszawa, 2006
- [8] Skarka W.: *Knowledge Acquisition for Generative Model Construction*. Concurrent Engineering CE2006, Antibes, France September 18-22, 2006 (accepted to be published)
- [9] Skarka W., Mazurek A.: *CATIA. Modeling and designing*. Helion. Gliwice 2005
- [10] Stokes M. (ed.): *Managing Engineering Knowledge; MOKA: Methodology for Knowledge Based Engineering Applications*. Professional Engineering Publishing, London, 2001.
- [11] *The ICAD system* www.ds-kti.com/our_products/icad.shtml



Wojciech SKARKA, Ph. D. Eng. is an assistant professor at the Department of Fundamentals of Machinery Design. Since 1998 he was employed as a designer and a researcher at Mining Mechanization Centre in Gliwice. He is the constructor of many mining machines and the author of a dozen of patents. Currently he carries research on Knowledge Based Engineering and Interactive Electronic Technical Manual at Silesian University of Technology.



Lukasz ZAJĄC is a M. Sc student of Computer –Aided Design and Maintenance of Machinery at the Silesian University of Technology at Gliwice. He works as a research assistant on the KBE project in Tenneco Automotive – Henrich Gillet GmbH. His current focus is CATIA API - VBA environment and transformation strategy development.

HEURISTIC MODELING OF OBJECTS AND PROCESSES USING DYNAMIC NEURAL NETWORKS

Piotr PRZYSTAŁKA

Silesian University of Technology, Department of Fundamentals of Machinery Design
Konarskiego Street 18a, 44-100 Gliwice, Poland, e-mail: pp@polsl.pl

Summary

The methodology of heuristic modeling is one of the subjects included in the activities developed by the Department of Fundamentals of Machinery Design [4, 6]. Among all the approaches of heuristic modeling some of the most common are artificial neural networks. There are many papers and books devoted to applications of neural networks for modeling dynamic systems [1, 2, 4, 5, 6, 7]. In this paper, known approach basing on dynamic neuron model is presented (dynamic neuron with IIR filter in the activation block [2]) but some developments are introduced. Locally recurrent networks which are composed of dynamic neural units described in [2, 5, 7] are able to model behavior of complex dynamic systems. Nevertheless, they have one major disadvantage, that is, neural networks composed of these neurons are not able to represent stochastic behaviors of some objects [4,6]. By introducing the ARMAX (or ARX) system into dynamic neuron model author has received dynamic neuron unit that never behaves in the same way (it brings an artificial neuron closer and closer to the biological model). In this paper the author presents formal description of dynamic neuron unit with ARMAX system in the feedback block. There are also described a general structure of dynamic neural network composed of these neurons, two known training methods and some commonly used quality measures. At the end of the paper three examples of applications are given.

Keywords: artificial neural networks, locally recurrent neural networks, linear, non-linear and chaotic dynamic systems, quasi-Newton methods, heuristic modeling.

HEURYSTYCZNE MODELOWANIE OBIEKTÓW I PROCESÓW PRZY POMOCY DYNAMICZNYCH SIECI NEURONOWYCH

Streszczenie

Metodologia heurystycznego modelowania obiektów i procesów jest jednym z kierunków badań rozwijanym przez Katedrę Podstaw Konstrukcji Maszyn [4, 6]. Spośród wielu metod modelowania heurystycznego duże znaczenie odgrywają metody bazujące na sztucznych sieciach neuronowych. Można wyróżnić wiele ciekawych prac badawczych prowadzonych w kierunku modelowania systemów dynamicznych z zastosowaniem tego typu narzędzia [1, 2, 4, 5, 6, 7]. W artykule zaprezentowano znane podejście bazujące na dynamicznych neuronach (dynamiczny neuron z filtrem IIR w bloku aktywacyjnym [2]) z pewnymi modyfikacjami. Lokalnie rekurencyjne sieci neuronowe złożone z dynamicznych neuronów opisane w [2, 5, 7] nadają się do modelowania zachowania złożonych systemów dynamicznych. Jednakże, posiadają one jedną główną wadę tzn. nie są zdolne do reprezentowania zachowania losowego niektórych obiektów [4, 6]. Poprzez wprowadzenie systemu typu ARMAX (ARX) do modeli dynamicznych neuronów autor otrzymał dynamiczny model neuronu, który nigdy nie zachowują się w ten sam sposób (przybliżył to model sztucznego neuronu do jego biologicznego wzoru). W artykule autor prezentuje formalny opis dynamicznego neuronu z systemem typu ARMAX w bloku sprzężenia zwrotnego. Opisuje również ogólną strukturę dynamicznej sieci neuronowej złożonej z tych neuronów, dwa znane algorytmy trenujące oraz powszechnie stosowane miary jakości. Przykładowe zastosowania opisywanych sieci zaprezentowane są w końcowym fragmencie opracowania.

Słowa kluczowe: sztuczne sieci neuronowe, lokalnie rekurencyjne sieci neuronowe, liniowe, nieliniowe i chaotyczne systemy dynamiczne, metody quasi-Newtonowskie, modelowanie heurystyczne.

INTRODUCTION

Nowadays there are many industrial plants that carry out complex processes. In general, a dynamic behaviour of them is difficult to be modelled with

the use of classical analytical methods [1, 2, 3, 4, 6]. Sometimes it is easier and faster (in some cases it is only one way) to use selected heuristic methods (soft modeling methods) in order to solve various engineering problems in modeling tasks [4]. The

paper describes locally recurrent neural networks applied to model non-linear dynamic objects and processes. Presented networks consist of artificial neurons with linear dynamic system blocks. The general concept is not new [2, 5], but some further developments are introduced. Generally, neural networks which are composed of those neurons [2, 5] are able to represent dynamic behaviour of some systems but they are not able to model their stochastic behaviour themselves [4, 6]. Therefore, there is need to elaborate much more general neuron models which can be used for modeling both dynamic and stochastic systems simultaneously.

In this paper, the author develops dynamic neuron model with the IIR or FIR filter in the activation block by introducing the ARMAX (or ARX) system into its structure. The behaviour of that neuron is not always deterministic (it is described in Section 1).

The paper is composed as follows. In Section 1 only some equations for formal descriptions of dynamic neurons are shown. Next, in Section 2 a few different variants of a dynamic unit are presented. Section 3 describes applications of two known quasi-Newton methods (the BFGS method and the LM method) for training depicted neural networks. At the end of the paper, there are three examples presented. They show modeling of dynamic systems, fault-tolerant application and modeling of an electric furnace.

1. ARTIFICIAL NEURON WITH IIR/FIR AND ARX/ARMAX DYNAMIC SYSTEMS

Developing dynamic neural units is one of the most common ways to improve the ability of artificial neural networks to model linear, non-linear and chaotic dynamic systems.

Dynamic and stochastic behaviour is embedded in the neuron by introducing the IIR or FIR filter and the ARX or ARMAX system into its structure.

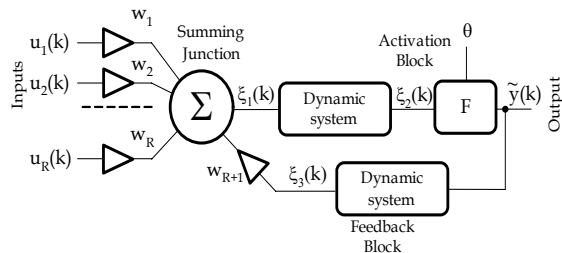


Fig. 1. Formal model of artificial neuron with dynamic systems in the activation and feedback block

The behaviour of the dynamic neural unit under consideration is described by the following equations:

$$\xi_1(k) = \sum_{i=1}^R u_i(k) \cdot w_i + \xi_3(k) \cdot w_{R+1}, \quad (1)$$

where: $\Theta_1 = [w_1 \ w_2 \ \dots \ w_{R+1}]$ is a vector with input weights w_i ; $u_i(k)$ are neuron inputs, $i=1,2,\dots,R$; R is a number of inputs.

$$\xi_2(k) = \sum_{\tau=0}^B b_{\tau} \cdot \xi_1(k-\tau) - \sum_{\tau=1}^A a_{\tau} \cdot \xi_2(k-\tau), \quad (2)$$

where: $\Theta_2 = [a_1 \ a_2 \ \dots \ a_A \ b_0 \ b_1 \ \dots \ b_B]$ is a vector describing the IIR filter placed between the summing junction and the activation block; a_{τ} and b_{τ} are the feedback and feed-forward filter parameters; (A,B) is its order, and

$$\xi_3(k) = \sum_{\tau=0}^C c_{\tau} \cdot \tilde{y}(k-\tau) - \sum_{\tau=1}^D d_{\tau} \cdot \xi_3(k-\tau) + \sum_{\tau=0}^E e_{\tau} \cdot \phi(k-\tau), \quad (3)$$

where: $\Theta_3 = [\mu \ d_1 \ d_2 \ \dots \ d_D \ c_0 \ c_1 \ \dots \ c_C \ e_0 \ e_1 \ \dots \ e_E \ \sigma_e^2]$ is a vector describing the ARMAX system located in the output feedback block; ϕ - vector of a white-noise random process where μ and σ_e^2 are the mean value and variance of a white noise; c_{τ} , d_{τ} , e_{τ} - system parameters; (C,D,E) is the structure representation of ARMAX system. The term $e_{\tau} \cdot \phi(k-\tau)$ is zero for deterministic dynamic systems and is a white-noise random process for stochastic systems.

There are three different ways of calculating its output:

- by using Gaussian activation function:

$$\tilde{y}(k) = F(\xi_2, \Theta) = e^{-\frac{1}{2} \left(\frac{\xi_2(k) - c}{\sigma} \right)^2}, \quad (4)$$

- by using hyperbolic activation function:

$$\begin{aligned} \tilde{y}(k) &= F(\xi_2, \Theta) = \\ &= \frac{2}{1 + \exp[-\beta \cdot (\xi_2(k) + \delta u)]} + b - 1, \end{aligned} \quad (5)$$

- by using linear activation function:

$$\tilde{y}(k) = F(\xi_2, \Theta) = a \cdot \xi_1(k), \quad (6)$$

where: $\Theta = [\sigma \ c \ \beta \ \delta u \ b \ a]$ - a vector depending on selected activation function.

2. DYNAMIC NEURAL NETWORK

This class of artificial neural networks (with dynamic neuron models) is well-known as *locally recurrent globally feed-forward networks* [2,5]. As we can read in [2] that is *somewhere in between a feed-forward and a globally recurrent architecture*. The dynamic unit described in Sec. 1 may be used differently.

Figure 2 shows three units: the first one is a static neuron (for $B=0$, $b_0=1$, $A=1$, $a_1=0$, $C=c_0=D=d_1=E=e_0=0$); the second unit is a completely dynamic neuron model with white-noise random process and hyperbolic (tan for short) activation function; the last one represents also

a dynamic neuron model with the finite impulse response filter and there is Gaussian (gauss for short) activation function. The topology of a dynamic neural network (described in Fig. 3) which has been used in this part of the research consists of three layers (a general structure was chosen based on information in the literature). The first layer has simple static neurons with a non-linear activation function. Hidden layer includes dynamic neurons with a non-linear activation function (dynamic neurons $iir^{(A,B)}armax^{(C,D,E)}-tan$, $fir^{(B)}arx^{(C,D)}-gauss$, $iir^{(A,B)}arx^{(C,D)}-tan$, etc. were tested). The last layer consists of simple static units but the activation function is linear.

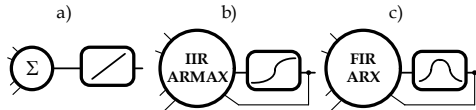


Fig. 2. Example of dynamic neural units (short notation): a) static neuron; b,c) dynamic neurons with the IIR or FIR block and ARMAX or ARX block

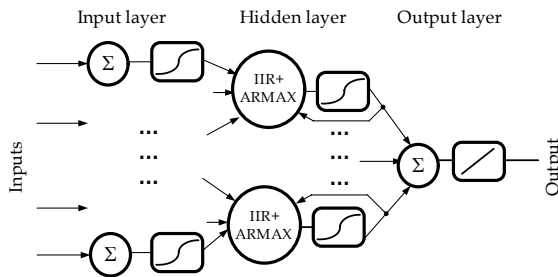


Fig. 3. Structure of a dynamic neural network

All unknown network parameters may be represented by the vectors: ${}^L\Theta = [{}^L\Theta_1 \ {}^L\Theta_2 \ {}^L\Theta_3 \ {}^L\Theta_n]$, where: L is a number of the layer, n is a unit number. For a three-layered network we are able to define: $L \in \{1, 2, 3\}$ and for each layers there are $n = n_L = \{1, 2, \dots, NL\}$, NL - the number of units in the layer L .

3. LEARNING ALGORITHMS AND QUALITY COEFFICIENTS

3.1. Learning methods

There are a lot of optimization problems in many engineering areas which can be formulated as follows: given a real-valued loss (cost) function $f: \mathcal{R}^n \rightarrow \mathcal{R}^1$, find a global minimum,

$$\Theta^* = \arg \min_{\Theta \in C} f(\Theta), \quad (7)$$

where: C is the predefined parameters space, usually a compact set in $\mathcal{R}^n$; Θ^* is the vector with optimal network parameters.

There are many solutions to this problem, but in general two kinds of high-performance algorithms are widely used: based on standard numerical optimization techniques (like gradient methods, quasi-Newton methods) or based on heuristic techni-

ques (like recursive random search, simulated annealing, genetic algorithms or even simple momentum technique in the backpropagation algorithm). In this paper the Broyden, Fletcher, Goldfarb, and Shanno (BFGS for short) quasi-Newton algorithm and the Levenberg-Maquardt (LM for short) algorithm are used. For the cost function that is described by the following equation:

$$f(k; \Theta) = \|y(k) - \hat{y}(k; \Theta)\|, \quad (8)$$

should be minimized basing on a given set of input-output patterns. All the line-search methods explore along the line containing current point $\Theta(k)$, parallel to the search direction $d(k)$ (Eq. 9).

$$\Theta(k+1) = \Theta(k) + \alpha \cdot d(k), \quad (9)$$

where: α is a scalar step length parameter. In the next part of this Section the two ways of definition of a direction will be discussed.

For the BFGS method the direction is given as follows:

$$d(k) = -H^{-1}(k; \Theta) \cdot \nabla f(k; \Theta), \quad (10)$$

where: the gradient information $\nabla f(k; \Theta)$ is derived by partial derivatives using the numerical differentiation method via finite differences (i.e. like in [3]). To avoid a large amount of computation during calculating H numerically the formula of BFGS for its approximation is used:

$$H(k+1) = H(k) + \frac{q(k) \cdot q^T(k)}{q^T(k) \cdot s(k)} - \frac{H^T(k) \cdot s^T(k) \cdot s(k) \cdot H(k)}{s^T(k) \cdot H(k) \cdot s(k)}, \quad (11)$$

where: $H(k=0)$ can be a set as the identity matrix I .

$$\begin{aligned} s(k) &= \Theta(k+1) - \Theta(k) \\ q(k) &= \nabla f(k+1; \Theta) - \nabla f(k; \Theta) \end{aligned} \quad (12)$$

On the other hand, for the LM method the direction can be calculated as follows:

$$d(k) = -[J^T(k; \Theta) \cdot J(k; \Theta) + \lambda \cdot I]^{-1} \cdot \nabla f(k; \Theta), \quad (13)$$

where: the Jacobian information is derived using also a numerical differentiation method, and scalar λ controls both the magnitude and direction.

A very important thing is that the term $e_{\tau} \cdot \phi(k-\tau)$ at each k -th algorithm step has to be fixed (for both algorithms).

3.2. Quality measures

The quality coefficients of neural network models under consideration are given below:

- Mean Absolute Percentage Error (MAPE for short):

$$MAPE = \frac{100}{n} \cdot \sum_{n=1}^N \left| \frac{y(n) - \hat{y}(n)}{y_{\max} - y_{\min}} \right|, \quad (14)$$

- Thiel's statistic (U for short):

$$U = \sqrt{\frac{\sum_{n=1}^N [y(n) - \hat{y}(n)]^2}{\sum_{n=1}^N [y(n) - y(n-1)]^2}}, \quad (15)$$

- conforming with Obuchowicz [2] is measure (J for short):

$$J = \frac{\sum_{n=1}^N [y(n) - \hat{y}(n)]^2}{\sum_{n=1}^N [y(n)]^2}, \quad (16)$$

where: N is the number of patterns; $\hat{y}(n)$ is the network output.

These measures may be employed for different tasks (i.e. the first and second coefficients are meaningful for control systems whilst the last one could be used for comparing with those described in the literature).

4. EXPERIMENTAL RESULTS

In the following parts of the paper some examples of applications will be presented. The first and second subsections show the results obtained using models which have been determined in the Matlab environment like Simulink and some other toolboxes. The last example describes the application of discussed neural structures to data gathered from the real process.

4.1. Modeling of dynamic systems

Problem 1

Given input-output pairs $\{u(k), y(k)\}$ generated by the non-linear system defined by (Narendra K, Parthasarathy K.):

$$f(x_1, x_2, x_3, x_4, x_5) = \frac{y(k-1) \cdot y(k-2) \cdot u(k) \cdot u(k-2) \cdot [u(k)-1] + u(k-1)}{1 + y(k-2)^2 + u(k)^2}$$

identify the topology of the network and its parameters.

For this task several different variants of dynamic and static neural structures were examined. Table 1 presents some results (one can compare different neural structures).

Table 1. Results obtained for feed-forward and locally recurrent neural networks

Type	No. Iter.	J	MAPE	U
	[-]	[-]	[%]	[-]
LM method				
<i>ffbp</i>	200	12e-05	2,76e-01	2,28e-01
<i>ffbp_{td}</i>	200	8,35e-05	2,28e-01	1,19e-01
<i>iir^(2,1)arx^(1,1)</i>	200	9,12e-05	2,56e-01	2,11e-01
BFGS method				
<i>ffbp</i>	200	32e-04	1,47	1,19
<i>ffbp_{td}</i>	200	22e-04	9,47e-01	7,8e-01
<i>iir^(2,1)arx^(1,1)</i>	200	35e-04	7,21e-01	8,1e-01

There are results for three-layered perceptron *ffbp* (the input layer has 16 neurons with hyperbolic activation function; 8 neurons in the hidden layer

with hyperbolic function; 1 neuron in the output layer), for time delay feedforward network *ffbp_{td}* with the same architecture as previous one but also with a delay block (0,1,2) in the input layer. The last one is dynamic neural structure *iir^(2,1)arx^(1,1)* (discussed in Sec. 2, three-layered network 5-12-1 with dynamic neurons and hyperbolic activation function in the hidden layer). As we can see locally recurrent network usually requires much less neurons to obtain almost the same quality result as in the case of a simple feed-forward network with static neurons. However, we must remember that dynamic neuron model is described by much more parameters than straightforward neural unit.

In the training stage the input-output patterns were represented as follows:

$$\{(\mathbf{P}(k), \mathbf{T}(k)) \mid k = 1, \dots, N\}$$

$$\mathbf{P}(k) = [u(k) \ u(k-1) \ u(k-2) \ y(k-1) \ y(k-2)]^T, \quad (17)$$

$$\mathbf{T}(k) = [y(k)]$$

and $u(k)$ was initialized with random values (1000 samples with a uniform distribution in the range $<-2; 1.2>$). In the testing stage three different input signals were used: polyharmonic signal, hard-limit signal and pseudo-random binary sequence (in Table 1 there are errors for the first signal only).

4.2. Fault-tolerant application

Problem 2

Design a virtual sensor $\hat{L}_2$ that could serve as a plain fault detection and identification block (based on the non-linear analytical redundancy technique).

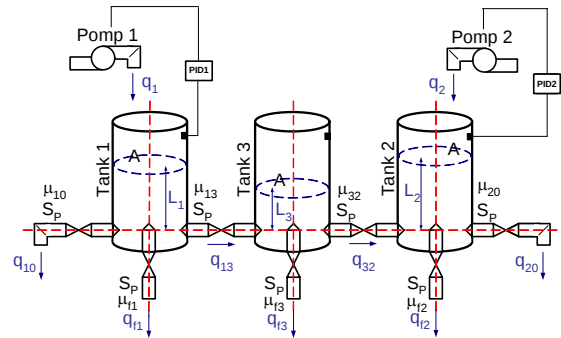


Fig. 5. The three-tank system

The three-tank system model as shown in figure 5 is written using well-known mass balance equations. The system can be represented like in [2] by:

$$\begin{cases} A \cdot \frac{d[L_1(t) + \Delta L_1(t)]}{dt} = q_1(t) - q_{13}(t) - q_{f1}(t) - q_{10}(t), \\ A \cdot \frac{d[L_2(t) + \Delta L_2(t)]}{dt} = q_2(t) + q_{32}(t) - q_{f2}(t) - q_{20}(t) \\ A \cdot \frac{d[L_3(t) + \Delta L_3(t)]}{dt} = q_{13}(t) - q_{32}(t) - q_{f3}(t) \end{cases}, \quad (18)$$

where: q_{ij} represents the water flow rate from tank i to j which is given by Torricelli's rule:

$$q_{ij}(t) = \mu_i \cdot S_p \cdot \text{sign}[L_i(t) - L_j(t) + \Delta L_{ij}(t)] \cdot \sqrt{2 \cdot g \cdot |L_i(t) - L_j(t) + \Delta L_{ij}(t)|} \quad (19)$$

where: ΔL_i represents fault of the measuring channel i , q_{fi} is undesirable leakage from the tank i .

The measured signals are: streams of the medium q_1, q_2 that flow into the first and second tank; control signals U_1, U_2 ; levels in tanks L_1, L_2, L_3 .

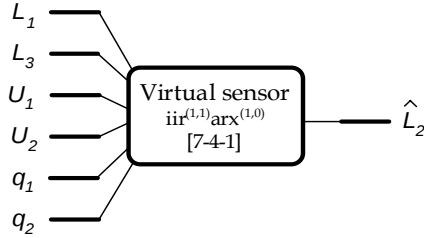


Fig. 6. Virtual sensor for fault detection and identification block

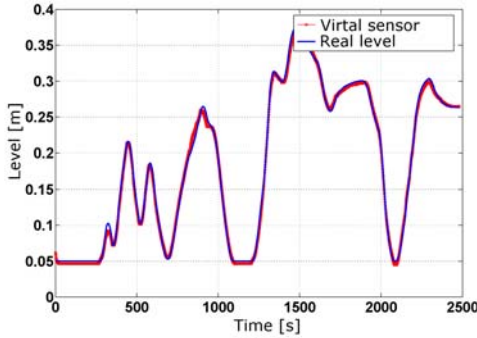


Fig. 7. Virtual and real sensor outputs (results for testing stage: $MAPE < 5\%$)

In order to create a virtual sensor the three-layered dynamic neural network was used (7 neurons with hyperbolic transfer function in the input layer, 4 hidden neurons with the $iir^{(1,1)} arx^{(1,0)}$ blocks and hyperbolic function, one output neuron with linear activation function, Fig. 6). The training and testing patterns were obtained by simulating model (Eq. 18) in the Matlab environment. The error ($MAPE$) was less than 5%, therefore it was possible to use the model in the fault detection and control modules what is shown in Figures 8 and 9.

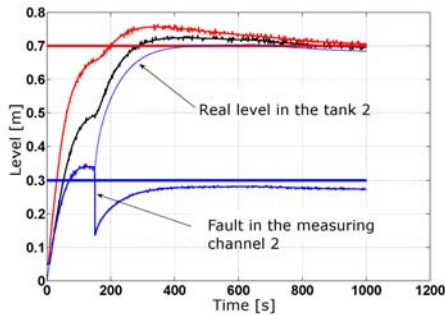


Fig. 8. Fault in the measuring channel 2 (without analytical redundancy)

The fault of the measuring channel no. 2 is detected by comparing value from the sensor L_2 with estimated value $\hat{L}_2$ obtains from the virtual sensor. When absolute difference between these values is greater than 10% of scale range then control system change its rules (Fig.9).

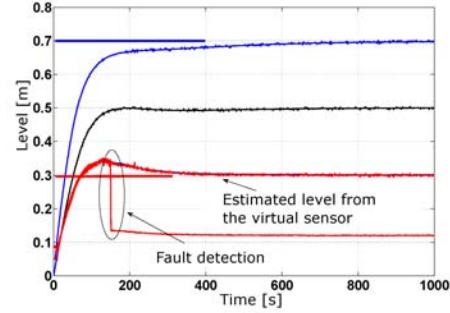


Fig. 9. Fault-tolerant control (with analytical redundancy)

4.3. Modeling of industrial furnace

Problem 3

The database concerning process of copper reduction from slag has been collected by SCADA system. There are more than 180 attributes collected: parameters of power supply, temperatures in many points of the furnace, and state parameters such as mass of charge, and its chemical analysis determined three times a cycle. In order to create a submodel of the analyzed process a non-linear discrete difference equation is proposed (Eq. 20, in the first stage of research inputs and output were selected based on information from staff maintaining this object and technical documentation dealing with this process):

$$\hat{y}_{Cu}(k + \Delta k) = \hat{f}[\mathbf{I}(k), \mathbf{L}(k), \hat{\mathbf{I}}(k), \hat{\mathbf{L}}(k), p(k), y_{Cu}(k)], \quad (20)$$

where: $\mathbf{I}(k)$ is a vector of three currents (for electrodes R, S, T), $\mathbf{L}(k)$ is a vector determines that positions of three electrodes; $p(k)$ is a position of a claw of the transformer; $y_{Cu}(k)$ is a current state parameter describing the copper concentration in slag; Δk is the time horizon of prediction (for this example it equals 5, 10, 15 or 20 min.); $\hat{f}$ represents a non-linear input-output relation of the neural network; $\hat{\mathbf{I}}(k), \hat{\mathbf{L}}(k)$ are suitably preprocessing currents and electrodes positions (detailed description was omitted in the paper).

In the training stage 8000 samples were used. The testing of the model was carried out using another set of 4000 samples. In Table 2 the results received for two neural networks are presented. A simple feed-forward with time delay lines network is presented $ffbptd$ (12 neurons with time delay lines (0,1,2) and hyperbolic activation function in the input layer, 7 neurons in the hidden layer with hyperbolic activation function; one output neuron with linear transfer function). There are also results for dynamic neural network $fir^{(2)} arx^{(1,0)}$

(6 neurons with hyperbolic activation function in the input layer; 5 hidden neurons with the $fir^{(2)}$ filter and the $arx^{(1,0)}$ system and hyperbolic function; one output neuron with linear transfer function).

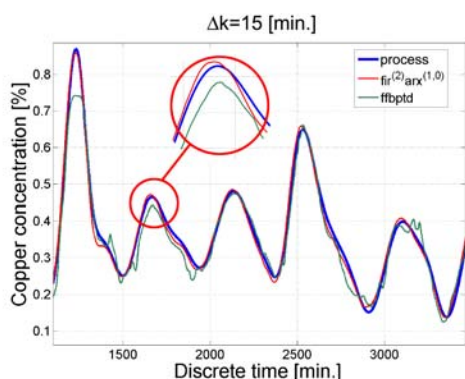


Fig. 10. Example of process and model outputs obtained for different neural structures

Table 2. Results obtained for locally recurrent neural networks

Type	Δk [min.]	J [-]	MAPE [%]	U [-]
LM method				
$ffbptd$	5	8,70e-04	1,17	1,58
$fir^{(2)}arx^{(1,0)}$	5	5,34e-04	0,25	0,50
$ffbptd$	10	2,00e-03	1,79	1,13
$fir^{(2)}arx^{(1,0)}$	10	1,1e-03	0,91	0,55
$ffbptd$	15	5,00e-03	2,09	1,14
$fir^{(2)}arx^{(1,0)}$	15	1,64e-03	1,20	0,49
$ffbptd$	20	9,28e-03	3,82	1,14
$fir^{(2)}arx^{(1,0)}$	20	1,4e-03	0,90	0,26

As we can observe, all the results ($MAPE \approx 1\%$) for each example are quite similar to each other. However, in the case of the $fir^{(2)}arx^{(1,0)}$ network Thiel's coefficients (U) are lesser than in the $ffbptd$ network. Consequently, these models may be useful in fault detection systems or even in control systems.

5. SUMMARY

This work demonstrates theoretical and practical aspects of heuristic modeling based on locally recurrent neural networks. As one can see it is possible to introduce into artificial neuron structure some modules which represent its dynamic and stochastic behaviour (it brings an artificial neuron closer and closer to the biological model). Moreover, some known local optimization algorithms may be used in order to train dynamic network networks. Nevertheless, these algorithms are still too slow for the reason that the gradient information is computed using a numerical differentiation method.

6. TO DO

Many problems are still open to solution, such as the automatic identification of dynamic neural topology. However, the most important tasks are to

find the gradient information using an analytical method and to apply some global optimization algorithms (e.g. genetic algorithms or recursive random search) for preliminary adjusting network parameters.

ACKNOWLEDGMENTS

Paper presents a selected part of the research supervised by Prof. W. Moczulski. This research has been partially supported by the Ministry of Education and Scientific Research under grant No. 4 T07B 018 27.

REFERENCES

- [1] Gobbak Anil K., Raghavendran H., Tapas Anand M.: *Internal Feedback Neuron Networks for Modeling of an Industrial Furnace*. The 1997 Ieee International Conference on Neural Networks. Volume: 2, On page(s): 700-705 vol.2. ISBN: 0-7803-4122-8.
- [2] Korbicz J., Koscielny J. M., Kowalczyk Z., Cholewa W. (Eds.): *Fault diagnosis. Models, Artificial Intelligence, Applications*. Springer-Verlag Berlin Heidelberg 2004 New York. ISBN 3-540-40767-7.
- [3] Majchrzak E., Mochnacki B.: *Metody numeryczne. Podstawy teoretyczne, aspekty praktyczne i algorytmy*. Wydawnictwo Politechniki Śląskiej, Gliwice 2004, ISBN 83-7335-231-7 (In Polish).
- [4] Moczulski W.: *Methodology of Heuristic Modelling of Dynamic Objects and Processes for Diagnostics and Control*. Recent Developments in Artificial Intelligence Methods. AI-METH 2005, p. 123 – 126.
- [5] Sinha N.K., Gupta M.M. and Rao D. H.: *Dynamic Neural Networks: An Overview*. Proceedings of IEEE International Conference on Industrial Technology 2000 (IEEE Cat. No.00TH8482). Volume: 1, On page(s): 491-496 vol.2. ISBN: 0-7803-5812-0.
- [6] Przystałka P., Moczulski W.: *Koncepcja metodyki tworzenia heurystycznych modeli obiektów przemysłowych: Pomiar Automatyka Kontrola*, 2005, Agencja Wydawnicza SIMP, ISSN 0032-4110, materiały VII Krajowej Konferencji Naukowo-Technicznej Diagnostyka Procesów Przemysłowych. Rajgród 2005, str. 136-138 (in Polish).
- [7] Yazdizadeh, A.; Khorasani, K.: *Identification of a class of non-linear systems using dynamic neural network structures*. Proceedings of International Conference on Neural Networks (ICNN'97). Page(s): 194-198 vol.1. Digital Object Identifier 10.1109/ICNN.1997.61162.

CONSTRUCTING MODEL OF KNOWLEDGE BASE ACCORDING TO MOKA METHODOLOGY

Wojciech SKARKA

Silesian University of Technology, Department of Fundamentals of Machinery Design
Konarskiego 18a, 44-100 Gliwice, Poland e-mail: wojciech.skarka@polsl.pl

Summary

The paper presents the structure of a model and the experience of model constructing according to MOKA methodology. The modeling process covers ontology creation of two parts of a model: formal and informal in Protégé-OWL system. Informal model is used for knowledge acquisition for the purpose of knowledge base creation for a Knowledge Based System. Interface of this part of a model are ICARE forms. Formal model includes target domain relations and ontology. Thus, the formal model orders ontology of informal model and defines fully the domain of a given knowledge base. The model has been tested in designing process of a group of machines and can be used in different phases of Knowledge Based System, first of all in designing but also in operation and maintenance. The model includes ontology for products and processes and complex relations between model elements require the usage of advanced tools such as Protégé-OWL for the purpose of model creation.

Keywords: Knowledge Based Engineering, Ontology, OWL, MOKA.

BUDOWANIE MODELU BAZY WIEDZY ZGODNEGO Z METODOLOGIĄ MOKA

Streszczenie

W artykule przedstawiono strukturę modelu oraz doświadczenia z budowy modelu zgodnego z metodologią MOKA. Proces modelowania obejmuje budowę ontologii w systemie Protégé-OWL dwóch części modelu: nieformalnego i formalnego. Model nieformalny służy do pozyskiwania wiedzy do celów budowy bazy wiedzy dla systemu opartego na wiedzy a zewnętrznym interfejsem tej części modelu są formularze ICARE. Natomiast model formalny zawiera docelowe relacje i słownik ontologii. Model formalny porządkuje więc ontologie modelu nieformalnego i definiuje w pełni dziedzinę zadanej bazy wiedzy. Model był testowany w procesie projektowo-konstrukcyjnym i może być stosowany w różnych fazach procesu tworzenia systemu opartego na wiedzy przede wszystkim w zakresie projektowania i konstruowania ale także w dziedzinie eksploatacji. Model zawiera ontologię dla produktów i procesów a złożone relacje między elementami modelu wymagają stosowania do budowy tego modelu zaawansowanych narzędzi budowy ontologii jakim jest Protégé-OWL.

Słowa kluczowe: Projektowanie oparte na wiedzy, ontologia, OWL, MOKA.

INTRODUCTION

MOKA methodology [8] (Methodology for Knowledge Base Engineering Applications) is a complex methodology which was published in 2001. It was made as a result of European Project aiming at elaboration of methodology for Knowledge Based Engineering - KBE. Among other methodologies [1], [5], [9] of creating Knowledge Based Systems it differs from others that it has specific domain of Knowledge Based System which is designing.

This methodology is the first widely spread solution facilitating in a complex way creation of KBE Applications. It covers first of all three most important phases of KBE system development:

- knowledge acquisition,
- knowledge formalization,
- creation of KBE applications.

One of the essential achievements of this methodology is ontology creation for designing domain, covering designed product as well as designing processes. In bibliographies [8] resulting from MOKA project the models are described as UML [10] (Unified Modeling Language) models and MML [8] (MOKA Modeling Language) which is a form of UML language, defined as UML language profile.

Despite the fact that UML language is commonly known and there is a great number of tools for modeling with the aid of this language, these tools are not aimed at solving problems connected with systems based on knowledge. Main usage of this language covers general modeling of environment requirements and issues connected with application modeling (CASE – Computer-Aided Software Engineering). It is favorable for broad usage of the methodology in practice to transfer ontology to

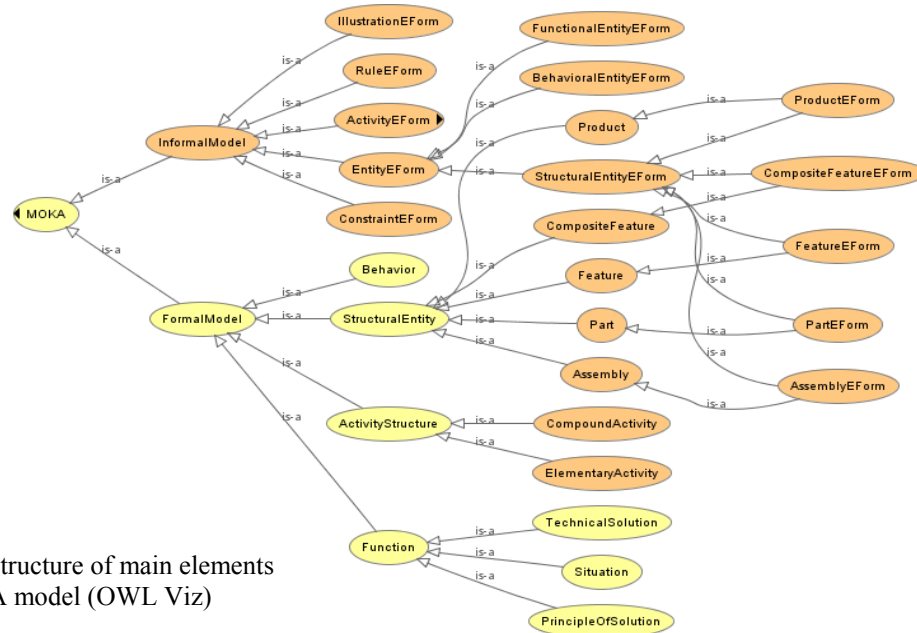


Fig. 1. The structure of main elements of MOKA model (OWL Viz)

a commonly used format of ontology record and use popular tool which aids knowledge base creation and Knowledge Base Systems. Protégé [3], with formal language OWL [11] (Web Ontology Language), which is one of the newest formal languages of ontology record, fulfills these criteria

2. CONSTRUCTING MOKA MODEL USING PROTÉGÉ

Main elements of MOKA model designed in Protégé system has been presented in Fig.1. Original nomenclature of MOKA system has not been preserved as a whole due to specific model structure. In Protégé system the whole model constitutes one coherent namespace, which is not congruent with MOKA model where there are a few namespaces. It results in a names conflict because in each space a name must be unique. Although in Protégé system it is possible to create separate namespaces and to import them, it causes additional problems and therefore, it has been decided to change names of some ontology elements as opposed to MOKA pattern. It is not the only change since some other changes have been introduced to the designed model. Majority of them have been forced by the change of representation from UML (MML) to OWL, however, some extra changes resulting from the target usage of methodology have been added.

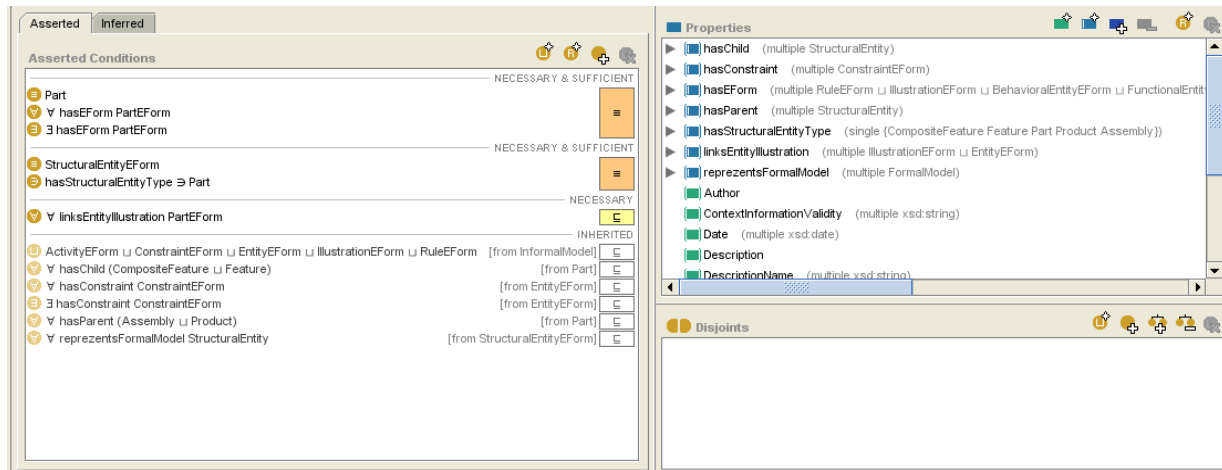
The main structure of MOKA model has been preserved and consists of two parts *InformalModel* and *FormalModel* represented in a system as classes, with *InformalModel* being defined class whereas *FormalModel* primitive class. *InformalModel* has five subclasses which form the most important tools of knowledge acquisition process. In MOKA model they are called *ICARE Forms* because they create forms for five categories of knowledge acquisition (Illustration, Constraint, Activity, Rule, Entity).

Due to the above mentioned conflict of names their notions have been changed by *EForm* suffix. In MOKA originally the forms have been given freedom in relation to formal model which played a role of a formal knowledge base for direct creation of KBE system. In the elaborated model the formal relations control the elements of ICARE forms and thus additional relations between formal and informal model are formed. In the further development of the model it should facilitate automatic creation of KBE system with the usage of CATIA system. For that purpose in the next phase it has been planned to develop the existing model with aspects connected with modeling in CATIA system and in particular Generative Model creation [6, 7].

For the created model it has turned out that the relations of *StructuralEntity* and *ActivityStructure* class and their subclasses are the most important. Relations from *StructuralEntity* subclasses control conceptually complex dependencies for product structure, reflecting them in forms, *ProductEForm*, *AssemblyEForm*, *PartEForm*, *FeatureEForm* and *CompositeFeatureEForm*.

As refers to relations from *ActivityStructure* class they control relations of *ActivityEForm* subclass, whereas *Behavior* and *Function* classes of formal model there is no need for these relations between formal and informal model to exist.

Additionally, apart from the above mentioned classes of formal model there are *Technology* and *Representation* classes in MOKA. Since they have relatively little influence on target KBE system (Generative model), it has been chosen to leave them out in this model, concentrating on general problems of product structure and designing process. The product representation model (*Representation* class) in MOKA is very general and of signal character. For the purpose of representation in CATIA system it is necessary to elaborate

Fig. 2. Conditions and features defining *Part* class

separate detailed formal model of high degree of complexity.

Global structure of class is only a tip of dependencies defined in a model. For each class a number of properties are defined (Fig. 2). These are properties of two kinds: object properties and data type properties. The first ones can be put in a hierarchy. They are not the only ones of defining properties but these two types are of the greatest importance while defining classes. Moreover, conditions are added to classes and there are two kinds of conditions: asserted and inferred.

Conditions can be divided to groups:

- necessary and sufficient,
- necessary,
- inherited.

The first group is important for including a class to defined category. Conditions define a domain in a very detailed way and the range of class properties. Additionally, class resolution can be determined in a given class group. In a formal OWL

record it is important since its fundamental assumption is open world assumption (OWA) [2]. Therefore, special attention must be paid to defining closure restrictions of sets of value ranges accepted by class properties. In some complex cases it can be helpful to declare class disjoints.

3. AN EXAMPLE OF KNOWLEDGE BASE ACCORDING TO DESCRIBED MOKA MODEL (PROTÉGÉ-OWL)

Purposely, MOKA model after the changes are introduced or the model reconstructed is aimed to be used for creation of KBE system with the use of Knowledgeware tools or VB programming [13] in CATIA system. However, up till now the model has been verified in creation of knowledge base for a group of devices using a system of a bolt-nut connection for creating lifting forces or pressure/thrust (screw lift and fly press) (Fig. 3)

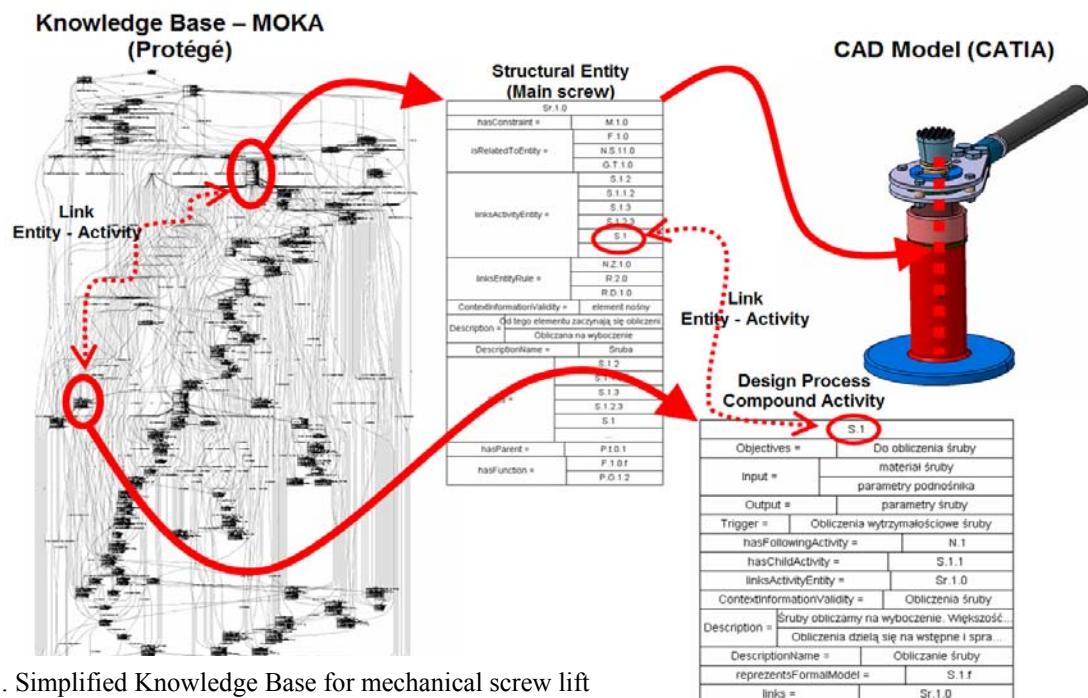


Fig. 3. Simplified Knowledge Base for mechanical screw lift

Due to knowledge base complexity it is not possible to present here the whole knowledge base even for so simple tools. Nevertheless Protégé system has tools which facilitate work with so complex knowledge bases as e.g. graphical representation of a net of relations, contextual functions facilitating recognition and getting to suitable elements of base and the most important – possibilities resulting from formal semantics of OWL and logical relations of ontology- reasoning which consists in generation of new relations between ontology elements derived from logical consequences of ontology in a form of inferred relations (taxonomy classification).

4. CONCLUSIONS

Creation of a model which is compatible with MOKA methodology in Protégé system and OWL language forms a very complex task. Similar task e.g. transferring CommonKADS methodology to Protégé system [4] also indicates high degree of complexity. Because of changes in model format it is not possible to reflect precisely the original MOKA model in Protégé system. UML model and graphical representation of this model has a more user friendly form. However, the lack of tools which would service UML for knowledge acquisition resulted in the choice of Protégé-OWL system. This system is very useful for ontology creation, complex interface requires an experienced designer in knowledge base or knowledge engineer. MOKA system built in Protégé system is very suitable for creation of knowledge bases connected with designing process. The construction of a system based on knowledge with the use of MOKA methodology as well as CATIA system necessitates considerable changes in MOKA model, especially in the aspect of using Generative Model [6], [7] and automation [12], [13] of its creation.

REFERENCES

- [1] Angele, J., Fensel, D., Landes, D., and Studer, R.: *Developing knowledge-based systems with MIKE*. Journal of Automated Software Engineering, 1998
- [2] Knublauch H., Rector A., Stevens R., Wroe C.: *A Practical Guide to Building OWL Ontologies Using The Protégé-OWL Plugin and CO-ODE Tools. Edition 1.0*. The University of Manchester, 2004
- [3] Noy N. F., McGuinness D. L.: *Ontology Development 101: A Guide to Creating Your First Ontology*. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05. March 2001.
- [4] Schreiber G., Crubezy M., Musen M.: *A Case Study in Using Protege-2000 as a Tool for CommonKADS*. In: R. Dieng and O. Corby (eds.) *Knowledge Engineering and Knowledge Management: 12th International Conference EKAW-2000, Juan-les-Pins, volume 1937 of Lecture Notes in Artificial Intelligence*, pp. 33-48, Berlin/Heidelberg, Springer-Verlag, 2000
- [5] Schreiber G. et al.: *Knowledge Engineering and Management. The CommonKADS Methodology*. MIT Press 2000
- [6] Skarka W.: *Automation of designing process using generative models*. (in polish) *Mechanik*. 11/2005 s.781-782, Agencja Wydawnicza SIMP, Warszawa, 2005
- [7] Skarka W.: *Knowledge Acquisition for Generative Model Construction*. Concurrent Engineering CE2006, Antibes, France September 18-22, 2006
- [8] Stokes M. (ed.): *Managing Engineering Knowledge; MOKA: Methodology for Knowledge Based Engineering Applications*. Professional Engineering Publishing, London, 2001.
- [9] Tu, S.W., Eriksson, H., Gennari, J.H., Shahar, Y., and Musen, M.A.: *Ontology-based configuration of problem-solving methods and generation of knowledge acquisition tools: The application of PROTÉGÉ-II to protocol-based decision support*. *Artificial Intelligence in Medicine*, 7(5), 1995
- [10] Website: *Object Management Group – UML* www.uml.org/
- [11] Website: *Web Ontology Language OWL* www.w3.org/2004/OWL/
- [12] Zajac Ł.: *Development of strategy of transferring Knowledge-Based CATIA model into API environment*. Master Thesis. Silesian University of Technology. Department of Fundamentals of Machinery Design, Gliwice, 2006.
- [13] Ziethen D. R.: *CATIA V5 – Makroprogrammierung mit Visual Basic Script*. HANSER, Berlin, 2006



Wojciech SKARKA, Ph.D. Eng. has been an assistant professor at the Department of Fundamentals of Machinery Design since 1998. For 10 years he was a designer and a researcher at Mining Mechanization Centre in Gliwice. He is the designer of many mining machines and the author of a dozen of patents. Currently he carries research on Knowledge Based Engineering and Interactive Electronic Technical Manuals at Silesian University of Technology.

IDENTIFICATION OF SUBMERSIBLE PUMP TEMPERATURE CHANGES MODEL USING KDD METHODS*

Dominik WACHLA

Silesian University of Technology, Department of Fundamentals of Machinery Design
Konarskiego Street 18A, 41-100 Gliwice, Poland, e-mail: dominik.wachla@polsl.pl

Summary

This paper deals with the problem of the autoregressive model identification using KDD methods. In the considered problem, the autoregressive models are applied to describe dynamics processes of various technical systems. In particular, a method of functional dependencies discovering was presented. The method was designed for exploring data sets gathered by industrial SCADA systems. For the problem of the identification of pump temperature changes model, the method was verified. For this particular reason, a set of data was used which was gathered by submersible pumping station SCADA system. The assumptions, the exemplary results of the conducted research and conclusions were presented, as well.

Keywords: databases, knowledge discovery in databases, system identification, genetic algorithm, support vector machines, attributes selection, SCADA systems.

IDENTYFIKACJA MODELU ZMIAN TEMPERATURY POMPY GŁĘBINOWEJ Z ZASTOSOWANIEM METOD ODKRYWANIA WIEDZY W BAZACH DANYCH

Streszczenie

W artykule poruszono problem identyfikacji modeli autoregresyjnych opisujących dynamikę obserwowanych procesów. W szczególności przedstawiono metodę odkrywania zależności funkcyjnych w zbiorach danych gromadzonych przez przemysłowe systemy SCADA. Opracowaną metodę zweryfikowano dla problemu identyfikacji modelu zmian temperatury pompy głębinowej. W tym celu zastosowano fragment danych zgromadzony przez system rejestracji danych współpracujący pompownią głębinową. Przedstawiono przyjęte założenia, fragmenty uzyskanych wyników oraz wnioski z przeprowadzonych badań.

Słowa kluczowe: bazy danych, odkrywanie wiedzy w bazach danych, identyfikacja systemów, algorytm genetyczny, metoda wektorów wspomagających, selekcja atrybutów, systemy SCADA.

INTRODUCTION

Currently, a lot of technical objects and industrial installations have a Supervisory Control and Data Acquisition (SCADA) systems [6]. They are used for the control and monitoring processes occurring in technical systems. One of the elements of SCADA systems is a database. The database is used for recording a measurement results of the observed system variables. The recorded data can be a source of valuable knowledge about a dynamics of processes occurring in the observed technical systems. A lot of available data and their imperfection require that the methods of Knowledge Discovery in Databases (KDD) should be used in a process of knowledge acquisition.

On the other hand, a typical analysis of process data consists of the identification of the quantitative relation occurring in a set of observed process variables. These relations are described by autoregressive models (Fig. 1) [6],[9]. However, the identified autoregressive models don't explain the physical essence of the observed

processes. Nonetheless, they can be applied for the control and diagnosis in a case when their accuracy is appropriate.

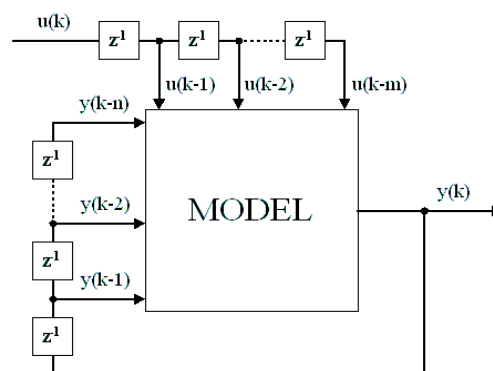


Fig. 1. The general structure of the autoregressive MISO models [10]

The main aim of conducted research was to develop a method of the autoregressive models identification of the observed processes in technical systems. The developed method constitutes

* The research was financed by the Minister of Science and Information Society Technologies (grant No 4 T07B 059 26).

a combination of selected KDD methods and has been designed for the analysis of the data gathered by SCADA systems. The verification of the proposed method was conducted with the application of data gathered by a SCADA system of a submersible pumping station.

1. THE METHOD

In the figure 2 the structure of the proposed method was presented. The method consists of two stages. In the first one, a data transformation into multidimensional space of regressors is done. The data transformation is conducted with the application of the TSI2SI algorithm [10]. The TSI2SI algorithm operation causes change of data description. As a result of TSI2SI algorithm operation, the data in the form of time series are transformed into a set of learning examples (learning vectors):

$$\{a_1(k), a_1(k-1), \dots, a_1(k-n), \dots, a_i(k), a_i(k-1), \dots, a_i(k-n), \dots, a_m(k), a_m(k-1), \dots, a_m(k-n)\} \quad (1)$$

where: m - number of process variables, n - time horizon. Then, the transformed data can be explored by adequate methods of Data Mining (DM).

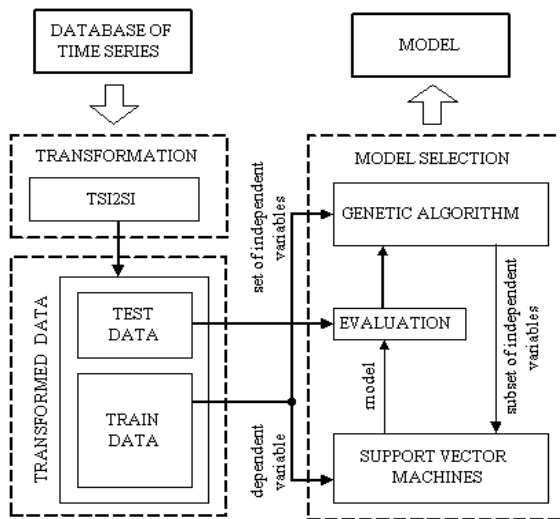


Fig. 2. The structure of the developed method

The application of high values of the parameters m and n in the TSI2SI algorithm induces a big growth in the dimensions of learning vectors (1). For that reason, the second stage of the proposed method is combination of the induction of functional dependencies and the selection of independent variable. This approach is consistent with the wrapper approach [5]. In particular, for the approximation of functional dependencies the Support Vector Machines (SVM) [8] method was chosen. Moreover, the Simple Genetic Algorithm (SGA) [4] for the selection of the independent variables was chosen.

The evaluation function (Fig. 2) is a significant element of the process of the independent variables selection. The general expression for the selection of independent variables using the evaluation function has the following form [6],[9]:

$$\tilde{A} = \arg \min_{A_i \in A} J(h(A_i)) \quad (2)$$

where: A - set of independent variables, A_i - subset of A , $h(A_i)$ - model $J(\cdot)$ - evaluation function.

A lot of model evaluation functions based on the theory of information in the domain of system identification was developed. They allow to evaluate identified models when one takes into account the structure complexity and accuracy. One of the most popular and used function is Akaike Information Criterion (AIC) [1],[9]:

$$J_{AIC} = N \log \text{MSE}(h(A_i)) + 2 \text{card}(A_i), \quad (3)$$

where: MSE - the Mean Square Error, $\text{card}(A_i)$ - a cardinal number of A_i .

Other approach are based on heuristic rules. In case when the two compared models have a similar accuracy and a different structure complexity, the general rule is to choose a model that have a lesser complexity of structure. Basing on this rule we can construct the heuristic function for models selection in the form:

$$J_{HF} = N \log \text{acc}(h(A_i)) - \rho \left(\frac{\text{card}(A) - \text{card}(A_i)}{\text{card}(A)} \right), \quad (4)$$

where: $\text{acc}(\cdot)$ - the accuracy function eg. Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE) and etc.

In the equation (4) the parameter ρ has an important role since its value influences on the process of models selection. The ρ parameter allows to determine a degree of the significance of two criteria considered in equation (4), i.e. the model accuracy criterion and model complexity criterion. For the increasing values of ρ the model complexity criterion is more important than the model accuracy criterion and vice versa for the decreasing values of ρ .

2. THE DATABASE

The economic situation in the Polish Mining Industry has caused liquidation of unprofitable coal-mines. As a result of liquidation a lot of pit shafts have remained. Some of the remaining pit shafts are flooded by underground water. This situation has caused a risk of flooding operative coal-mines. To prevent this, submersible pumping stations were installed. Each of the submersible pumping stations consist of four pumps and has SCADA systems for controlling and monitoring water levels and pumps operation. Each of the installed SCADA system has a database. The database is used for recording

measurements of 34 observed process variables. The measurements are conducted with 1 second interval. For the verification of the method a part of data from the database of a submersible pumping station SCADA system was used. The shared data describe a operation of pumping station within a 3,5 months.

From the point of view of the machinery diagnostics, the following attributes are interesting:

- T - the pump temperature,
- C - the pump capacity,
- P - the power of pump motor,
- E - the current intensity taken by pump motor,
- S - the state of pump operation.

Basing on consultations with experts [3] and our own observations, the pump temperature T was established as significant process variable. In connection with this, a review of available data was carried out. As a results of the review, it was stated that data connected with one of the pumps can be used for the identification of pump temperature changes model because these data describe a selected pump in a „good" technical state. Then the identified model can be used for the diagnosis of a change of the pump technical state. The basis of the diagnosis is a residual analysis here. Additionally, a correlation analysis among all attributes in the shared database was carried out. The correlation analysis has shown that the pump capacity C, power P, current intensity E and sate of pump operation S are a strongly correlated. In the connection with this a set of considered variables was limited to the temperature T and to the state of pump operation S. Then these variables were used for the identification of a submersible pump temperature changes model.

3. MODEL IDENTIFICATION

For identification of submersible pump temperature changes model the following values of parameters in the proposed method was chosen:

- Transformation *TSI2SI*:
 - $n=10$,
 - $\Delta t=1$,
- Genetic Algorithm [4]:
 - crossover probability $p_c = 0.7$,
 - mutation probability $p_m = 0.01$,
 - size of population: 100,
 - number of iteration: 1000,
 - selection method: roulette,
 - type of succession: trivial.
- Support Vector Machines [7],[8]:
 - SVM algorithm: ν -SVR ($\nu=0.54, C=1$),
 - classes of considered SVM models: (Tab. 1).
- Selection criterion: (Tab. 2).

For calculating a value of $acc()$ function (Tab. 2), the Hold-Out method [2] was chosen. In Hold-out method a set of learning data

was divided into two subset. The first subset of learning data is used as a training set and the second subset of learning data is used as a testing set. In the conducted research, sets of the training and testing data had the same number of elements. The complete set of the obtained results of the identification of submersible pump temperature changes model was presented in [10]. In the figure 3, an example of one of discovered models with linear SVM kernel and his statistical evaluations was presented.

Tab. 1. The considered classes of SVM models [10]

Type of model	Kernel	Parameters
MP1	Linear	—
MP2	RBF	$\gamma=0.001$
MP3	RBF	$\gamma=0.01$
MP4	RBF	$\gamma=0.1$
MP5	RBF	$\gamma=1$

Tab. 2. The plan of pump temperature changes model identification [10]

Criterion	$acc()$	ρ	Type of model
J_{AIC}	MSE	—	MP1, MP2, MP3, MP4, MP5
J_{HF}	MAPE	0.1	
J_{HF}	MAPE	0.5	
J_{HF}	MAPE	1.0	

4. SUMMARY

In this paper the method of the data exploration was presented. The method has been designed for analyzing of data gathered by industrial SCADA systems. The essence of the developed method is the data transformation into multidimensional space of regressors. The applied transformation of the process data causes an increase in a number of considered variables. In the connection with this, a variables selection is necessary. For this reason the process of variables selection basing on wrapper approach [5] was applied.

For the identification of the submersible pump temperature changes model the developed method was applied. The part of the data gathered by SCADA system of a submersible pumping station was used as well.

As a results of the conducted research, a lot of submersible pump temperature changes models were obtained. All of the discovered models have a large deviation error in heating and cooling phases of pump operating. Probably, it is caused by using linear and RBF kernels in SVM method only. Nevertheless, the SVM models with linear kernels are more adequate than SVM models with RBF kernel. In the case of nonlinear models (RBF kernel) only the models of MP5 type have a similar degree of reduction of independent variables. The values of

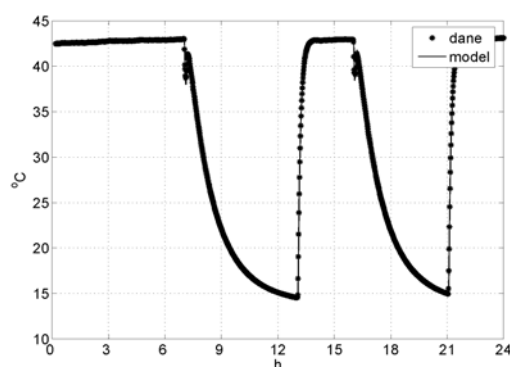
calculated statistics are similar too. However, this accuracy has resulted from an increase in support vectors.

For all of the considered classes of SVM models, the greater values of ρ parameter have caused a decrease in the number of independent variables.

The presented research will be continued to verify the usefulness of different methods of states space searching and other kernels of SVM than linear and RBF. Additionally, we are going to identify a diagnostic model for the recognition of a pump technical state using the knowledge of the submersible pumping station staff.

- SVM model
 - kernel function : linear
 - kernel parameters : —
 - number of SV : 786
 - independent variables : $T(k-1)$, $T(k-2)$,
 $T(k-5)$, $S(k-4)$
 - dependent variable : $T(k)$
- Model evaluation
 - MSE : 3.1416 E-2
 - MAPE : 0.1469
 - correlation : 0.9998

a)



b)

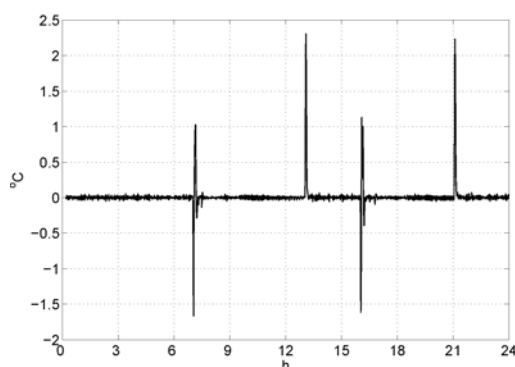


Fig. 3. The exemplary results of the pump temperature changes model identification with the use of SVM linear kernel and $J_{HF}^{\rho=0.5}$ criterion: (a) model (b) residuum [10]

REFERENCES

- [1] Akaike H.: A new look at the statistical model selection. *IEEE Transactions on Automatic Control*, 19:716–723, 1974.
- [2] Ciupke K.: *Selection and reduction of information in machinery diagnostics*. Vol. 118. Department of Fundamentals of Machinery Design, Silesian University of Technology, Gliwice, 2001 (in Polish).
- [3] Gibiec M.: Intelligent health prognostics of machines used in mining industry. *VI National Conference "Diagnostics of Industrial Processes" DPP'05, Rajgród*, ss. 243–245, Warsaw University of Technology, 2005 (in Polish).
- [4] Goldberg D.E.: *Genetic Algorithms in Search, Optimization and Machine Learning*. WNT, Warszawa, 1998 (in Polish).
- [5] Kohavi R., John G. H.: Wrappers for feature subset selection. *Artificial Intelligence*, 97(1-2):273–324, 1997.
- [6] Korbicz J., Kościelny J.M., Kowalczyk Z., Cholewa W. (Eds.): *Diagnostics of Processes. Models, Methods of Artificial Intelligence, Applications*. WNT, Warszawa 2002 (in Polish).
- [7] Schölkopf B., Bartlett P. L., Smola A. J., Williamson R.: Shrinking the tube: a new support vector regression algorithm. vol. 11, ss. 330 – 336, Cambridge, MA, 1999. MIT Press.
- [8] Schölkopf B., Smola A. J.: *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, Cambridge, MA, USA, 2001.
- [9] Söderström T., Stoica P.: *System Identification*. PWN, Warszawa, 1997 (in Polish).
- [10] Wachla D.: *Identification of Dynamic Diagnostic Models Using Methods of Knowledge Discovery in Databases*. Phd Thesis, Silesian University of Technology, Department of Fundamentals of Machinery Design, Gliwice, 2006 (in Polish).



Dominik WACHLA (Phd Eng.) is an assistant at the Faculty of Mechanical Engineering at Silesian University of Technology at Gliwice. His research is focused on the application of methods of artificial intelligence in the technical diagnostics of machinery and processes.

COST-SENSITIVE FEATURE SELECTION*

Krzysztof CIUPKE

Silesian University of Technology, Faculty of Mechanical Engineering
Konarskiego Street 18A, 44-100 Gliwice, Poland, e-mail: kciupke@polsl.pl

Summary

The paper concerns the selection of features in the technical diagnostics domain. The author focused his attention on a wrapper approach. In this approach an application of the ant algorithm as a search engine is proposed. The proposed method of so-called ant wrapper approach is presented. The method takes advantage of cost of features, where the cost is connected with the cost of sensors. The algorithm as a pseudo-code and some results of a verification experiment are shown. The verification was carried out on data derived from an active diagnostic experiment concerning a rotating machine. The obtained results show, that the proposed method could allow to reduce the number of used sensors.

Keywords: feature selection, ant algorithm, machine learning, artificial intelligence, technical diagnostics.

SELEKCJA CECH Z UWZGLĘDNIENIEM KOSZTU ICH POZYSKANIA

Streszczenie

W artykule opisano metodę selekcji cech z zastosowaniem algorytmu mrówkowego. Metoda pozwala także na uwzględnienie kosztu atrybutu, przy czym jego koszt związany jest z kosztem pozyskanie sygnału diagnostycznego. W przypadku gdy sygnał ten jest już wykorzystywany uznaje się, że koszt wyznaczenia danej cechy jest pomijalnie mały. Metodę przedstawiono w postaci pseudo-kodu i zweryfikowano dla danych pochodzących z czynnego eksperymentu diagnostycznego. Uzyskane wyniki pokazują, że istnieje możliwość ograniczenia liczby stosowanych czujników.

Słowa kluczowe: selekcja cech, algorytm mrówkowy, uczenie maszynowe, sztuczna inteligencja, diagnostyka techniczna.

1. INTRODUCTION

The main problem in a process of constructing expert systems is knowledge acquisition. One of the sources of knowledge could be databases. The author focused his attention on problems of knowledge acquisition from databases using *machine learning* methods. The collected data should be pre-processed. The pre-processing includes, among others, *feature subset selection*.

The data collected in the domain of technical diagnostics is usually result of observations and measurements carried out on investigated objects. Contemporary monitoring systems of machinery allow to observe plenty of diagnostic signals simultaneously. For each signal many features can be estimated. So that the total number of features describing a state of a machine can be very high. Not all of the features are important from the point of view of reasoning about a state of the machine. Besides, as shown in [5], quality of the acquired knowledge could deteriorate while using or adding

an improper feature. That is why the selection of a set of considered features becomes a crucial task.

2. THE PROBLEM OF COST OF FEATURES IN THE FEATURE SELECTION TASK

In many problems, for each feature a cost of its acquisition can be specified, and every feature can have its own cost. But there are problems, such as in technical diagnostics, that the cost of features can be defined in different way. As was mentioned above, plenty of signals can be observed and for each signal many features can be estimated. And the main cost which should be taken into consideration is the cost of measuring the signal (among others the cost of sensor, its mounting on the machinery, transmitting the signal) and not the cost of estimation its features.

Let $Attr$ be the set of considered features $a \in Attr$, n be the number of considered features and let the cost of obtaining the features is $C = \{C_1, \dots, C_n\}$. Besides let us assume, that the number of observed signals is m , so the set of features could be grouped

* The research was partially financed by the Ministry of Science and Higher Education (grant No. 4 T07B 018 27)

into m groups $G=\{G_1, \dots, G_m\}$. Thus the groups can be defined as:

$$G_j = \{a \in Attr \mid sensor(a) = j\} \quad (1)$$

where $sensor(.)$ specifies the sensor (signal) number for which the feature a was estimated.

Next, we could make an assumption, that the cost of features in the group is the same, i.e. if the cost of feature $a_i \in G_j$ is C_i then:

$$C_i = C_j, \forall a_i \in G_j \quad (2)$$

On the basis of cost of features defined in such a way becomes an idea of its application into the feature selection problem. The idea could be formulated as follows:

if the feature $a \in G_j$ is chosen by the selection algorithm as a relevant one, then the cost of the other features $b \in G_j$ can be omitted.

Assuming that the cost of features estimation from the given signal is negligible the above mentioned rule becomes true.

3. SELECTION OF FEATURES

In the research the wrapper approach [7] was applied to selection of the set of relevant features. There are many possibilities while defining an search engine using this approach. In the paper an application of ant algorithm is presented.

3.1. Ant algorithm

Ant algorithms are inspired by real ants behaviour and are one of the most successful examples of swarm intelligence systems. The problem representation is suitably mapped on a graph. To each arc (i, j) of the graph is associated a variable τ_{ij} called *pheromone trail*. Pheromone trails are read and written by ants. The utility of the arc (i, j) to build good solutions is proportional to the amount of pheromone τ_{ij} [4].

An ant k located in node i uses the pheromone trails to choose the next node to move from the set of one-step neighbours N_i with some probability. The probability, in the simple ant colony optimisation algorithm, is computed as follows [4]:

$$p_{ij}^k = \begin{cases} \tau_{ij} & j \in N_i \\ 0 & j \notin N_i \end{cases} \quad (3)$$

While moving from node i to j each ant deposit some pheromone $\Delta\tau_{ij}$ on the arc (i, j) . In the cycle ant colony optimisation algorithm ants deposit a constant amount of pheromone after reaching the goal state. So each ant will change the value of τ_{ij} in time t :

$$\tau_{ij}(t) = \tau_{ij}(t-1) + \Delta\tau \quad (4)$$

Similarly to real pheromone trails, artificial pheromone trails „evaporate“. The evaporation is carried out in an exponential way:

$$\tau(t) = (1 - \rho)\tau(t-1), \quad \rho \in (0,1] \quad (5)$$

at each iteration of the algorithm. So finally, the amount of pheromone on the arc (i, j) could be calculated as follows:

$$\tau_{ij}(t) \leftarrow (1 - \rho)\tau_{ij}(t-1) + \sum_{k=1}^K \Delta\tau_{ij}^k, \quad (6)$$

where: K - number of ants, $\Delta\tau_{ij}^k = 1/L_k$ if the ant k chosen arc (i, j) otherwise $\Delta\tau_{ij}^k = 0$, L_k - length of the k -th ant route.

The probability (3) could also be defined using not only the information about pheromone trails, but also heuristic information η_{ij} [1,6]:

$$p_{ij}^k = \frac{[\tau_{ij}]^\alpha [\eta_{ij}]^\beta}{\sum_{l \in N_i} [\tau_{il}]^\alpha [\eta_{il}]^\beta}, \quad (7)$$

where α and β are parameters that control the relative weight of pheromone trail and heuristic information.

In the research as heuristic information the cost of features was used.

3.2. Ant wrapper approach

Using the ant algorithm as a search engine in the wrapper approach some assumption should be made. Let $Attr$ be the set of input attributes, and let K denotes cardinality of the set $Attr$. Each node represents an attribute, and each arc (i, j) corresponds to choosing the j -th attribute as the next attribute after the i -th one.

Moreover, let the length of each arc be the same and be equal one. Then, the length of the k -th ant route L_k is equal the number of visited nodes, i.e. the number of chosen attributes.

Assuming this, the procedure of attribute selection using wrapper approach with the ant algorithm as a search engine, here called *ant wrapper approach* [3], could be formulated. The algorithm as a pseudo-code is presented in Fig. 1.

The procedure *AntWA* begins computations with the whole set of attributes $Attr$. At the first step (line 1) the set $Attr$ is evaluated, e.g. the quality q_0 of classification using all the attributes is calculated. Then, ants are located in nodes (line 2) and arcs are initiated with some small amount of pheromone (line 3). Also the cost is initiated (line 4). For each ant k the set of neighbour nodes N (i.e. not visited yet) is calculated (line 9) and the probability P of choosing each neighbour node is computed (line 10). The next node to move to is selected by the *NextAttr()* procedure at line 11. If as the next attribute an attribute a is chosen then – according to

```

procedure AntWA(Attr)
1  q_0 = SetOfAttrEvaluation(Attr);
2  M = AntsInit();
3  A = PheromonInit();
4  K = NumberOfAttr(Attr);
5  C = CostOfAttrInit(K);
6  while(terminat. crit. not satisf.)
7    for(k=1 to K)
8      while(M(k) ≠ Attr)
9        N = Attr \ M(k);
10       P = NextAttrProb(A,N,C(k));
11       a = NextAttr(P);
12       C(k) = ChangeCosts(C(k),a);
13       M(k) = AddAttr(M(k),a);
14       q = SubsetOfAttrEval(M(k));
15       if(CloseEnough(q, q_0))
16         A = UpdatePheromone(M(k));
17         DeleteAnt(k);
18         NewAnt(k);
19         C(k) = CostOfAttrInit();
20       end
21     end
22   end
23   PheromoneEvaporation(A);
24 end
end procedure

```

Fig. 1. Ant wrapper approach pseudo-code

the idea described above – the costs of attributes from the group G_j where $a \in G_j$ are set to one (line 12). Each ant builds its own solution. So that, the costs should be defined for each ant separately ($C(k)$ at line 12). A list of nodes (attributes) visited by the ant is stored in the memory matrix M . Next, at line 14, the subset of attributes is evaluated (the same procedure as in line 1 is applied), and if the quality q is close enough to q_0 (e.g. $q=q_0$) then the pheromone trial is updated (line 16), the ant is removed and new ant begins its route with initial costs of attributes. The evaporation process is done by the procedure at line 23, and the computation is repeated till termination criteria (specified at line 6) are not satisfied.

3. VERIFICATION OF THE PROPOSED METHOD

To confirm the usefulness of the considered method for the diagnostic knowledge acquisition process, verification research was required. The research was carried out for a data set derived from an active diagnostic experiment.

3.1. The investigated object and the data set

As an object a model of a rotor machine called *RotorKit* [2] was used (see Fig. 2). The object allows to observe a rotor-bearings system during its operation at different rotating speeds and allows to introduce a few malfunctions. Two discs mounted on the rotor allow to produce the system imbalance and additional equipment allows to produce other malfunctions as e.g. overload.

Four sensors were used to observe relational vibrations in the points A and B (see Fig. 2). The

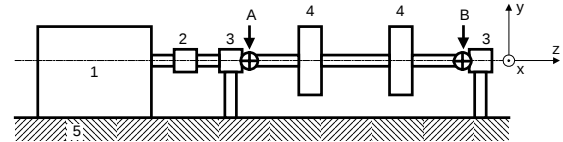


Fig 2. *RotorKit* - a rotor machine model. 1 - DC motor, 2 - coupling, 3 - bearings, 4 - discs, 5 - foundation, A, B - measuring points of relative vibrations

vibrations were observed in two directions: horizontal (x) and vertical (y).

Five technical states (classes) of the model were considered:

- one-plane imbalance (IMB1),
- two-plane imbalance (IMB2),
- overload (OVL),
- rub (RUB),
- whirl (WRL).

The number of examples in each class is presented in Table 1.

Table 1. Number of measurements per class

Class	IMB1	IMB2	OVL	RUB	WRL	Total
Number	40	98	95	60	48	371

On the ground of signals measurements a set of data was obtained. The initial set of diagnostic signal features was specified on the basis of literature. Displacement amplitudes of points A and B in x and y directions for different frequency components (e.g. 1X, 1.5X,...) were mainly used. Finally a set of 38 features was applied in the research.

The features were grouped according to the signal for which they were estimated. Four signals were used, so the features were grouped to four groups. Two features: phase difference for the components 1X observed in points A and B in x or y directions were grouped into successive two groups.

3.2. Obtained results

The ant wrapper approach was applied to the data set described above. As the machine learning method the See5 system was applied [8]. The procedure had only one termination criterion: the number of steps which was fixed to 300.

The results describing minimal and maximal length of obtained subsets of attributes are presented in Fig. 3. In Fig. 4 the number of subsets obtained in every step of the procedure is shown as well as the average length of the subsets.

As one could see (Fig. 3a) the smallest subsets were obtained for the parameters $\alpha=1$, $\beta=0$. The subsets were obtained very quickly and as shown in Fig. 4a there are many subsets of relevant features. By increasing the importance of the cost of features ants were “forced” to change their “natural” route. Much fewer subsets were found (see Fig. 4b) and the cardinality of the smallest subsets varied from 3 to 5 (see Fig. 3b).

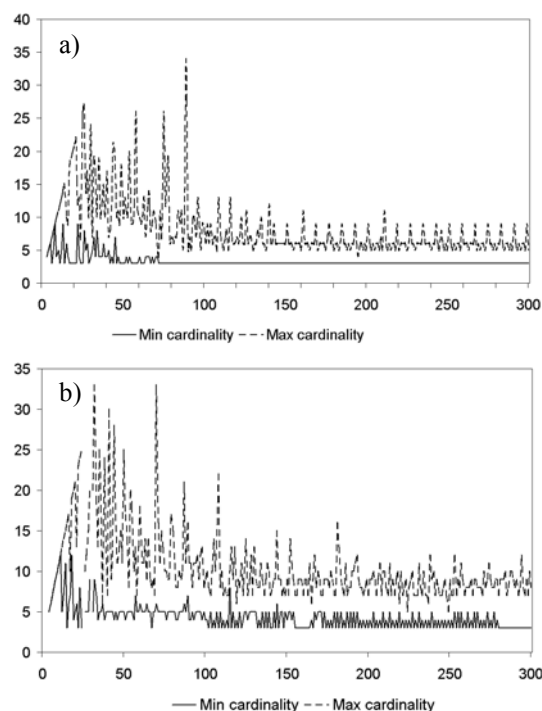


Fig. 3. Minimal and maximal cardinalities of subsets of features for a) $\alpha=1, \beta=0$ and b) $\alpha=1, \beta=1$

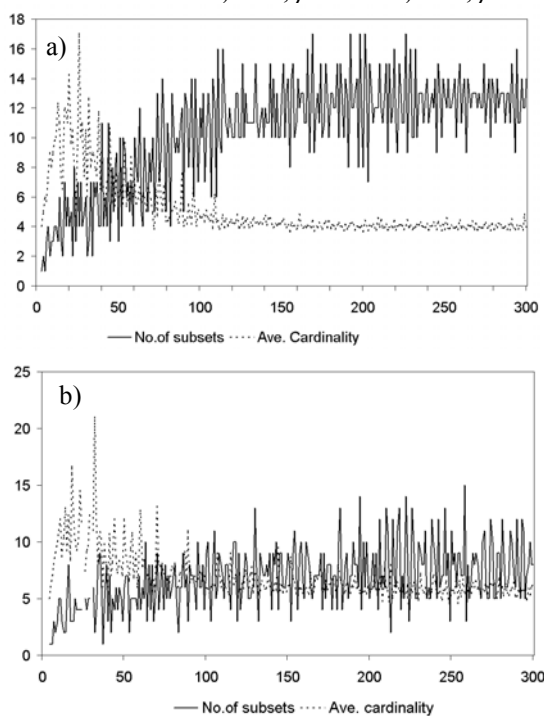


Fig. 4. Average cardinalities and number of subsets of features for a) $\alpha=1, \beta=0$ and b) $\alpha=1, \beta=1$

The most important is the answer to the question whether the increasing the value of the parameter β produced desired results considering groups of chosen features. The answer is positive, i.e. using the parameters $\alpha=1, \beta=0$ the most frequently and the smallest subsets of features contains features from three groups (three sensors are needed), while in other cases ($\alpha=1, \beta>0$) mainly features from two groups were chosen, so that only two sensors are needed to identification of the technical state of the

considered object. The features allowing the state identification are: maximal relative displacement amplitude of the shaft in the point B in y direction for the rotating frequency range 0.2-0.5 and displacement amplitude of the shaft for frequency 3X in A and B points x direction (where X denotes the rotating frequency of the shaft). The efficiency of the classification using ten-fold cross-validation technique is 71.5%.

3. SUMMARY

In the article a new method of cost-sensitive feature selection is presented. The cost of features is strictly connected with the cost of acquiring the diagnostic signal for which the feature is evaluated. The verification of the method was carried out for the data derived from the active diagnostic experiment. The obtained results show that the method allows not only to reduce the number of features but also the number of signals, and simultaneously, the number of used sensors.

REFERENCES

- [1] Boryczka U.: *Co-operation through stigmergy in ant colony optimization for TSP*, Proc. of the 7th Intern. Conf. on Soft Computing Mendel 2001, Brno, Czech Republic, pp. 191-196.
- [2] Ciupke K. *Application of ant algorithm to selection of diagnostic information.*, (in Polish). Pomiary, Automatyka, Kontrola, 2005, No. 9bis, pp. 134-135.
- [3] Dorigo M., Di Caro G., Gambardella L.: Ant algorithms for discrete optimization. *Artificial Life*, 1999, Vol. 5, No. 3, pp. 137-172.
- [4] Hand D. J.: *Discrimination and classification*. Chichester, J. Wiley and Sons, 1981
- [5] Jensen R., Shen Q.: Finding Rough Set Reducts with Ant Colony Optimization. *Proceedings of the 2003 UK Workshop on Computational Intelligence*, pp. 15-22, 2003.
- [6] Kohavi R., John G.: The wrapper approach. In: H. Liu and H. Motoda: *Feature extraction, construction and selection: A data mining perspective*. Kluwer Academic Publisher, 1998.
- [7] C5/See5 Home page, url: www.rulequest.com.



Krzysztof CIUPKE - Assistant Professor at the Department of Fundamentals of Machinery Design, Silesian University of Technology at Gliwice. His research activities include application of methods of artificial intelligence in technical diagnostics and machinery designing.

AN OPTIMIZATION OF HEURISTIC MODEL OF WATER SUPPLY SYSTEM WITH GENETIC ALGORITHM

Bogdan WYSOGLĄD, Ryszard WYCZÓŁKOWSKI

Silesian University of Technology, Department of Fundamentals of Machinery Design
Konarskiego 18a, 44-100 Gliwice, Poland, e-mail: bwysoglad@polsl.pl

Summary

The paper presents the method of optimization of heuristic model of water supply system with Genetic Algorithm. This model is an essential part of intelligent diagnostic system of local water supply system. The main task of this system is water leakage detecting and localization. For inputs, this system uses information from pressure or flow sensors, mounted on the pipeline network, the output is a piece of information about leakage detection and localization. The main advantage of this system is a possibility of approximate leakage localization using only a limited number of installed sensors. The first problem which should be solved to apply this system in practice is to find the best localization of sensor which should be installed on water pipeline. The method of solving this problem was described in the paper.

Keywords: diagnostic, genetics algorithm, neural network, water pipeline.

OPTIMALIZACJA MODELU HEURYSTYCZNEGO SIECI WODOCIĄGOWEJ Z ZASTOSOWANIEM ALGORYTMU GENETYCZNEGO

Streszczenie

W artykule przedstawiono sposób optymalizacji heurystycznego modelu sieci wodociągowej, opartego o sztuczne sieci neuronowe, z zastosowaniem algorytmu genetycznego. Model ten stanowi zasadniczy element inteligentnego układu diagnostyki sieci wodociągowej, mającego za zadanie wykrywanie przecieków sieci wodociągowej. Wejściami do modelu są dane z czujników ciśnienia lub przepływu zainstalowanych na sieci, zaś wyjściami informacja o lokalizacji potencjalnych przecieków. Podstawową zaletą tej koncepcji systemu diagnozowania sieci wodociągowej jest możliwość przybliżonej lokalizacji uszkodzeń sieci w oparciu o ograniczoną liczbę czujników na niej zainstalowanych. Istotnym problemem który należy rozwiązać w pierwszym etapie budowy systemu, jest wybór lokalizacji ograniczonej liczby czujników, zapewniających jednak identyfikowanie jak największej liczby potencjalnych awarii. Artykuł przedstawia sposób rozwiązania tego problemu poprzez optymalizację rozmieszczenia czujników z zastosowaniem algorytmu genetycznego.

Słowa kluczowe: diagnostyka, algorytm genetyczny, sztuczne sieci neuronowe, sieci wodociągowe.

1. INTRODUCTION

Water supply systems are one of the most essential parts of the urban and rural technical infrastructure. It is necessary for them to be reliable, especially because of counteraction of water loss. Finding leaks is one of the typical problems connected with water pipelines maintenance. This task is not simple, because quite often leaking water can run deep into ground and therefore pipe failure does not show up on the ground surface. Bearing this in mind one can expect that a diagnostic system, supporting leakage finding would be very useful, especially on an industrial area with coal mining, where leakages are often encountered.

To avoid necessity of using measuring system which is big, complex and spread out at significant area of the country, with big number of on line measuring points, the concept of diagnostic system, which uses artificial neural network (ANN) for modeling the pipeline network and recognizes a leak

of water was suggested [9]. The idea of this system is based on methods known from model-based process diagnostics where a model of the object being monitored is used for fault detection. Based on measuring pressure and flow in chosen points on pipeline networks, diagnostic system will suggest if and where the leakage is.

The methods of optimal location of sensors for optimum estimation of unknown parameters of distributed systems are still developed. For example some methods, based on optimum experimental design, where state of observed object is described with partial differential equation was described in [8].

In presented paper authors described other method which take advantage of genetic optimization.

2. PROBLEM DESCRIPTION

The first problem which should be solved to apply this conception in practice is to find the best localization of sensor which should be installed on water pipeline. The considered water supply system has more than 50km of pipes with different diameters, and working for about one thousand water consumers.

In the first stage of the diagnostic system creation, because of economical and technical reasons, the owner of the supply system decided to limit the number of used flow sensor to six. The number of pressure sensors, which are cheaper and much easier to mount, was not limited.

But the problem which point of network should be observed as an input to neural network when we have a limited number of sensor still should be considered.

3. THE PROPOSED METHOD OF OPTIMIZATION

The task that must be solved is an optimal choice of a location of 6 sensors in the water supply network. The values of pressures and flows measured by these sensors make up input vector of the neural networks, which recognize a leak of water and its localization.

The algorithm of the solution to the problem was showed below.

1. Randomly choose a few dozen of measuring points.
2. Randomly choose a few dozen of points in which water leaks will be simulated. The ANN classifier will recognize these points.
3. Generation of data (training and test) with the use of pipeline network model made in the EpaNet simulation environment [10].
4. Represent the problem variable domain as a chromosome of a fixed length. Define a fitness function. Randomly generate an initial population of chromosomes.
5. Calculate the fitness of each chromosome. As the measure of chromosome's fitness the efficiency measure of the ANN classifier was used (correct classification ratio [1]).
6. Typical operation for GA (selection, crossover, mutation) to create a new population.
7. Go to point 5, and repeat the process until the termination criterion is satisfied.
8. Verification of usefulness of the ANN classifier chosen in point 7 with the use of a new set of data.

4. CLASSIFIER

4.1. Generation of data

To train the neural network, which is the essential part of fault detector of diagnostic system, the set of data was prepared. Because there is no sensor on the water network now, to generate data the numerical model of water supply system was

used. To build this model, the EpaNet [10] program was applied. Running this model it was possible to calculate flows and pressure in all points of network with leakage located in any point of it or without leakage at all.

Because this is an huge number of possibilities, allowed sensor localizations was limited to main junction only. In all main junction one potential pressure sensor and flow sensors for each connected link was localized. It provided 16 pressure and 45 flow sensors.

To prepare data for ANN training, 33 points of leakages were chosen at random. For each leakage pressures and flows for all potential sensors were calculated. Simulations were repeated for every hour for 30 days. During simulation theoretical daily pattern of water consumption was considered [3]. For every simulation, individual water consumption for each consumer was randomly modified, but changes never exceed 20% of the expected value. The same procedure was repeated for no leakage case.

In this way, more than 27.000 sets of training data, for different leakages and hours of the day were collected. Every set contain data from all potential sensor localizations, and was used for training and testing different ANNs with any subsets of initial set of sensors as an input.

4.2. Verification of usefulness of data

In the first step, the usefulness of data to build ANN classifier was checked. During the research a *Neural Network Toolbox* of MatLab [11] was used.

From the economical point of view the usage of only pressure sensors would be the best solution, because it is easy and cheaper to install a pressure sensor in a pipe than a flow sensor. But during the preliminary research it was observed that the changes of pressures in measuring points of network both with leakage and without leakage were very small and because of measurement inaccuracy, these small pressures changes are practically insignificant.

To check data usefulness, three classifiers were built. The input vectors were different for each classifiers and contained:

- 61 values of pressures and flows of water,
- 45 values of flows of water,
- 16 values of pressures of water.

To build classifier a two-layer network, with tan-sigmoid transfer function in the hidden layer and a log-sigmoid transfer function in the output layer was used. This is a useful structure for classification problems [7].

The network had $n = 34$ output neurons because there are 33 points of leak. On the basis of information available in books [5], concerning the recommended number of neurons, one used $k = 3 \cdot n$ neurons in the hidden layer.

Before training the preprocessing routine procedure was used to scale the inputs so that they are in the range $[-1,1]$. The data was to divide into

training and test subsets. One-fourth of the data for the test set, and three-fourths) for the training set.

The backpropagation training algorithms was used for training. The training was stopped after 300 iterations.

The result of training was the neural networks, which recognized leaks. In the Table 1 the results of research of three classifiers were presented. The best results were obtained in case of classifier which had input vector composed only from values of water flows in pipes.

Tab. 1.

Input vector include:	Number of inputs	Number of outputs	Correct classification ratio
pressures and flows	61	34	87,8 %
pressures	16	34	67,4 %
flows	45	34	92,8 %

4.3. The structure of neural classifier

Basing on the results of research obtained in chapter 4.2 a decision to take only water flows during further research was taken.

In order to find the best localization for flow sensors, with the use of genetic algorithms, a two-layer network with tan-sigmoid transfer function in the hidden layer and a log-sigmoid transfer function in the output layer was used. The network had an input vector containing 6 values of flows, 34 output neurons and 18 neurons in the hidden layer.

Practical application of genetic algorithm requires the fitness calculating for each testing chromosome. It was expected that more than 1000 values of correct classification ratios of the ANN classifiers should be calculated. In this case appropriate training algorithm of ANN has a significant influence on efficiency of research.

The resilient backpropagation (Rprop) training algorithm is one of the fastest algorithms on pattern recognition problems [6]. Its performance also degrades as the error goal is reduced. The memory requirements for this algorithm are relatively small in comparison to the other algorithms.

5. GENETIC ALGORITHM

5.1. Representation of the problem domain as a chromosome

At the beginning we represented the problem domain as a chromosome which consists 34 genes represented by 0 or 1. Each gene represents one point of the model of the water supply network in which pressure and flow of water can be measured. If a gene number i is equal 1 that means that value of flow measured in the point number i is an input to the ANN. As mentioned above, it was assumed that the number of measuring points is equal 6. It is

cause each chromosome contains 6 genes represented by 1 and 28 genes represented by 0.

5.2. Selection

At the beginning population of $N = 50$ chromosomes was chosen. As a fitness function, the efficiency measure of the ANN classifier (correct classification ratio) [1] was used.

During research the roulette wheel selection [4], one of the most commonly used chromosome selection techniques, was used.

5.3. Crossover and mutation operators

Because each chromosome ought to contains exactly 6 genes represented by 1, typical crossover and mutation operators were modified for using in described problem.

At first, the crossover operator randomly chooses 3 crossover points where two parent chromosomes 'break', and then exchanges the chromosome parts. As a result, two new offspring's are created. If number of genes represented by 1 is different then 6, the chromosome is deleted and crossover operation is repeated.

The mutation operator changes the value of a randomly selected gene in a chromosome. After first operation chromosome contains 5 or 7 genes represented by 1.

In next step, if chromosome contains 5 genes represented by 1, the mutation operator randomly selects gene in a chromosome which value is 0 and changes its value for 1. And accordingly, if chromosome contains 7 genes represented by 1, the mutation operator randomly selects gene in a chromosome which value is 1 and changes its value for 0.

6. RESULTS

During the research crossover with the probability p_c equal to 0.7 and mutation with the probability p_m equal to 0.01 was used. The values chosen for p_c and p_m are fairly typical in GAs [2].

A common practice in GAs is to specify the number of generation. During the research one assumed that the desired number of generations is 50.

The most serious problem in the use of GAs is concerned with the quality of the results, in particular whether or not an optimal solution is reached. It depends on the number of chromosomes in population and the value of p_m . Fig. 1 shows plots of the best and average values of the fitness function across 50 generations (performance graphs). The best and average curves represented here are typical for GAs [4].

In the Tab. 2 the obtained results of research were presented. The best result of fitness function

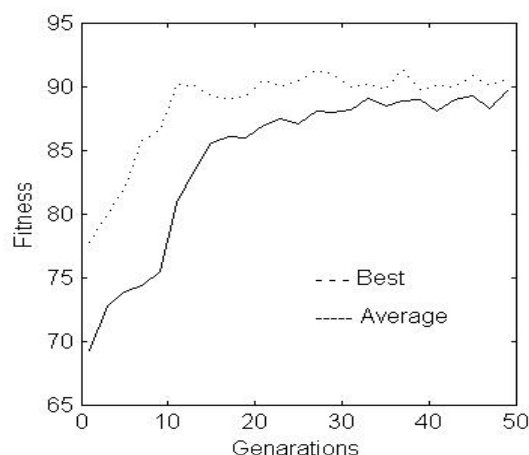


Fig. 1. Performance graphs created in a population of 50 chromosomes, mutation rate $p_m=0.01$

(a correct classification ratio) obtained in case of classifier which had input vector composed from 6 values of water flows in pipes was equal to 90,8 %. In comparison with the result obtained in the case of classifier which had input vector composed from 45 values (correct classification ratio was equal to 92,8 %) the final result of research one can consider as satisfactory.

Tab. 2

	Value	Generation number
Max. fitness	90,8 %	37
Max. average fitness	89,4 %	50

7. CONCLUSIONS

Because it is the first stage of diagnostic system creation, and the main task of the research was to determine where sensors should be mounted, described classifier still wasn't verified on a real water supply system.

In the near future, in collaboration with the local water supply system holder, six water flow sensors will be mounted in the points of water pipeline network chosen during the research. Then it will be possible to test the numerical model of water supply system with the measures obtain from real object.

After the verification of the numerical model new sets of data and classifiers will be generated. If obtained results will be good enough, the future diagnostic system will be based on six chosen measuring points.

The described classifier was prepared to distinguish only 33 points of leakage (and leakages located near them). After sensor installing the classifier will be developed to increase its accuracy.

Decreasing number of simulated points of leakage during ANN training seems to be the simplest solution to do it. But in this way a number of neurons in ANN will be significantly reduced too.

Using fuzzy classifier seems to be more promising method of developing described system, and authors are going to verify this way in the nearest future.

REFERENCES

- [1] Cholewa W., Kaźmierczak J.: *Data Processing and Reasoning in Technical Diagnostics*. WNT, Warsaw, 1995.
- [2] Goldberg D.E.: *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, 1989.
- [3] Mielcarzewicz W.: *Obliczanie systemów zaopatrzenia w wodę*. Arkady, Warszawa 2000.
- [4] Negnevitsky M.: *Artificial Intelligence. A guide to Intelligent Systems*. Addison-Wesley, 2004.
- [5] Osowski S.: *Sieci neuronowe do przetwarzania informacji*. Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa, 2000.
- [6] Riedmiller M., Braun. H.: *A direct adaptive method for faster backpropagation learning: The RPROP algorithm*. Proceedings of the IEEE International Conference on Neural Networks, 1993.
- [7] Rutkowski L.: *Metody i techniki sztucznej inteligencji*. PWN, Warszawa, 2005.
- [8] Uciński D.: *Optimal sensor location for parameter estimation of distributed processes*. Internacjonal Journal of Control, 73(13): 1235-1248, 2000.
- [9] Wyczółkowski R., Moczulski W.: „Concept of intelligent monitoring of local water supply system”. Materials of AI-METH 2005 – Artificial Intelligence Methods. November 16-18, 2005, Gliwice, Poland.
- [10] www.epa.gov/ORG/NRMRL/wswwd/epanet.html.
- [11] www.mathworks.com/access/helpdesk/help/pdf_doc/nnet/nnet.pdf



Bogdan WYSOGLĄD is an assistant professor in the Department of Fundamentals of Machinery Design at Silesian University of Technology. He conducts research in the field of machine building and maintenance. His researches are focused on: technical diagnostics of machinery, signal analysis and application of methods of Artificial Intelligence.



Ryszard WYCZÓLKOWSKI is an assistant professor in the Department of Fundamentals of Machinery Design at Silesian University of Technology. His researches are focused on: design and operation of the machines, computer science, particularly machinery diagnostics and application of methods and means of Artificial Intelligence.

ARTIFICIAL INTELLIGENCE ALGORITHMS COMBINED WITH THE PIES IN IDENTIFICATION OF POLYGONAL BOUNDARY GEOMETRY

Eugeniusz ZIENIUK, Andrzej KUŹELEWSKI, Wiesław GABREL

University of Białystok, Faculty of Mathematics and Physics, Institute of Computer Science
Sosnowa Street 64, 15-887 Białystok, Poland, e-mail: {ezieniuk, akuzel, wgabrel}@ii.uwb.edu.pl

Summary

Identification of a shape of a boundary belongs to a very interesting part of boundary problems called inverse problems. Various methods were used to solve these problems. Therefore in practice, there are two well-known methods widely applied to solve the problem: the FEM and the BEM. In this paper a competitive meshless and more effective method – the PIES combined with artificial intelligence (AI) methods is applied to solve the shape inverse problems. The aim of the paper is an examination of two popular AI algorithms (genetic algorithms and artificial immune systems) in identification of the shape of the boundary.

Keywords: the PIES, identification of a boundary shape, genetic algorithms, artificial immune systems.

ALGORYTMY SZTUCZNEJ INTELIGENCJI POŁĄCZONE Z PURC W IDENTYFIKACJI KSZTAŁTU WIELOKĄTNEJ GEOMETRII BRZEGU

Streszczenie

Identyfikacja kształtu brzegu należy do bardzo interesującej grupy zagadnień brzegowych nazywanej zagadnieniami odwrotnymi. Istnieje liczna grupa metod służących rozwiązywaniu takich problemów. Jednakże w praktyce do rozwiązywania zagadnień odwrotnych szeroko wykorzystywane są dwie metody: MES i MEB. W niniejszej pracy zaproponowano zastosowanie alternatywnej bezelementowej i bardziej efektywnej metody – PURC połączonej z algorytmami sztucznej inteligencji (SI) do identyfikacji kształtu brzegu. Celem pracy jest zbadanie efektywności dwóch popularnych algorytmów SI (algorytmów genetycznych i sztucznych systemów immunologicznych) w identyfikacji kształtu brzegu.

Słowa kluczowe: PURC, identyfikacja kształtu brzegu, algorytmy genetyczne, sztuczne systemy immunologiczne.

INTRODUCTION

The problem of identification of a shape of a boundary consists in reproducing unknown part of the boundary using experimental values of measurements obtained in a few points of the domain or on the boundary. In practice, the problem is reduced to finding the whole domain, where solutions in a few points only are known. Identification of a shape of the boundary belongs to ill-posed problems [6], that is small fluctuations in input data might lead to obtaining significantly different results.

The commonly applied method of solving these kind of problems is based on minimization of a functional. From a practical point of view, it is equivalent to multiple solving of analysis problem with modification of a shape of the boundary geometry. Various numerical methods have been used to solving analysis problems, however the commonly applied methods are the Finite Element Method (FEM) [1, 12] and the Boundary Element Method (BEM) [2, 12]. These methods there is assumed, that the shape of the boundary geometry is

considered using the discretization process. Effectiveness of the BEM or the FEM can be noticed in solving of analysis problems, that is in searching solutions of differential equations in given domains and known boundary conditions. However, use of discretization is very inconvenient when it should be repeated in every iteration of the identification process.

The alternative method of solving analysis problems is the Parametric Integral Equation System (the PIES) [8,10]. The shape of the boundary geometry is inserted into mathematical formalism of the PIES and continuously defined, that is the discretization process is not required. Definition of the boundary geometry is reduced to assigning a few corner points. Therefore, the identification process is reduced to finding only a few corner points.

Another part of identification process using above-mentioned method is an optimization algorithm applied to modifying the shape of the boundary geometry. In previous paper [11] Newton's iterative method was successfully adapted in the case of boundary problems described by Laplace's equation. However, this algorithm

requires to start the identification process from close neighbourhood of a searching point. In practice, it is rarely occurred situation. This problem can be solved using heuristic methods of optimization. Artificial neural networks (ANN), genetic algorithms (GA) and artificial immune systems (AIS) belong to widely used AI methods of optimization.

The aim of this paper is to compare results obtained from two methods of identification of the shape of the boundary geometry. Both are based on the PIES, however different methods of heuristic optimization (GA and AIS) are applied.

1. FORMULATION OF THE PROBLEM AND PROPOSED METHOD OF SOLVING

The identification of unknown part of boundary geometry in 2D boundary problem described by Laplace's equation is considered. In practice, the problem is reduced to finding an unknown part of the boundary on the basis of experimentally measured values in a few measurement points ($u_1, u_2, \dots, u_l$) in domain Ω or on given part of the boundary.

In process of reproducing an unknown part of the boundary the PIES is applied. The manner of defining and modifying of polygonal boundary geometry in the PIES is presented in fig. 1.

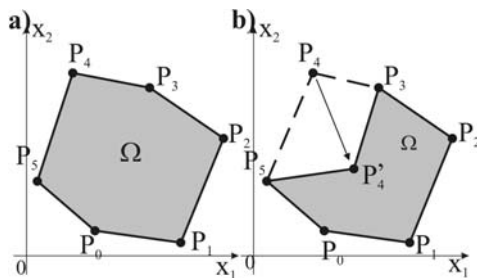


Fig.1. a) Defining the shape of polygonal boundary geometry using corner points, b) Modifying the shape of polygonal boundary geometry by corner point P_4

Corner points P_i are designed for defining of a shape of the boundary geometry and, contrary to the traditional element methods, they are not making discretization of the boundary. After moving an arbitrary corner point P_i , great part of the boundary is changed. Using that point we can either modify or identify the shape of the boundary geometry. Therefore, in practice the identification process is reduced to finding coordinates of corner points.

The identification process is closely connected with experimental values of measurements obtained in measurement points on the boundary or in the domain. The commonly applied method of solving identification problems is based on minimization of the functional:

$$P(s, P_p) = \frac{1}{m} \sum_{i=1}^m [\tilde{u}_l(s_i, P_p)^* - u_l(s_i, P_p)]^2, \quad (1)$$

where: $\tilde{u}_l(s_i, P_p)^*$ are results of experiments, $u_l(s_i, P_p)$ are results obtained from numerical method (i.e. the PIES), P_p – corner point.

The minimization is an iterative process and can be executed using various number of method. We take into consideration two heuristic methods: GA and AIS.

2. THE PIES FOR LAPLACE'S EQUATION

In order to solve the analysis problem the PIES for 2D Laplace's equation was applied. The system is presented by the following formula [8]:

$$0.5u_l(s_1) = \sum_{j=1}^n \int_{s_{j-1}}^{s_j} \left\{ \tilde{U}_{lj}^*(s_1, s) p_j(s) - \tilde{P}_{lj}^*(s_1, s) u_j(s) \right\} J_j(s) ds, \quad l=1, 2, \dots, n \quad (2)$$

where: n – number of segments, s and s_1 – parametric variables $s_{j-1} < s < s_j$, $s_{l-1} \leq s_1 \leq s_l$. Integrands are presented as follows:

$$\begin{aligned} \tilde{U}_{lj}^*(s_1, s) &= \frac{1}{2\pi} \ln \frac{1}{[\eta_1^2 + \eta_2^2]^{0.5}}, \\ \tilde{P}_{lj}^*(s_1, s) &= \frac{1}{2\pi} \frac{\eta_1 n_1^{(j)}(s) + \eta_2 n_2^{(j)}(s)}{\eta_1^2 + \eta_2^2} \end{aligned} \quad (3)$$

where: $S_k(k=j, l)$ – functions of any degree, which define boundary segments, $\eta_1 = S_l^{(1)}(s_1) - S_j^{(1)}(s)$ and $\eta_2 = S_l^{(2)}(s_1) - S_j^{(2)}(s)$. The boundary geometry in the PIES is inserted into integrands (3) and defining using curves (such as Beziér curves, B-spline, v-spline) of arbitrary degree.

The solution in the domain $x \in \Omega$ can be obtained using new integral identity [9] presented by the following relation:

$$u(x) = \sum_{j=1}^n \int_{s_{j-1}}^{s_j} \left\{ \hat{U}_j^*(x, s) p_j(s) - \hat{P}_j^*(x, s) u_j(s) \right\} J_j(s) ds, \quad x \in \Omega \quad (4)$$

Integrands $\hat{U}_j^*$, $\hat{P}_j^*$ are presented by formulas:

$$\begin{aligned} \hat{U}_j^*(x, s) &= \frac{1}{2\pi} \ln \frac{1}{[\tilde{r}_1^2 + \tilde{r}_2^2]^{0.5}}, \\ \hat{P}_j^*(x, s) &= \frac{1}{2\pi} \frac{\tilde{r}_1 n_1^{(j)}(s) + \tilde{r}_2 n_2^{(j)}(s)}{\tilde{r}_1^2 + \tilde{r}_2^2} \end{aligned} \quad (5)$$

where $\tilde{r}_1 = x_1 - S_j^{(1)}(s)$ and $\tilde{r}_2 = x_2 - S_j^{(2)}(s)$.

3. PROPOSED ALGORITHMS OF OPTIMIZATION

3.1 Application of GA

GA consist in the emulation of nature [5, 6]. The elementary processes occurring in nature (enabling individual species to survive and continuously adapt themselves to the environment, such as natural selection and reproduction) are finding application in GA. The idea of GA can be described as follows:

1. Create the initial population randomly.

2. Calculate the fitness of each individual from the population (using the PIES).
3. At each iteration, do:
 - 3.1 Select individuals from population depending on their fitness (selection).
 - 3.2 Recombination (cross-over and mutation).
 - 3.3 Calculate the fitness (using the PIES).
 - 3.4 If solutions are close enough to measured values or fixed number of iterations are achieved stop the iteration process.

The algorithm applied in this paper differs from traditional genetic algorithm described in [5]. Proposed algorithm uses the floating-point representation, ranking selection, simple and arithmetical crossover, and uniform and non-uniform mutation [6].

3.2 Application of AIS

Formal definition of AIS is presented in [3] and reads as follows: "Artificial immune systems are adaptive systems, inspired by theoretical immunology and observed immune functions, principles and models, which are applied to problem solving".

In practice, optimization algorithms based on AIS exploit two processes from theoretical immunology: cloning and changing of the shape of antibodies (hypermutation). Combination of these mechanisms is applied in clonal selection algorithm [4]. Application of this algorithm combined with the PIES can be described as follows:

1. Create a random initial population in given domain of searching and calculate the fitness of each individual from the population (using the PIES).
2. At each iteration, do:
 - 2.1 Reproduce the whole population, that is create a number of clones of the each individual from the population.
 - 2.2 Perform hypermutation on the each clone.
 - 2.3 Calculate fitness for the each clone (using the PIES) and find only the best members from all groups of the clones.
 - 2.4 If solutions are close enough to measured values or fixed number of iterations are achieved stop the iteration process.

4. RESULTS OF EXPERIMENTS

From practical point of view, we may consider any physical problem, which can be modelled by Laplace's equation.

The first problem of identification of the shape is presented in figure 2. Coordinates of two points P_0 and P_3 are unknown. Two domain of searching are considered: $a=4$ and $a=12$.

All experiments were repeated 100 times and mean values were calculated. In order to compare two methods of heuristic optimizations the number of calculations of fitness function in one iteration was fixed and equals 80. Thus, the size of population in GA was 80, and in AIS was 20 with 4 clones for

an individual. In order to assess the quality of searching the distance between founded and real point was calculated. It was assumed that maximum number of iterations equals 200. Measurement of time of executing of one iteration was carried out on the same PC and applications were compiled by the same compiler. Results of experiments are presented in the table 1.

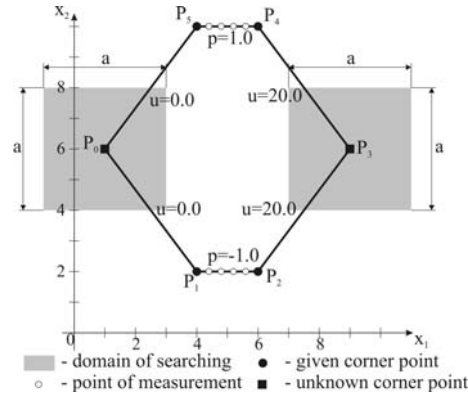


Fig. 2. Identification of the shape of the boundary (searching corner points do not belong to one segment)

Table 1. Results of the first identification problem

a	GA + the PIES		AIS + the PIES	
	min	mean	min	mean
number of generations				
4	21	35.56	24	51.76
12	45	66.00	35	99.26
value of fitness function				
4	5.1e-6	8.3e-6	5e-6	8.7e-6
12	4.9e-6	7.2e-6	5.1e-6	7.6e-6
distance between founded and real point				
4	0.0018	0.012	0.0008	0.011
12	0.0022	0.010	0.0002	0.0097
time of executing of one iteration [s]				
-	15.76		2.8	

As can be seen results obtained using GA and AIS are very similar. Only the application with AIS is about four times faster then the one with GA.

In the second example more complicated problem is considered (fig. 3). Two searching points are neighbours, that is the points create one segment.

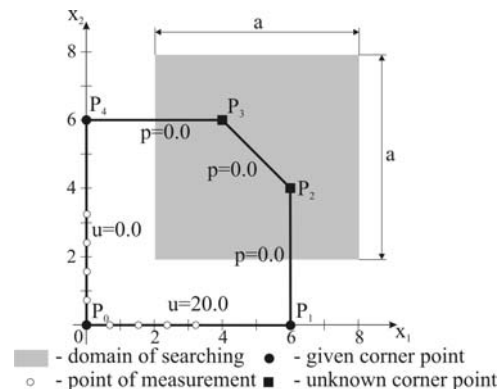


Fig.3. Identification of the shape of the boundary (searching corner points belong to one segment)

All experiments were repeated 100 times and mean values were calculated. The size of population in GA was 200, and in AIS was 50 with 4 clones for an individual. It was assumed that maximum number of iterations equals 1000. Results obtained in the second example are presented in the table 2.

Table 2. Results of the second identification problem

a	GA + the PIES		AIS + the PIES	
	min	mean	min	mean
number of generations				
6	863	964.2	1000	1000
value of fitness function				
6	6.1e-6	8.9e-6	4.1e-6	4.2e-6
distance between founded and real point				
6	0.0217	0.182	0.0146	0.106
time of executing of one iteration [s]				
-	28.46		4.5	

As can be seen similarity between results obtained using GA and AIS is much less than in first example. The application with AIS is about six times faster and more accurate then the one with GA.

5. CONCLUSIONS

Combination of the PIES and two methods of heuristic optimization in order to identify boundary geometry were presented in the paper.

Advantages of the PIES (in comparison with the BEM or the FEM) allow to reduce either the size of input data or the number of searching points. Hence, it increases the effectiveness of the presented algorithms (comparing to the traditional mesh methods).

Application of AIS seems to be more effective and accurate than GA, therefore it is not a big difference. Accuracy of both methods might be insufficient in more complicated problems. Therefore, in the authors opinion, the next researches should be oriented to searching hybrid methods, which combine heuristic algorithms of optimization with traditional ones.

REFERENCES

- [1] Beer G., Watson J. O.: *Introduction to finite and boundary element methods for engineers*. John Wiley & Sons, New York, 1992.
- [2] Brebbia C.A., Telles J.C.F., Wrobel L.C.: *Boundary element techniques, theory and applications in engineering*. Springer-Verlag, New York, 1984.
- [3] De Castro L. N., Timmis J.: *Artificial Immune Systems: A New Computational Approach*. Springer-Verlag, London, 2002.
- [4] DeCastro L. N., von Zuben F. J.: *The clonal selection algorithm with engineering applications*. GECCO'00, pp. 36-37.
- [5] Goldberg D. E.: *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison

– Wesley Publishing Company, Massachusetts, 1989.

- [6] Michalewicz Z.: *Genetic Algorithms + Data Structures = Evolution Programs*. Springer – Verlag, Berlin, Heidelberg, 1996.
- [7] Tikhonov A. N., Arsenin V. Y.: *Solution of Ill-posed problems*. John Wiley & Sons, New York, 1977.
- [8] Zieniuk E.: *Potential problems with polygonal boundaries by a BEM with parametric linear functions*. Engineering Analysis with Boundary Elements, 2001, Vol. 25, No. 3, pp. 185-190.
- [9] Zieniuk E.: *A new integral identity for potential polygonal domain problems described by parametric linear functions*. Engineering Analysis with Boundary Elements, 2002, Vol. 26, No. 10, pp. 897-904.
- [10] Zieniuk E.: *Bézier curves in the modification of boundary integral equations (BIE) for potential boundary-values problems*. International Journal of Solids and Structures, 2003, Vol. 40, No. 9, pp. 2301-2320.
- [11] Zieniuk E., Bołtuć A.: *Identification of polygonal domains using PIES in inverse boundary problems modeled by 2D Laplace's equation*, TASK QUARTERLY, 2005, Vol. 9, No. 4, pp. 415-426.
- [12] Zienkiewicz O. C.: *The Finite Element Methods*. McGraw-Hill, London, 1977.



Eugeniusz ZIENIUK is an associated professor in Institute of Computer Science, University of Białystok. His main research interests include computer methods in boundary problems. He is the author of the parametric integral equations system.



Andrzej KUŹELEWSKI is an assistant in Institute of Computer Science, University of Białystok. His research interests include interval and fuzzy arithmetic, artificial intelligence methods in connection with boundary problems.



Wiesław GABREL is an assistant in Institute of Computer Science, University of Białystok. His main research interests include artificial intelligence methods, especially genetic and evolutionary algorithms, in connection with boundary problems.

APPLICATION OF EVOLUTIONARY ALGORITHMS TO IDENTIFICATION OF DIAGNOSTIC INVERSE MODELS

Grzegorz URBANEK

Silesian University of Technology, Department of Fundamentals of Machinery Design,
Konarskiego 18a, 44-100 Gliwice, Poland, e-mail: grzegorz.urbanek@polsl.pl

Summary

Determination of two-stage models with auxiliary signals is one of possible method of inverse models identification based on examples. First stage of this diagnostic model can be a classifier based on selected features that classifies examples to predefined auxiliary signals classes. The efficiency of the model identified in this way depends on set of the selected features. The proposed method of evolutionary search of relevant features set and the obtained results of the research were described in the paper.

Keywords: multi-stage models, diagnostic inverse models, evolutionary algorithms.

ZASTOSOWANIE ALGORYTMÓW EWOLUCYJNYCH W IDENTYFIKACJI ODWROTNYCH MODELI DIAGNOSTYCZNYCH

Streszczenie

Jedną z metod identyfikacji odwrotnych modeli diagnostycznych na podstawie przykładów jest wyznaczanie ich jako modeli dwustopniowych z użyciem sygnałów dodatkowych. Pierwszy stopień takiego modelu może być rozpatrywany jako klasyfikator, który na podstawie wybranych cech sygnałów diagnostycznych klasyfikuje przykłady do klas zdefiniowanych w przestrzeni cech sygnałów dodatkowych. Jakość tak identyfikowanego modelu zależy w głównej mierze od użytego zbioru cech sygnałów diagnostycznych. W artykule przedstawiono zaproponowaną metodę ewolucyjnego poszukiwania zbioru cech relewantnych oraz wybrane wyniki przeprowadzonych badań.

Słowa kluczowe: modele wielostopniowe, odwrotne modele diagnostyczne, algorytmy ewolucyjne.

INTRODUCTION

Possession of a diagnostic model has a great significance in the technical diagnostics. This model converts diagnostic signals features to object state features. So, such a model is an inverse model of an object model that converts state features to diagnostic signals features.

One of the method of inverse models identification is identification on the basis of examples obtained using numerical object models. This method allows to identify a global model or a set of local models.

The solution of the task of a global model identification is difficult for the sake of large size of this task. The size of this task is large because of a huge number of the diagnostic signals features that

are possible to determine and because of a large number of the state features [2]. The size of the task of identification of an inverse model can be reduced provided that the inverse model will be identified as a set of independent local models.

1. DECOMPOSITION OF THE GLOBAL OBJECT MODEL

A simple decomposition of a global model to local models connected with assemblies, subassemblies or parts of the object could lead to obtaining a set of dependent local models. The method that allows to obtain independent local models is a decomposition of the object model into a two-stage model using auxiliary signals features (see Fig. 1).

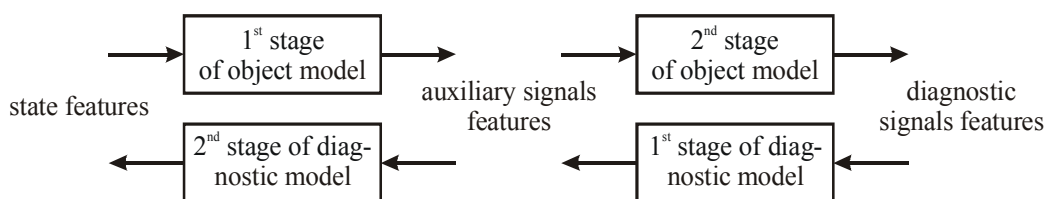


Fig. 1. Decomposition of object model and diagnostic model

Then the object model can be considered as a system that consists of two subsystems [2]:

- First-stage system that transforms state features to auxiliary signals features;
- Second-stage system that transforms auxiliary signals features into observed signals features (symptoms).

Each stage of the two-stage model is a set of independent local models.

Identification of an inverse model of the decomposed object may be realized also as an identification of a two-stage inverse model. This task consists of three subtasks:

- determining the first stage of the diagnostic model as an inverse model of the second stage of considered object model,
- determining the second stage of the diagnostic model as an inverse model of the first stage of the considered object system,
- determining an associative diagnostic model by serial links between both stages of the diagnostic model where the output of the first stage is linked with the input of the second stage.

2. IDENTIFICATION OF THE FIRST STAGE OF DIAGNOSTIC MODEL

The paper concerns the identification of the first stage of the diagnostic model only.

The task of searching the dependency between inputs (values of observed signals features) and outputs (values of auxiliary signals features) of local models (see Fig. 2) of the 1st stage of diagnostic model is difficult.

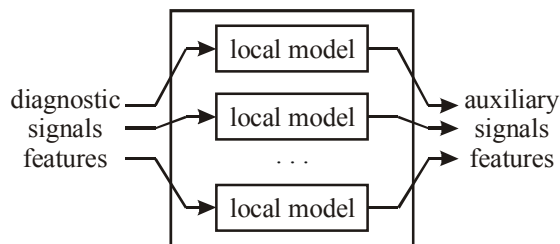


Fig. 2. 1st stage of diagnostic model

The solution of the problem could be searching a set of relevant diagnostic signals features that guarantee similar distribution of the learning examples in this features space and in the auxiliary signal features space.

The proposed method of the identification of the 1st stage of the diagnostic model consists of two steps:

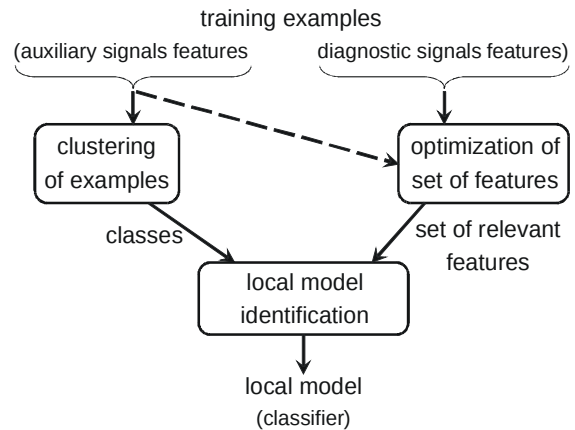


Fig. 3. Scheme of the method of local models identification

- selection of the auxiliary signals features,
- identification of local models that consist of the following tasks (see Fig. 3):
 - determination of the clusters in the auxiliary signals space,
 - selection of the relevant features,
 - designing the classifiers.

3. CONSIDERED OBJECT

The considered object was the turbomachine. The scheme of the machine is presented in Fig. 4. The identified inefficiencies were the displacements of the bearings (marked with numbers 1-7) from the reference position [4].

The learning examples were obtained from Institute of Fluid-Flow Machinery Polish Academy of Science in Gdańsk. The set of learning examples consists of two subsets:

- first subset containing 57 training examples:
 - the reference state: 1 example,
 - small displacements of the one from seven bearings: 28 examples
 - large displacements of the one from seven bearings: 28 examples
- second subset containing 32 testing examples:
 - simultaneous displacements of the bearings 5 and 6.

Each example is represented by values of 2212 features (316 for each bearing).

4. LOCAL MODELS

Local models were assigned to each of the 7 bearings. As the auxiliary signals features the bearing loads were used. Using the bearing loads allow the separation of the local models. These loads

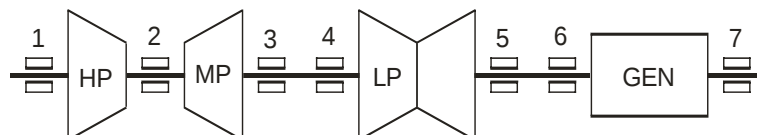


Fig. 4. Scheme of the machine

are not observed directly during object exploitation but they may be determined during simulation researches. The loads depend on displacements of the associated bearing as well as on displacements of other bearings. The loads also influence the shaft location and the values of vibration signals.

It was assumed that the local models should identify based on selected signal features each bearing as a classifier for the bearing loads. The following classes of bearing loads were considered (Fig. 5):

- R – if the end of the reaction vector is on the right to the end of the reference reaction vector,
- L – if the end of the reaction vector is to the left of the end of the reference reaction vector,
- U – if the end of the reaction vector is above the end of the reference reaction vector,
- D – if the end of the reaction vector is below the end of the reference reaction vector,
- C – the reference reaction point.

These classes were defined on the basis of clusters of training examples in the bearing loads space (see Fig.6). The clusters were determined on the ground of analysis of figures which present training examples as points in the bearing loads space (Fig.6 is an example of such a figure).

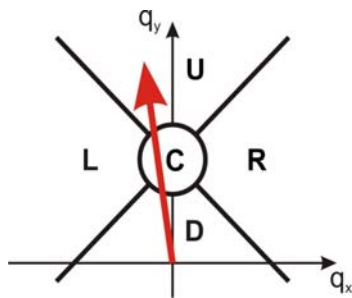


Fig. 5. Classes in bearings loads space (big arrow is an example of bearing load vector)

The classifier is defined by the classes and by the classification criterion. The classes were defined by coordinates of they centers. The nearest neighbor

was considered as a classification criterion.

4.1. Selection of the relevant features

Over two thousand features were determined for observed diagnostic signals of examined object. For the purpose of the task size reduction only few relevant features for each local model should be selected.

For the selection of an optimal features set the evolutionary algorithm was used (Fig. 7). The number f_n of selected features is the parameter of evolutionary algorithm $f_n \in \{4,5,6,7,8,9,10\}$. It was assumed that the population consists of 20 individuals. The genotype of the individual is a row vector with f_n numbers (identifiers) of selected features. This vector defines a feature space, which should be compared with the fixed auxiliary features space. The comparison of spaces is done on the cluster basis. The clusters in the space of features are the phenotypes. The value of a fitness function for an individual is proportional to mean number of the same nearest neighbors in both spaces.

$$fit = \frac{1}{t} \sum_{i=1}^t \frac{n_i}{n_c} \quad (1)$$

t – number of training examples,

n_c – number of considered nearest neighbors, $n_c=20$,

n_i – number of neighbors of i -th example repeated among n_c nearest neighbors in the relevant features space and in the auxiliary signals features space.

The algorithm uses four operators [1],[3],[5]:

- proportional selection,
- two-point crossover (features replacement),
- uniform mutation,
- elitist succession.

Also, the 'repair' operator that prevents features repeating in a genotype is used.

Two termination conditions was assumed:

- achieving of 500 generations,
- 250 generations without increasing the value of fitness function.

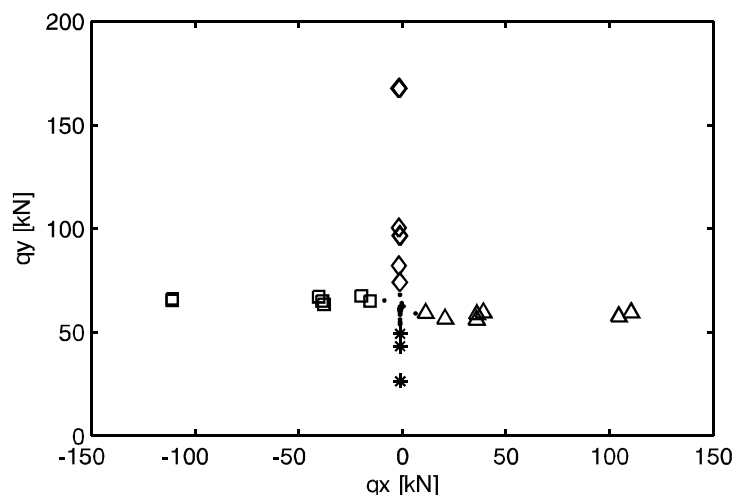


Fig. 6. Clusters in reaction space; training examples, bearing 2

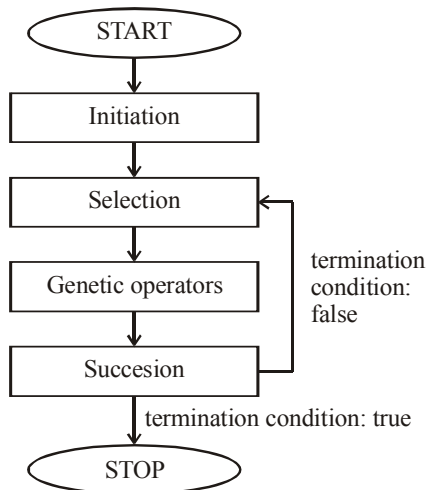


Fig. 7. Scheme of the evolutionary algorithm

4.2. Quality assessment of classifiers

Quality of classifiers was assessed in 4 ways:

- relative number of correctly classified examples:

$$O_1 = \frac{l_1}{l_w} \cdot 100\%, \quad (2)$$

l_1 – number of correctly classified examples,

l_w – number of all classified examples,

- relative number of error “unserviceability undetection”:

$$O_2 = \frac{l_2}{l_w} \cdot 100\%, \quad (3)$$

l_2 – number of examples classified to class C, which should be classified to classes L, R, U or D,

- relative number of error “non-existent unserviceability detection”:

$$O_3 = \frac{l_3}{l_w} \cdot 100\%, \quad (4)$$

l_3 – number of examples classified to classes L, R, U or D, which should be classified to class C,

- relative number of error “a wrong direction of bearing load vector displacement detection”:

$$O_4 = \frac{l_4}{l_w} \cdot 100\%, \quad (5)$$

l_4 – number of examples classified to one of the classes L, R, U or D, which should be classified to another of these classes.

5. SELECTED RESULTS OF THE RESEARCH

The classifiers for each of 7 bearings was identified based on the selected relevant diagnostic signals features. The values of classifiers assessment is shown in tab.1.

The best results was achieved for $f_n=5$ – the range of f_n variation was assumed correctly.

Tab.1. Quality assessment of the obtained classifiers

f_n	O_1	O_2	O_3	O_4
4	68,3%	4,9%	7,6%	19,2%
5	83,0%	7,1%	2,7%	7,1%
6	77,2%	6,3%	8,5%	8,0%
7	75,0%	6,3%	8,9%	9,8%
8	69,6%	6,7%	11,6%	12,1%
9	49,1%	5,8%	19,6%	25,4%
10	76,3%	6,3%	6,3%	11,2%

6. SUMMARY

Good quality of obtained classifiers certify the correctness of the elaborated method of identification of the first stage of the inverse diagnostic model.

The method of identification of the second stage of the described diagnostic model was developed in PhD Thesis of J. Wojtusik [7].

The researches described in the paper completed within a framework of PhD Thesis [6].

REFERENCES

- [1] Arabas J.: *Lectures on Genetic Algorithms* (in Polish), WNT, Warszawa, 2001.
- [2] Cholewa W.: *Diagnostic object models*. [in:] Kiciński J. (ed.): *Modelling and diagnosis of mechanical, aerodynamic and magnetic interactions in turbomachinery*, (in Polish) IMP PAN, Gdańsk, (2005).
- [3] Goldberg D. E.: *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley Publishing Company, Inc., 1989.
- [4] Kiciński J. (ed.): *Modelling and diagnosis of mechanical, aerodynamic and magnetic interactions in turbomachinery*, (in Polish) IMP PAN, Gdańsk, (2005).
- [5] Michalewicz Z.: *Genetic Algorithms + Data Structures = Evolutionary Programs*. Springer Verlag, Berlin, (1992).
- [6] Urbanek G.: *Application of evolutionary algorithms to identification of diagnostic inverse models*, (in Polish) PhD Thesis, Silesian University of Technology, Gliwice, 2006.
- [7] Wojtusik J.: *Diagnostic Multi-Models of Rotating Machineries*, (in Polish) PhD Thesis, Silesian University of Technology, Gliwice, 2006.



PhD. Eng. **Grzegorz URBANEK** is an assistant in the Department of Fundamentals of Machinery Design of Silesian University of Technology in Gliwice. His main fields of interest are heuristic methods, technical diagnostic and CAD.

THE INFLUENCE OF SOME CONTROLLING PARAMETERS ON THE EFFICIENCY OF AN IMMUNE ALGORITHM

Agnieszka KIMACZYŃSKA*, Sebastian KIMACZYŃSKI**, Ryszard ROHATYŃSKI***

University of Zielona Góra, Department of the Management,
Podgórna 50, 65-001 Zielona Góra,

*A.Kimaczynska@wz.uz.zgora.pl, ** S.Kimaczynski@wz.uz.zgora.pl, *** R.Rohatynski@wz.uz.zgora.pl

Summary

In the paper the influence of the probability of the mutation and size of population on the efficiency of the immune algorithm is considered. The influence was investigated on a simple knapsack problem example. A remarkable lack of the results repeatability was observed. To quantitative description of this phenomenon statistical measures were applied. The best combinations of the parameter values of the immune algorithm were found by means of the full plan of numerical experiments.

Keywords: immune algorithm, knapsack problem, efficiency.

WPLYW PARAMETRÓW STERUJĄCYCH NA EFEKTYWNOŚĆ ALGORYTMU IMMUNOLOGICZNEGO

Streszczenie

W artykule opisano badania wpływu prawdopodobieństwa mutacji i liczebności populacji na skuteczność algorytmu immunologicznego. Jako przykładowe zadanie systemu immunologicznego wybrano problem plecakowy. Zaobserwowano brak powtarzalności wyników. Do opisu tego zjawiska zastosowano ujęcie statystyczne. Doświadczalnie określono najlepszą kombinację wartości parametrów algorytmu immunologicznego.

Słowa kluczowe: algorytm immunologiczny, problem plecakowy, efektywność.

INTRODUCTION

In an earlier paper an investigation of genetic algorithms efficiency in task of a mechanism optimization was presented [1]. Algorithms were tested at different values of their main controlling parameters: crossover and mutation probabilities and the size of population. As a result of numerous computational experiments values of these parameters which assure the best efficiency have been obtained. The iteration numbers to finding the best solution and their statistical dispersion have been accepted as the measures of efficiency. In the present paper the similar concept has been assumed for the immune algorithms investigation.

Artificial immune system performance depends significantly on two parameters: the population size and the mutation probability. Proper selection of these parameters is a prerequisite of the algorithm convergence and achievement of correct results. The large number of population causes long time of calculations (i.e. large number of iterations), while too small one can result in incorrect solutions. On the other hand, too high probability of the mutation can result in disruption of the search direction, whereas too low one causes the convergence of the evolution process extremely slow.

Consequently, there must be optimum values of the population size and mutation probability for

a certain class of optimization problems. The paper aims at the demonstration of this statement on a relative simple example.

1. OUTLINE OF THE ALGORITHM AND PROBLEM TO BE SOLVED

a. Immune algorithm

Artificial immune [2] is relatively new discipline which started about a decade ago. Nevertheless, it has gained many successful applications, e.g. in diagnostics of computer viruses [3], in security systems of offices, in recognizing signals warning about a potential failure in aeroplanes, in expert systems for advising physicians in diagnosing diseases from symptoms [4], and many other. Artificial immune systems have been inspired by natural immune systems of living organisms. The main function of a natural immune system is protection of the organism from toxic cells and pathogens. This is performed by a two-stage process: (i) the detection of pathogens and, subsequently, (ii) their efficient elimination. The detection consists in recognition whether the particular cell is a familiar one or a stranger. The stranger can be undesirable virus, a bacterium, or even a degenerated cell. Lymphocytes of B-type and T-type are responsible for the pathogens identification. After having the pathogens detected

the lymphocytes react sending antibodies to neutralize menacing cells. The lymphocytes which have been activated are subjected to cloning. Thanks to this they remember pathogens and the next time they react to the pathogens more quickly. In order to increase the efficiency of the lymphocytes' mission they are also subjected to genetic operations: crossover, inversion, and mutation [2].

b. Knapsack problem

The classical knapsack problem is an abstraction of many real problems occurring in logistics, in digital coding, in aiding managers to undertake decisions, etc. [6]

The conventional formulation of this problem looks as follows [5]:

There is number n of objects (parcels, goods, etc.), each having a certain size w_i and a particular value c_i . There is also given a container („the knapsack”) of a capacity W . The point of the task is to place the objects in the knapsack with some constraints: to maximize the sum of the values of objects packed into knapsack but not exceeding the capacity of the knapsack.

Formally the knapsack problem is recorded as bellow:

$$f(x) = c_i x_i \rightarrow \max$$

$$w_i x_i \leq W$$

Where:

$f(x)$ – the objective function,

c_i – value of item of i 's number,

x_i – binary variable which is '1' if an item is packed to the knapsack and '0' otherwise.

w_i – a characteristic feature (size, volume, weight) of the item number i ,

W – the entire capacity of the knapsack.

2. PROBLEM FORMULATION

The research was aimed at finding the influence of the immune algorithm controlling parameters, i.e. the population size and the mutation probability, on the probability of achieving correct results in various number of iterations. For this purpose a knapsack problem was formulated and for 25 combinations of the parameters with 15 repetitions of each combination, numerical experiments were carried out.

a. The knapsack problem to be solved

The exemplary knapsack problem was formulated as follows:

We have 15 things which have to be packed into a knapsack of the capacity up to 70 units.

Weights of things (in some units) are given in order below:

[1 2 3 4 5 6 7 8 9 10 11 12 13 14 15]

Values of things (in some units) are given in order below:

[8 1 12 5 11 9 4 4 7 15 4 6 10 3 2]

The optimal solution of this problem is known. It is shown in a vector form:

[1 1 1 1 1 1 1 1 1 0 1 1 0 0 0],

where '1' means that the thing has been packed and '0' means the given thing has been ommitted.

Optimal solution of this knapsack problem is known. The optimal value of the knapsack is 54, and its weight is 68 units.

b. Plan of experiments

The full plan of experiments for the considered problem was assumed that resulted in 25 combinations of the population size and mutation probability values. These combinations are listed in table below:

Table 1 Combinations of the population size and mutation probability values.

No. of the combinations	Size of population	Mutation probability
1	10	0,2
2	10	0,3
3	10	0,5
4	10	0,7
5	10	0,9
6	20	0,2
7	20	0,3
8	20	0,5
9	20	0,7
10	20	0,9
11	50	0,2
12	50	0,3
13	50	0,5
14	50	0,7
15	50	0,9
16	75	0,2
17	75	0,3
18	75	0,5
19	75	0,7
20	75	0,9
21	100	0,2
22	100	0,3
23	100	0,5
24	100	0,7
25	100	0,9

3. RESULTS OF COMPUTER SIMULATION

The immune algorithm which was used for the experiment was adapted from [2]. The original algorithm was converted for Pascal on Matlab. A simple procedure of sorting the objective function values was added. Also, there was changed the way the parameter M (which is responsible for number of clones) values was modified. The parameter M decreased by 1, if during 20 iterations there was no improvement of the value of the objective function. When the value of the parameter M reached 0, the

algorithm was programmed to stop. This stop condition was meant to prevent repetition of the algorithm iterations which would not bring any improvements of the objective function value. It turned out that often the algorithm stopped prematurely, not reaching the optimum result.

For the statistical evaluation of the algorithm efficiency, calculations for each combination were repeated 15 times.

The following data were recorded:

- the number of the iteration,
- the weight of the knapsack obtained for each member of the population,
- the value of the knapsack obtained for each member of the population.

These data were used for calculation of the average numbers of the iterations, the average weights of the knapsack, the average values of the knapsack, and for evaluation of the scattering of the results (by the standard deviation), for each combination of input data. The results are shown on figures 1 to 5.

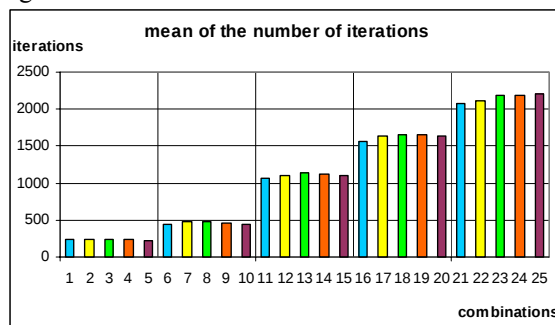


Figure 1. The mean of the number of the iterations

The mean of the number of algorithm iterations is shown on figure 1. The number of the iterations depends significantly on the size of the population. It increases with the population size.

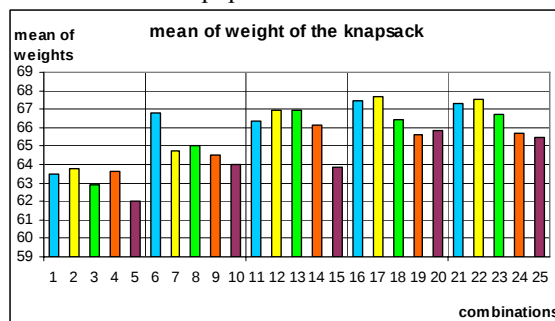


Figure 2. The mean of the weight of the knapsack

The mean of the weight of the knapsack as shown on figure 2, is strongly dependent on quantity and value of things packed to the knapsack: the greater quantity and value of things packed the greater weight of the knapsack. The maximum weight of the knapsack - 70 units, was never violated.

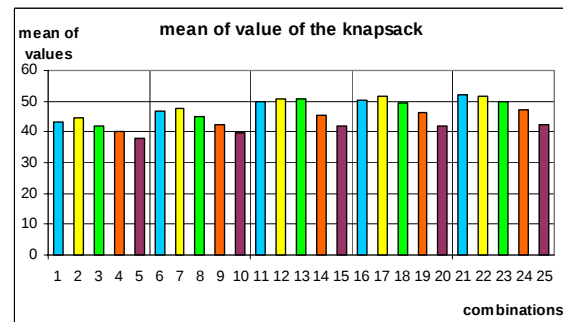


Figure 3. The mean of the value of the knapsack

Figure 3 depicts the mean of the knapsack values. For a given population size all members of the populations, which were close to the optimum, were obtained at the mutation probability 0,3. These parameter combinations which yielded the average value of the knapsack greater than 50 (the best solution is 54 units) deserve special attention. Numbers of such the combinations are 12, 13, 16, 17, 21, 22, 23.

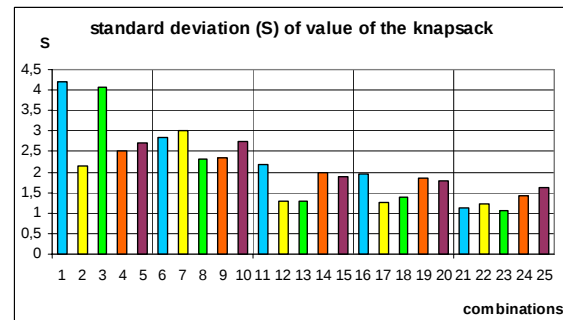


Figure 4. The standard deviation of the knapsack

Except of the mean values of the knapsack, also the standard deviations of the values were evaluated. It is shown on figure 4. It can be seen that the standard deviation of the values is fairly high. This phenomenon of strong randomization of results is a characteristic feature of the immune algorithms. Of course, the lesser deviation, the better. Thus, the special attention deserve these parameter settings which yield the standard deviation below 1.5; namely the combinations 12, 13, 17, 18, 21, 22, 23. These combinations resulted in a relative stability of the results of the investigated immune algorithm. For the user it means a relatively high credibility of the results i.e. it is not necessary to run the algorithm many times.

Because relations between the mean values of the knapsack and the standard deviations are of great importance so they are demonstrated on fig. 5.

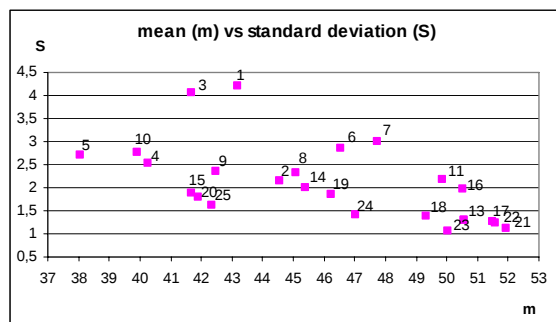


Figure 5. The mean of the results versus their standard deviation

This figure confirms that the best settings are the combinations 21, then 22 and 17. They have population size 100 or 75, and the mutation probability equal to 0.2 or 0.3. For these parameter values, however, amount of the iterations is quite large, which means a long calculation time. For the population of 75 members the algorithm needs about 1600 iterations, whereas at 100 members – as many as 2100 iterations. However, this number of the population members and the probability of mutation equal 0.2 or 0.3 guarantee receiving of the most credible results.

There is explicit need for the trade off between size of population, time of calculation and probability of getting optimal solution when using artificial immune algorithms. Recognition of this fact opens an interesting field for further investigations.

4. CONCLUSIONS

Experiments carried out by means of the immune algorithm on the knapsack problem have shown that there are some combinations of the values of the algorithm controlling parameters ensuring its relatively highest efficiency. Our research confirmed that in the artificial immune systems the population size plays the same role as in natural ones; the greater, however, it is the greater is the number of iterations needed for the algorithm to arrive at the best solution. On the other hand, the probability of the mutation has not shown monotonical impact on the efficiency of calculations; there seem to be a range of the best mutation probability values. Best operation of the immune algorithm was observed

when probability of the mutation was about 0.3. It should be noted that this figure is much greater than the one recommended for the genetic and evolutionary algorithms.

Any generalization of the results presented in this paper should be taken with caution. Different kinds of problems have own specificity which certainly would require another set of values of controlling parameters, and perhaps even another set of controlling parameters. However, it is very probable that for any specific kind problem there exist optimum values of some parameters. The awareness that such values exist creates a promising area for further research on the more efficient use of the artificial immune algorithms in searching for innovative solutions of problems which are difficult for other methods.

5. REFERENCES

- [1] R. Rohatyński, T. Klekiel, K. Białas-Heltowski: *Investigation of evolutionary algorithm effectiveness in optimal synthesis of certain mechanism*, *Evolutionary methods in mechanics*, Kluwer Academic Publishers, Cracow 2004, pp. 13 – 22.
- [2] S. T. Wierchoń: *Sztuczne Systemy immunologiczne – teoria i zastosowania*, EXIT, Warszawa 2001.
- [3] J. O. Kephart: *A biologically inspired immune system for computers*, *Artificial Life IV*, MIT Press 1994, pp. 130 – 139.
- [4] M. Bereta, T. Burczyński: *Hybrid immune algorithm for feature selection and classification of ECG signals*, *AI-METH 2005*, Gliwice 2005, pp. 25 – 28.
- [5] M. M. Sysło: *Algorytmy*, WSiP, Warszawa 2002;
- [6] P. Kamiński: *Formulation objective function of the decision – making problem in ship power plant*, *International Conference Design 2006*, Dubrovnik, vol 1.
- [7] M. Święcicki, W. Wais, P. Wais, L. Nowak: *Metody optymalizacji immunologicznej w rozwiązywaniu problemów kombinatorycznych należących do klasy NP*, *Zastosowania teorii systemów*, WIMiR AGH, Kraków 2005, pp. 349–363.

APPLICATION OF EQUIVALENT SOURCE METHOD TO IDENTIFICATION OF MODELS OF NOISE SOURCES IN AUTOMOTIVE AIR CONDITIONING SYSTEMS

Anna BZYMEK

Silesian University of Technology, Department of Fundamentals of Machinery Design
Konarskiego 18a Street, 44-100 Gliwice, Poland, e-mail: anna.bzymek@polsl.pl

Summary

The paper deals with the experimental analysis of noise sources in Automotive Air Conditioning Systems (AACS), especially a part of the system – Heating Ventilating Air Conditioning (HVAC) unit. Presented approach was based on Equivalent Source Method (ESM) which is an inverse method of identification of models of noise sources. In order to model the acoustic pressure above the main outlet of the HVAC unit, a model consisting of a few equivalent sources was elaborated.

Keywords: equivalent source method, HVAC unit, acoustic pressure.

ZASTOSOWANIE METODY EKWIWALENTNYCH ŹRÓDEŁ W CELU IDENTYFIKACJI MODELI ŹRÓDEŁ HAŁASU W UKŁADZIE KLIMATYZACJI SAMOCHODOWEJ

Streszczenie

Artykuł przedstawia eksperymentalną analizę źródeł hałasu w systemie klimatyzacji samochodowej, szczególnie jej części – w module HVAC. Prezentowane podejście bazuje na metodzie ekwiwalentnych źródeł w celu identyfikacji modeli źródeł hałasu. W artykule przedstawiono główne założenia metody, opisano przebieg i sposób realizacji badań. Uzyskane wyniki opatrzone komentarzem.

Słowa kluczowe: metoda ekwiwalentnych źródeł (ESM), moduł HVAC, ciśnienie akustyczne.

INTRODUCTION

Nowadays, cars are most common means of transport, so that demands dealing with driving safety become higher. During the maintenance of a car, passengers are exposed to a harmful influence of noise coming from several sources. The main source of noise in vehicles are: engine, tyres and air conditioning system.

Automotive Air Conditioning System (AACS) consists of numerous components. It must be stressed that all of them generate some kinds of noises. The part of AACS, which is placed close to the driver, is the Heating, Ventilating and Air Conditioning (HVAC) unit (Fig.1). Therefore, the first step of decreasing noise generated by the system is to limit noise generated by this unit. Because of that manufacturers focused their attention on the investigations devoted to this problem. In order to measure noise and vibration generated by the HVAC unit and model such a noise source, special methods are required to be applied [10]. The model lets us to investigate different parameters of operation of the unit, and finally obtained results can help with decrease of a level of the noise. The main goal of the approach

presented in the paper consists in identification of a model of sound pressure emitted by the HVAC unit. In the case of presented approach Equivalent Source Method was employed.

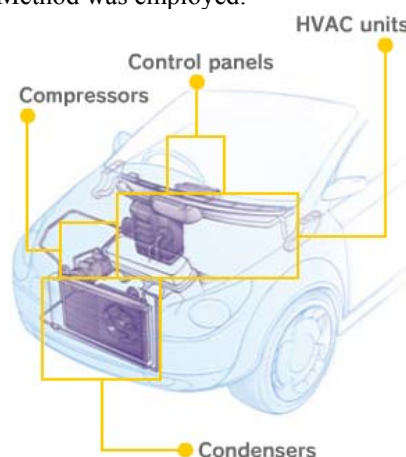


Fig. 1. Automotive Air Conditioning System with the HVAC unit [11]

In the paper results of MSc Thesis, completed in the Department of Fundamentals of Machinery Design at the Silesian University of Technology at Gliwice, are presented. The measurements and investigations were conducted by the author during

her 4 month fellowship at Universite de Technologie de Compiègne in France. All applied procedures were elaborated in MATLAB.

1. EQUIVALENT SOURCE METHOD

The ESM theory is mainly applied to acoustic and vibroacoustic analysis [1, 5, 6, 7], as well as in geophysical research [4]. The main assumption of the method is to replace an investigated object with a set of equivalent sources represented by monopoles. In order to model the object, measurement of its acoustic radiation is necessary to be done. The process of recording and analysing data was described in the further part of the paper and in [1].

The basis of Equivalent Source Method is eq. 1.

$$G(M, S) = \frac{i\omega\rho_0}{2\pi r} \cdot e^{-ikr}, M = 1, \dots, m; S = 1, \dots, s \quad (1)$$

Where: $G(M, S)$ is a matrix containing values of acoustic pressure generated by each source s and measured by each microphone m , in function of m - s distance (r). An exemplary microphone - source configuration, used during undertaken investigation, was presented in Fig. 2.

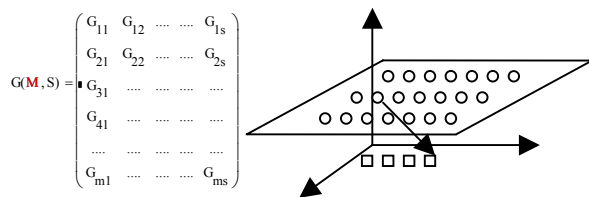


Fig. 2. An example of the m - s configuration, and $[G]$ matrix

Each single m - s configuration requires that calculation of acoustic pressure for each source (monopole) and each microphone are to be performed. Let's assume that the model consists of S monopoles and M microphones and $M \geq S$. For example, G_{11} indicates the value of pressure generated by the 1st modeled monopole measured by the 1st microphone, and G_{23} indicates the pressure generated by the 3rd monopole measured by the 2nd microphone.

For each m - s configuration, and simultaneously for each new $[G]$ matrix, the condition number (CN) has to be calculated. This parameter is represented by a numerical value, which determines correctness of inversion of the matrix. This value entails that the matrix is well or ill-conditioned [8] to be inverted. The inversion of $[G]$ is possible when $m=s$ (the number of microphones is equal to the number of sources). Otherwise the matrix is pseudoinverted. During the investigation the value of CN parameter was obtained with the use of MATLAB. The analyses of variation of the CN value in frequency function were made and described in [1]. The results were presented and discussed in the paper in section 2.

ESM theory assumes that it is possible to build the model of acoustic radiation of an investigated object [2] on the basis of a vector of acoustic pressure measured around the object. A model consisting of a set of equivalent sources, is characterized by its parameters such as optimal source strength q_{opt} , and acoustic pressure calculated p_{calc} . The values of the parameters were assigned with the use of (2,3).

$$q_{opt} = [G]^+ \cdot p_{mes} \quad (2)$$

$$p_{calc} = [G] \cdot q_{opt} \quad (3)$$

Where p_{mes} is a vector of acoustic pressure measured, p_{calc} is a vector of acoustic pressure calculated, $[G]^+$ is the pseudoinverted matrix of G and q_{opt} is a vector of optimal strength of equivalent sources. At first the vector of p_{mes} , for a given frequency (the same as the frequency selected in order to determine the matrix $[G]$), was used in order to compute q_{opt} , which is the first parameter of the model. Then the p_{calc} (the second parameter) was obtained, and finally, in order to identify the quality of the model, the error e (eq. 4) between the pressure calculated and measured was estimated:

$$e = \frac{\|p_{mes} - p_{calc}\|}{\|p_{calc}\|} \quad (4)$$

where: e - error
 $\|p_{mes} - p_{calc}\|$ - norm ($p_{mes} - p_{calc}$).

These operations were made with the use of procedures elaborated in MATLAB. Obtained results were visualised as 3D plots. Investigations and their results were presented in the next section.

2. DESCRIPTION OF THE LABORATORY STAND

The HVAC unit was investigated in the anechoic chamber. The acoustic pressure above the HVAC unit was measured by means of a microphone mounted on a special frame (Fig. 3) and moved along the 5x11 grid [1]. Calculations obtained in further part of this work are expressed as frequency functions. Because of complex structure of the HVAC, signal analysis at this stage of the research was limited to one part of the unit. This part was the main outlet, called also slide door.



Fig. 3. Measurement set

In the experiment nine different tests were conducted under three different states of HVAC power supply (4V,7V,12V) and m - s distances (100mm, 50mm, 5mm.). Frequency range of recorded signals was 0-5 kHz, with sampling frequency equal to 6,25 Hz.

3. RESULTS OF A SELECTED EXAMPLE

The method was tested with taking into account numerous configurations. In the paper only one of them, characterized by 7V and 100mm for 1kHz, was presented. Results of other experiments were presented in [1].

The investigations of 2D and 3D m - s configurations were performed [1] in order to check how the number of sources and distances between sources and microphones influences the evolution of condition number of the matrix [G]. It was proven that the condition number increases when a number of sources and microphones also increase. At the same time, it increases when the distance between sources and microphones increases. As the result of described discussion [1], the model consisting of 5 sources and 5x11 microphone grid was used in further investigations.

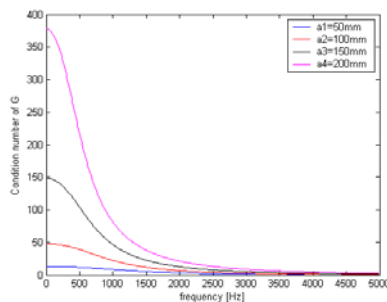


Fig. 4. Condition number curve in frequency function

The values of acoustic pressure are presented in the form of a surface (Fig. 5), whose dimensions are equal to dimensions of the HVAC outlet. It is important that data acquisition was performed for such the configuration of slide door where one in the center was open. It means that the acoustic pressure was higher in the center and it is noticeable in (Fig. 5). Such the way of presenting results allows us to compare them. The shape of the surfaces of pressure calculated (Fig 5 right) and measured (Fig 5 left) are similar. However, in the case of the surface of acoustic pressure calculated, there is an area with higher values of pressure, which are not present in the case of acoustic pressure measured. One can state that it can be caused by calculation errors. It have to be mentioned that in this case only 5 sources were taken into account.

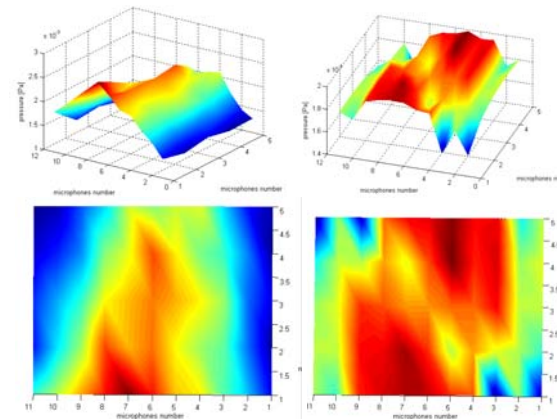


Fig. 5. P_{mes} (left) P_{calc} (right) surfaces

In Fig. 6 (left) amplitudes of equivalent source strengths (which were calculated on the basis of [G] and p_{mes}) in comparison with surface of pressure measured were presented. It is noticeable that the shape of the surface (right) corresponds to a value of computed equivalent source strength (left).

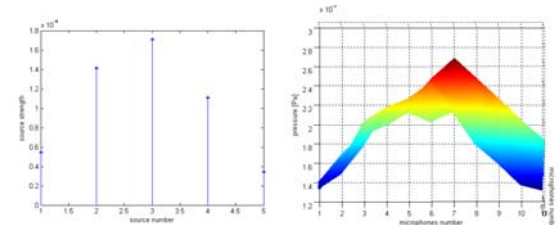


Fig. 6. Optimal source q_{opt} strength (left) in comparison to pressure measured p_{mes} (right)

The mean values of p_{mes} and p_{calc} in frequency function were shown in Fig. 7 (left). It is noticeable that both of these courses are very close to each other. It means that the error e (right) is really small and quality of the model is acceptable.

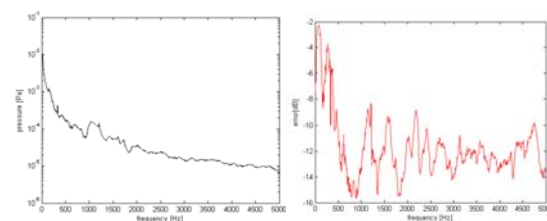


Fig. 7. Mean value of pressures (left) and error (right) in frequency function

On the basis of these plots one can state that this small difference between pressures lets obtain the correct results in this frequency. Obtained results proved that the method can be used as the basis of identification of models consisting of monopoles.

4. CONCLUSIONS

The main problem of the application of ESM method was to determine the assumption of the m - s configuration and criteria for numerical simulation, which let us obtain the best conditioned matrix to

be inverted. The author tested the model containing 5 sources, 5x11 grid of microphones and m -s distance equal to 100 mm. One can state that under these conditions the sound field radiated by the HVAC unit can be well reconstructed. The grid consisting of 55 microphones is sufficient to properly scan the pressure above the outlet and 5 monopoles are sufficient to represent a sound field emitted by the real object for low frequencies. In the case of lower frequencies, the error reaches lower value. It means that the pressure calculated is closer to the measured pressure. In highest frequencies the error is bigger and it is noticeable that surface of p_{mes} is not correctly modeled by the p_{calc} surface.

For different frequencies different kinds of noises can be detected. Since that shapes of surfaces was different. The study of the kind of noise at this level of investigations was not conducted and it should be distinguished within further stages of the investigations.

Analyzing obtained results one can state that some mistakes can be caused by calculation errors. It has to be mentioned that in this case only 5 sources were taken into account. One predicts that the more applied monopoles the more similar surfaces to the one obtained on the basis of measurements. However, it is not unquestionable, since as it was shown, when the number of sources increases the matrix $[G]$ becomes bigger and worse conditioned to be inverted. It also means that in this case the error can be bigger. However, it was not verified and it is a problem required to be investigated within the framework of further experiments.

REFERENCES

- [1] Bzymek A.: *Identification of models of noise sources in automotive airconditioning system*. MSc Thesis, Silesian University of Technology at Gliwice, 2005.
- [2] Batko W., Dąbrowski Z., Engel Z., Kiciński J., Weyna S.: *Nowoczesne metody badania procesów wibroakustycznych*, Wydawnictwo Instytutu Technologii Eksploatacji PIB, Radom 2005.
- [3] Bécot F. X., Kropp W.: *Iterative Equivalent Sources model of the tyre emission over impedance surfaces*, Euronoise Conference, Naples 2003.
- [4] Dymont, J. Arkani-Hamed J.: *Equivalent Source magnetic dipoles revisited*, Geophysical research letters, vol 25,no11, pages 2003-2006,1998.
- [5] Johnson M. E., Elliott S. J., Baek K - H., Garcia-Bonito J.: *An equivalent source technique for calculating the sound field inside an enclosure containing scattering objects*, Journal of the Acoustical Society of America, vol.104,no3 (1998) p1221-1231.
- [6] Laine A. L., Johnson M. E.: *Development of an efficient binaural simulation for the analysis of structural acoustic data* Summary report of the research performed under the NASA Langley Research Center, NCC1-01029; 2002.
- [7] Moorhouse A. T.: *Prediction of the sound output of an assembled machine driven by an internal electric motor*. (<http://www.enableacoustics.org.uk/Expertise/IOA%20paper%20Andy%20Moorhouse.pdf>).
- [8] Vieira Phino M. E., J. R de Franca Arruda *On the use of the Equivalent source method for nearfield acoustic holography*, ABCM Symposium Series in Mechatronics Vol 1 p.590-599, 2004.
- [9] Piscoya R., Ochmann M., Brick H., Koltzsch P., *Modelling of the Combustion noise by means of the equivalent source method (ESM)*, CFADAGA2004/126.
- [10] Schillemeit B., Cucuz S., Article *Comparison of experimental NVH analysis techniques on Automotive HVAC systems*, SAE 2002 World Congress, Detroit, Michigan
- [11] <http://www.zexel.co.jp/english/product/item/heater.html>.



Anna BZYMEK is PhD student in the Department of Fundamentals of Machinery Design at Silesian University of Technology at Gliwice. She deals with methods of acoustic analysis and methods of image processing and recognition.

ZASTOSOWANIE ANALIZY MODALNEJ DO DETEKCJI USZKODZENIA WIADUKTU

Krzysztof MENDROK

Akademia Górniczo – Hutnicza, Katedra Robotyki i Mechatroniki
Al. Mickiewicza 30, 30-059 Kraków, fax. 012 634 35 05, e-mail mendrok@agh.edu.pl

Streszczenie

W pracy przedstawiono główne problemy eksploatacyjnej analizy modalnej dużych obiektów budownictwa lądowego. Skupiono się na trudnościach związanych ze specyfiką obiektów tj. gabarytami i niekorzystnymi wpływami otoczenia. Opisano badania modalne przeprowadzone na wiadukcie drogowym w celu określenia jego stanu technicznego. Wiadukt ten nie był wcześniej badany więc nie dysponowano modelem dla obiektu w stanie nieuszkodzonym co znacznie utrudniło ocenę stopnia jego uszkodzenia.

Słowa kluczowe: eksploatacyjna analiza modalna, ocena stanu konstrukcji.

DAMAGE DETECTION OF THE VIADUCT WITH USE OF THE MODAL ANALYSIS

Summary

In the paper the main problems of the operational modal analysis of large civil engineering objects are presented. Author focused on the difficulties which come from the specificity of the objects, that is the large dimensions, negative influence of the environment. The modal analysis performed on the road viaduct for its damage assessment is described. There are no undamaged model data available, what made the damage assessment particularly difficult.

Keywords: operational modal analysis, damage assessment.

1. WPROWADZENIE

W dzisiejszych czasach dąży się do stałego monitoringu odpowiedzialnych obiektów inżynierii lądowej takich jak mosty, wysokie budynki, maszty czy kominy. Dotychczas stosowana inspekcja wizyjna wykonywana okresowo przez odpowiednie służby jest narzędziem o niewielkiej skuteczności. Ze względu na okresowość inspekcji istnieje niebezpieczeństwo nie wykrycia inicjacji i wzrostu uszkodzenia w okresie między przeglądami, a także brak możliwości stwierdzenia powstania uszkodzeń niewidocznych z zewnątrz. Nowoczesne systemy monitoringu obiektów inżynierii lądowej bazują na metodach opartych o zmianę własności dynamicznych (opartych na modelu). Do najczęściej wykorzystywanych w detekcji uszkodzeń parametrów dynamicznych należą parametry modelu modalnego obiektu tj. częstotliwości drgań własnych [1], współczynniki tłumienia modalnego [2] oraz wektory modalne [4], [5], [6]. Niejednokrotnie analizuje się też zmiany w charakterystykach częstotliwościowych monitorowanych obiektów [3]. W metodach wykrywania uszkodzeń bazujących na modelu modalnym, porównuje się wybraną wielkość zidentyfikowaną dla obiektu w stanie nieuszkodzonym, z wielkością bieżącą. Jeżeli wystąpią różnice wartości, istnieje podejrzenie że w układzie powstało uszkodzenie. W nowoczesnych

systemach monitoringu dla mostów jako wskaźnik uszkodzenia wykorzystuje się głównie zmianę częstotliwości drgań własnych obiektu. Podejście takie jest wygodne, gdyż wymaga zastosowania niewielkiej liczby czujników, ale obarczone jest dużym ryzykiem wystąpienia błędów związanych z wpływem warunków zewnętrznych. Zmiana tychże, może spowodować przesunięcie częstotliwości drgań własnych porównywalne z wywołanym przez wystąpienie w układzie uszkodzenia. Na przykład w wysokich letnich temperaturach sztywność obiektu maleje, a więc częstotliwości drgań własnych również się obniżają. Zjawisko to jest szczególnie wyraźne dla mostów drogowych z grubą warstwą asfaltu. Jego sztywność najbardziej zmienia się wraz z temperaturą. Podobny wpływ na własności dynamiczne ma zmiana wilgotności powietrza. Konstrukcja betonowa absorbuje wilgoć i jej masa rośnie, a co za tym idzie częstotliwości drgań własnych maleją. Dlatego, aby skutecznie i wiarygodnie wykorzystać metodę detekcji uszkodzeń opartą o zmiany częstotliwości drgań własnych, należy najpierw określić wrażliwość obiektu na zmiany warunków zewnętrznych. W niektórych pracach podejmowane są próby analitycznego wyznaczania wpływu otoczenia na zmianę parametrów modalnych obiektu [1]. Najczęściej jednak stosuje się tablicę, w której zestawia się własności obiektu zidentyfikowane dla

różnych temperatur i wilgotności powietrza. Taka tablica jest unikatowa i sporządzana niezależnie dla każdego obiektu w wyniku szeregu eksperymentów. Bardziej zaawansowaną metodą wyeliminowania wpływu warunków pogodowych na skuteczność działania systemu monitoringu jest zastosowanie tzw. filtru środowiskowego [7]. Jest to najczęściej model autoregresyjny o ruchomej średniej (ARMA) zidentyfikowany na podstawie serii danych eksperymentalnych. Bez względu jednak na rodzaj wybranej metody detekcji uszkodzenia, niezbędne jest posiadanie modelu dla obiektu nieuszkodzonego. Danych takich nie ma dla większości omawianych obiektów. W Polsce sporą grupę obiektów drogowych typu mosty, wiadukty, estakady stanowią jeszcze konstrukcje z lat 80-tych, 70-tych, a niekiedy i starsze. Z drugiej strony w ostatnich 15 – 20 latach obserwuje się gwałtowny wzrost ilości pojazdów na drogach. Zwłaszcza liczba dużych ciężarówek – TIR'ów jest teraz nieporównywalnie wyższa. Wspomniane obiekty drogowe nie były przewidziane na taką intensywność użytkowania i dlatego mogą stanowić potencjalne zagrożenie dla użytkowników. Dla obiektów takich trudno jest przeprowadzić ocenę stanu technicznego z wykorzystaniem metod diagnostycznych opartych na modelu. Powodem jest tutaj wspomniany brak danych dla obiektu w stanie nieuszkodzonym. Jeden z takich przypadków został opisany w tym artykule. Jest to wiadukt drogowy, który na podstawie wizualnej oceny został zakwalifikowany do remontu. Badania modalne miały określić jak pilnie ów remont należy przeprowadzić.

2. PROBLEMY ANALIZY MODALNEJ OBIEKTÓW MOSTOWYCH

Dla wyznaczenia parametrów modelu modalnego obiektu w praktyce najczęściej wykorzystywana jest eksperymentalna analiza modalna (EMA) [8]. Istnieje szereg algorytmów EMA realizowanych zarówno w dziedzinie czasu jak i częstotliwości [8]. Wymagają one przeprowadzenia czynnego eksperymentu modalnego. W eksperymencie takim mierzone są odpowiedzi układu, najczęściej w postaci przyspieszenia drgań na zadane, mierzone wymuszenie. Wymuszenie realizowane jest najczęściej przy pomocy wzbudnika drgań elektrodynamicznego, hydraulicznego, lub młotka modalnego. Na podstawie zmierzonych przebiegów odpowiedzi i sił wymuszających estymowane są widmowe funkcje przejścia bądź odpowiedzi impulsowe. Dają one stosunkowo dokładne estymaty modelu modalnego, postaci drgań własnych (wektory modalne) uzyskane w ich wyniku są skalowane. Oznacza to, że przy ich pomocy można dokonać syntezy widmowych funkcji przejścia, przewidywać odpowiedź układu na zadane wymuszenie. Model modalny ze skalowanymi postaciami drgań własnych nadaje się też do symulacji modyfikacji strukturalnych

własności obiektu. Kolejną zaletą EMA jest, fakt, iż przy odpowiednim dobraniu wymuszenia można uzyskać wszystkie postacie drgań własnych w określonym paśmie częstotliwości. Wady metod EMA wynikają z konieczności zastosowania czynnego eksperymentu z mierzalnym wymuszeniem. Podejście takie wymaga wyłączenia badanego urządzenia z eksploatacji i odizolowania go od innych źródeł wymuszenia. Dla obiektów inżynierii lądowej takich jak mosty jest to bardzo trudne i najczęściej niemożliwe do zrealizowania. Pierwszą niedogodnością jest duża masa i sztywność układu. Skuteczne wymuszenie konstrukcji mostowej wymaga zastosowania bardzo drogiego hydraulicznego wzbudnika drgań dużej mocy. Na świecie istnieje bardzo niewiele tego typu urządzeń i zakup bądź wynajęcie jednego z nich w celu przeprowadzenia kilku pomiarów byłoby ekonomicznie nieuzasadnione. Równie trudne i kosztowne w realizacji byłoby odizolowanie obiektu od innych źródeł drgań. Stwarza to konieczność zatrzymania ruchu nie tylko na moście ale i na otaczających go ulicach, aby drgania gruntu wywołane przejazdem ciężkich samochodów nie przenosiły się na badany obiekt. Praktycznie niemożliwe jest odizolowanie badanego mostu od wpływu wiatru. Drgania wywołane wiatrem mogą w istotny sposób zakłócić czynny eksperyment modalny. Jedyną receptą jest wtedy przeprowadzenie pomiaru w dzień bezwietrzny, a to z kolei stoi w sprzeczności z zatrzymaniem ruchu, które trzeba planować i zgłaszać w odpowiednich instytucjach z dużym wyprzedzeniem.

Rozwiązaniem powyższych problemów jest zastosowanie eksploatacyjnej analizy modalnej (OMA) [9]. Algorytmy OMA umożliwiają wyznaczenie parametrów modelu modalnego na podstawie danych eksploatacyjnych, to znaczy, że pomiar odbywa się w czasie normalnej pracy obiektu. Rejestrowane są przyspieszenia drgań w wybranych punktach maszyny, jedna z odpowiedzi traktowana jest jako odniesienie, do którego liczone są funkcje korelacji lub wzajemnej gęstości widmowej mocy. Takie podejście umożliwia badanie obiektu w warunkach eksploatacyjnych, przy eksploatacyjnych wymuszeniach. Nie ma więc konieczności stosowania drogich urządzeń wymuszających, ani izolowania obiektu badań od innych źródeł drgań. Wadą OMA jest to, że w wyniku otrzymuje się nieskalowane postacie drgań własnych. Należy także pamiętać, że w wyniku OMA estymowane są jedynie te postacie, które są wzbudzane wymuszeniem eksploatacyjnym. W przypadku badań opisanych w pracy dla celów diagnostycznych istotne były właśnie te postacie, które wymuszane są ruchem pojazdów i wiatrem, a znajomość skalowanych postaci drgań nie była konieczna. Mimo uniknięcia trudności związanych z przeprowadzeniem EMA, w czasie realizacji eksperymentu biernego napotkano na szereg innych utrudnień.

Gabaryty omawianego obiektu są duże (około 100 m długości), w związku z tym zachodziła konieczność stosowania długich, ciężkich przewodów i wielokrotnej zmiany pozycji sprzętu pomiarowego. Mimo zaangażowania dużego zespołu badaczy, alpinistów i samochodu z wysięgnikiem pomiar trwał powyżej 8 godzin. W związku z tym zmieniały się warunki eksperymentu wraz z upływem czasu. Rano i wieczorem temperatura powietrza była niższa o kilka stopni niż w godzinach południowych, ponadto w związku z przesuwaniem się słońca kolejno nagrzewały się różne partie mostu. Również ruch pojazdów w ciągu dnia zmieniał się. Duża ilość samochodów przejeżdżających przez wiadukt jednocześnie zmieniała masę obiektu. Kolejną konsekwencją faktu, iż pomiar odbywał się na obiekcie było narażenie badaczy i sprzętu na opady atmosferyczne. Podczas badań napotkano też na poważne trudności związane z zasilaniem aparatury pomiarowej. Przygotowane akumulatory podczas pomiarów bardzo szybko się rozładowywały. Innym sposobem zasilania w warunkach polowych jest zastosowanie generatora. Należy jednak zwrócić uwagę aby generator był przystosowany do pracy z urządzeniami elektronicznymi, tzn. posiadał odpowiedni układ redukujący poziom zakłóceń.

3. OBIEKT ANALIZY

Obiektem analizy był wiadukt drogowy nad ulicą Brzeźnicką w Bochni. Wiadukt ten jest częścią międzynarodowej drogi E 40 pomiędzy Krakowem i Tarnowem. Ruch na tym odcinku drogi jest szczególnie intensywny. Wiadukt składa się z pięciu przęseł o rozpiętości 18,94 + 19,53 + 19,62 + 19,61 + 18,70 = 96,40 m. Całkowita szerokość wiaduktu wynosi 11,40 m, w tym: jezdnie - 7,00 m oraz pobocza - 2 * 2,00 m. Zdjęcie omawianego obiektu pokazano na rysunku 1.



Rys. 1. Widok badanego wiaduktu

Schematem statycznym wiaduktu jest rama pięcioprzęsłowa. Konstrukcja nośna jest wykonana z prefabrykowanych belek sprężonych o rozpiętości

18,00 m. Nad filarami belki te są uciągłone i połączone z odczepami filarów. W przekroju poprzecznym występuje 8 belek w rozstawie osiowym 1,50 m. Z belkami prefabrykowanymi współpracuje żelbetowa płyta pomostu o grubości 12 ÷ 18 cm.

Podpory wiaduktu są słupowe: przyczółki dwusłupowe, filary jednosłupowe. Wszystkie podpory są posadowione na palach żelbetowych. Wizualna inspekcja wiaduktu określiła jego stan techniczny jako bardzo zróżnicowany. Podczas oceny tej stwierdzono, że w dobrym stanie znajdują się podpory. O wiele gorszy jest stan belek głównych, zwłaszcza skrajnych. Wystąpiły w nich ubytki betonu oraz zarysowania. Bardzo zróżnicowany jest stan żelbetowej płyty pomostu – od złego (wystąpiła konieczność lokalnych wzmocnień) do dobrego.

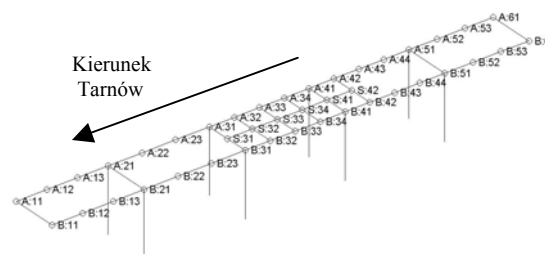
Zróżnicowanie stanu technicznego oraz stopień zagrożenia jego dalszej eksploatacji skłoniły administratora obiektu do przeprowadzenia szczegółowych badań. Ich wyniki są przedstawione w artykule.

4. EKSPERYMENT MODALNY

Eksperyment modalny na przedstawionym wiadukcie przeprowadzono 14 lutego 2006 roku. W czasie pomiarów wykorzystano wymuszenie eksploatacyjne w postaci ruchu pojazdów po jezdni wiaduktu. Sieć punktów pomiarowych składała się z 42 punktów i została podzielona na trzy komponenty (grupy):

- 18 punktów rozmieszczonych wzdłuż lewego skrajnego dźwigara,
- 18 punktów rozmieszczonych wzdłuż prawego skrajnego dźwigara,
- 6 punktów rozmieszczonych wzdłuż czwartego dźwigara od prawej strony na środkowym przęśle.

Przyjętą sieć punktów pomiarowych jest przedstawiona na rysunku 2.



Rys. 2. Rozkład punktów pomiarowych na wiadukcie

Czujniki mocowano do spodniej powierzchni dźwigarów. Z uwagi na wysokość obiektu do pomiaru zostali zatrudnieni dwaj alpinisci i ciężarówka z wysięgnikiem. Grupa czujników zamocowana na wiadukcie została pokazana na rys. 3.

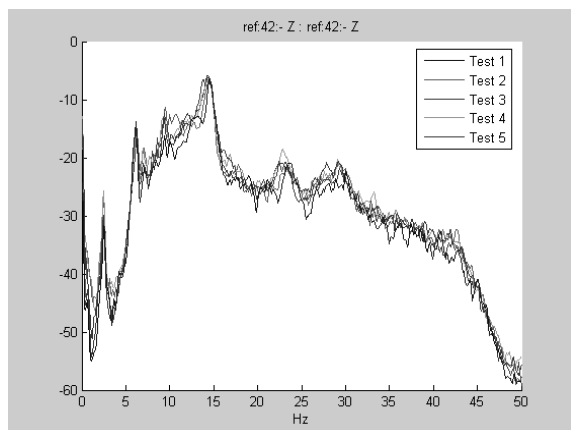


Rys. 3. Czujniki zamocowane na dźwigarze wiaduktu

Rejestrowano 900 sekundowe przebiegi czasowe przyspieszenia drgań z częstotliwością próbkowania 100 Hz. Tak długi czas rejestracji był niezbędny aby uzyskać wysoką rozdzielczość częstotliwościową do dalszych analiz oraz aby zapewnić możliwie jednorodne warunki wymuszenia w każdym z pięciu eksperymentów częściowych. W każdej rejestracji mierzono 9 sygnałów odpowiedzi i jeden przebieg odniesienia. Do akwizycji danych zastosowano rejestrator cyfrowy TEAC GX-1 w konfiguracji 16-to kanałowej oraz 10 akcelerometrów sejsmicznych PCB 393A03.

5. EKSPLOATACYJNA ANALIZA MODALNA OBIEKTU

Estymację parametrów modalnych poprzedziła analiza jakości danych pomiarowych w oparciu o przebiegi zmierzone w punkcie pomiaru referencyjnego we wszystkich przeprowadzonych eksperymentach częściowych. Na rys. 4 pokazano przebiegi widm amplitudy sygnału referencyjnego.



Rys. 4. Widma amplitudy przyspieszenia drgań (dB m/s^2) w punkcie referencyjnym w 5 eksperymentach częściowych

Przebiegi te pokazują, że we wszystkich eksperymentach częściowych przebieg widma jest podobny co oznacza, że wymuszenie było odpowiednie w każdym z testów.

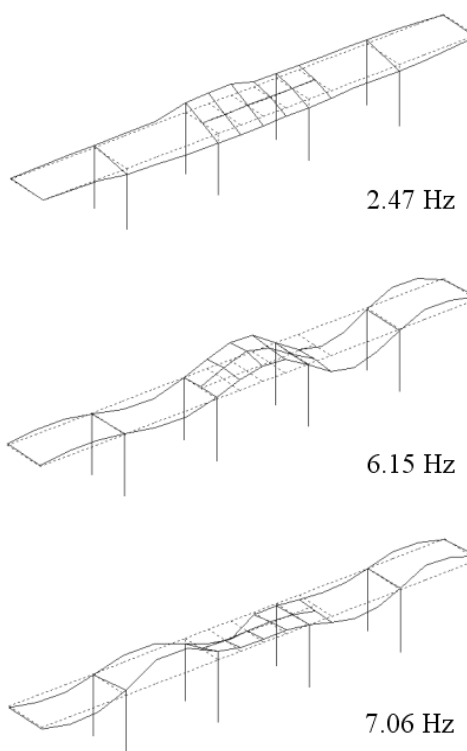
Estymacja parametrów była przeprowadzona

przy pomocy modułu oprogramowania MDOF pakietu oprogramowania do analizy modalnej VIOMA [9]. Zastosowano algorytmy LSCE i BR. Przeprowadzono zbiór procedur estymacji, z których wyników wybrano najbardziej reprezentatywne postacie drgań przy pomocy autorskiej procedury konsolidacji modelu modalnego [10] będącej składnikiem oprogramowania VIOMA.

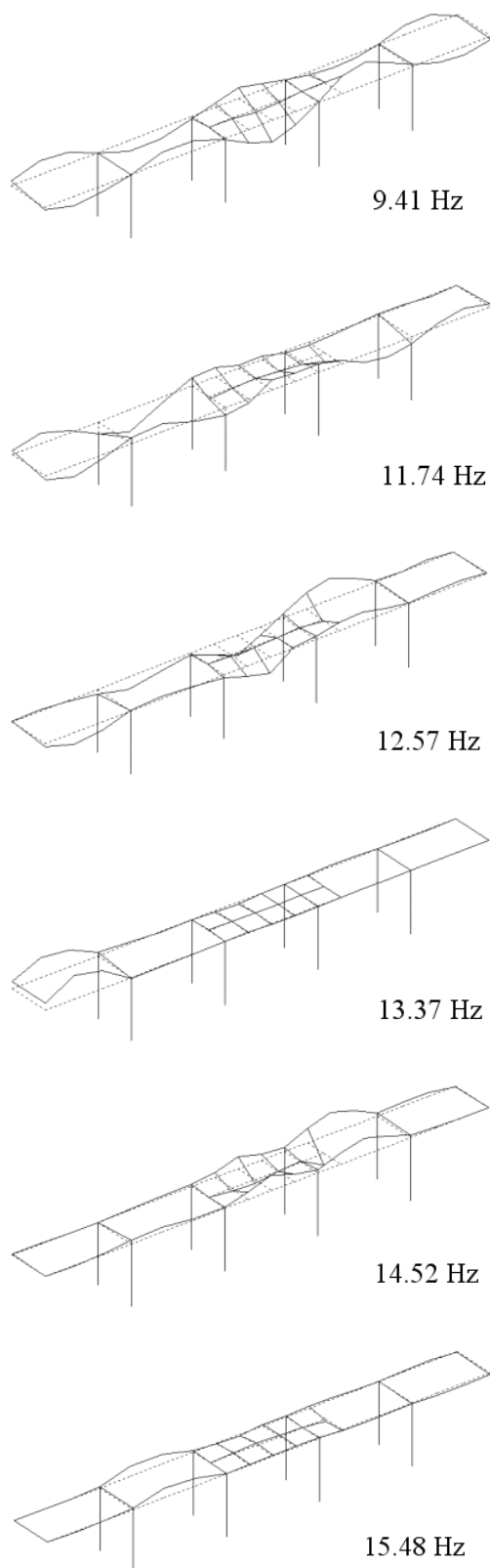
Wyniki estymacji parametrów modelu modalnego zestawiono w tabeli 1. Zidentyfikowane postacie drgań własnych przedstawione są na rys. 5a i rys. 5b.

Tab. 1. Wyniki identyfikacji modelu modalnego wiaduktu

Nr PDW	Częstotliwość [Hz]	Współczynnik tłumienia [%]
1	2.47	4.65
2	6.15	2.64
3	7.06	3.18
4	9.41	2.46
5	11.74	2.96
6	12.57	1.41
7	13.37	1.45
8	14.52	1.91
9	15.48	0.29



Rys. 5a. Zidentyfikowane postacie drgań własnych wiaduktu



Rys. 5b. Zidentyfikowane postacie drgań własnych wiaduktu

Dla badanego wiaduktu przeprowadzono dodatkową analizę. Częstotliwości drgań własnych i współczynniki tłumienia modalnego zostały wyznaczone osobno dla każdego z trzech

zmierzonych dźwigarów. Porównanie otrzymanych wyników zamieszczono w tabeli 2.

Tab. 2. Porównanie parametrów modelu modalnego wyznaczonych osobno dla każdego z badanych dźwigarów

Skrajny lewy dźwigar (A)		4-ty dźwigar od prawej (S)		Skrajny prawy dźwigar (B)	
CzDW [Hz]	Wsp. tł.	CzDW [Hz]	Wsp. tł.	CzDW [Hz]	Wsp. tł.
2.49	3.29	2.49	3.15	2.50	3.53
6.09	2.46	6.18	2.64	6.14	2.46
7.08	3.15	7.11	3.83	7.01	2.96
9.29	2.34	9.56	2.66	9.34	2.85
14.51	1.56	14.57	2.65	14.53	1.58

Analizując wartości częstotliwości drgań własnych poszczególnych dźwigarów zawarte w tabeli 2 można zauważyć, że zewnętrzne dźwigary są mniej sztywne od wewnętrznego. Może to być symptomem znacznie większego zużycia tych elementów. Aby dokładniej określić przyczynę mniejszej sztywności skrajnych dźwigarów należy zastosować inną metodę nieniszczących badań uszkodzenia obiektu np. metodę dostrajania modelu elementów skończonych. Podobne wnioski, dotyczące mniejszej sztywności zewnętrznej części wiaduktu, można wyciągnąć obserwując kształt postaci drgań własnych wiaduktu. Widoczna jest wyższa względna amplituda przemieszczeń tych części obiektu (na postaciach przedstawiających zginanie).

W tabeli 3 zawarto współczynniki MAC policzone pomiędzy odpowiednimi postaciami drgań własnych badanych dźwigarów w celu ilościowej oceny różnicy względnych amplitud dźwigarów w postaciach zginania.

Tab. 3. Zestawienie współczynników MAC dla kolejnych dźwigarów.

Nr PDW	MAC pomiędzy A^1 i B^2	MAC pomiędzy A^1 i S^3	MAC pomiędzy B^2 i S^3
2	0.8015	0.4916	0.7218
3	0.8484	0.9376	0.9344
5	0.5242	0.0649	0.4090
6	0.7210	0.4841	0.4501
8	0.9756	0.2973	0.2873
1) skrajny lewy dźwigar 2) skrajny prawy dźwigar 3) 4-ty dźwigar od prawej			

Współczynniki MAC zestawione w tabeli 3 potwierdzają różnicę w kształcie postaci wyestymowanych dla kolejnych dźwigarów. Wartości MAC'ów pomiędzy skrajnymi dźwigarami są wysokie natomiast pomiędzy skrajnymi a środkowym niskie. W połączeniu z obserwacją kształtu postaci można na tej podstawie wnioskować

o utracie sztywności skrajnych dźwigarów, na które działają większe obciążenia w trakcie eksploatacji.

6. PODSUMOWANIE

Wyniki zaprezentowane w artykule ukazują przydatność eksploatacyjnej analizy modalnej do oceny stanu technicznego obiektów inżynierii lądowej takich jak mosty, czy wiadukty. Metoda została użyta do takiej oceny intensywnie eksploatowanego wiaduktu drogowego, aby wspomóc zarząd dróg i mostów w podjęciu decyzji o terminie i zakresie prac remontowych.

Metoda może też być zastosowana do ciągłego monitoringu stanu omawianych obiektów. Jeżeli znany będzie trend zmiany parametrów modalnych, możliwe będzie także określenie czasu jego bezpiecznej eksploatacji. Częstotliwości drgań własnych i współczynniki tłumienia modalnego mogą być wykorzystywane jako wskaźniki wystąpienia uszkodzenia.

LITERATURA

- [1].Betti, R., Testa, R. B.: *Vibration and Damage Control for Long-Span Bridges*, Journal of Smart Materials and Structures, Vol. 4, s. 91-100, 1995.
- [2].Kawiecki, G.: *Modal Damping Measurements for Damage Detection*, European COST F3 Conference on System Identification and Structural Health Monitoring, Madrid, Spain, 651-658, 2000.
- [3].Agneni A., Crema, L., and Mastroddi F.: *Damage Detection from Truncated Frequency Response Functions*, European COST F3 Conference on System Identification and Structural Health Monitoring, Madrid, Spain, 137-146, 2000.
- [4].Ho, Y. K., Ewins, D. J., *Numerical Evaluation of the Damage Index*, Structural Health Monitoring 2000, Stanford University, Palo Alto, CA, 995-1011, 1999.
- [5].Wang, M. L., Xu, F. L., Lloyd, G. M.: *A Systematic Numerical Analysis of the Damage Index Method Used for Bridge Diagnostics*, Smart Structures and Materials 2000: Smart Systems for Bridges, Structures, and Highways, Proceedings of SPIE, Vol. 3988, Newport Beach, CA, pp. 154-164, 2000.
- [6].Carrasco, C., Osegueda, R., Ferregut, C., Grygier, M.: *Localization and Quantification of Damage in a Space Truss Model Using Modal Strain Energy*, Smart Systems for Bridges, Structures, and Highways, Proceedings of SPIE, Vol. 3043, pp. 181- 192, 1997.
- [7].Peeters B., Maeck J., De Roeck. G.: *Dynamic monitoring of the Z24-bridge: separating temperature effects from damage*, In Proceedings of the European COST F3 Conference on System Identification and Structural Health Monitoring, s. 377-386, Madrid, Spain, Lipiec 2000.
- [8].Uhl T.: *Komputerowe wspomaganie identyfikacji modeli konstrukcji*, WNT, Warszawa, 1997.
- [9].Uhl T., Lisowski W.: *In-operation modal analysis*, Wydawnictwo KRI DM AGH, Kraków, 2001.
- [10].Lisowski W.: *Przykład algorytmu konsolidacji modelu w autonomicznej analizie modalnej*, Zagadnienia analizy modalnej konstrukcji mechanicznych praca zbiorowa pod redakcją Tadeusza Uhl, Wydawnictwo KRI DM AGH, Kraków, 2003.

Dr inż. **Krzysztof**

MENDROK jest adiunktem w Katedrze Robotyki i Dynamiki Maszyn, Akademii Górniczo – Hutniczej w Krakowie. Jego zainteresowania skupiają się na zagadnieniach dynamiki strukturalnej, a zwłaszcza na algorytmach analizy modalnej i analizy dróg rozchodzenia się energii drgań w konstrukcjach. Zajmuje się wykorzystaniem tych metod do wykrywania uszkodzeń i identyfikacji sił. Jest autorem prac podejmujących tę tematykę.



WPŁYW WYMIANY CIEPŁA W WĘZŁACH ŁOŻYSKOWYCH NA SYMPTOMY DIAGNOSTYCZNE MASZYN WIRNIKOWYCH

Jan KICIŃSKI

Instytut Maszyn Przepływowych PAN w Gdańsku
Wydział Nauk Technicznych, Uniwersytetu Warmińsko-Mazurskiego w Olsztynie
e-mail: kic@imp.gda.pl

Streszczenie

Praca przedstawia wpływ wymiany ciepła w filmie smarnym łożysk ślizgowych i w efekcie wpływ deformacji termosprężystych panwi i jej utwierdzeń zewnętrznych na pracę dwupodporowego wirnika ułożyskowanego ślizgowo. Analizę teoretyczną przeprowadzono w oparciu o tzw. 3-D elastodiatermiczny model węzła łożyskowego. Przeprowadzona została analiza wpływu dwóch rodzajów utwierdzeń zewnętrznych i zacisku wstępnego panwi na zachowanie się układu w stabilnym jak i niestabilnym zakresie pracy. Prześledzony został rozwój wirów i bicia olejowego w łożyskach i opisane zostały wyróżniki diagnostyczne takiego stanu.

Słowa kluczowe: łożyska ślizgowe, wiry i bicie olejowe, dynamika wirników, symptomy diagnostyczne.

THE INFLUENCE OF HEAT TRANSFER IN BEARINGS KNOTS ON DIAGNOSTICS SYMPTOMS OF ROTOR MACHINES

Summary

The paper presents the influence of heat transfer in oil film of journal bearings and thereby the influence of thermoelastic deformation of bearing bush and its external fixings on two supported rotor characteristics. Analysis has been performed on the basis of 3-D elastodiathermic model of journal bearing. Carried out has been the analysis of influence of two kinds of external fixations and initial clamping of bearing bush on properties of rotor set in stable and unstable operation regime. Traced has been the propagation of oil whirl and oil whip in journal bearings and described has been the diagnostic determinant of such state.

Keywords: slide bearings, oil whirl, oil whip, rotor dynamic, diagnostic symptoms.

1. MODEL TEORETYCZNY

W niniejszym artykule przedstawimy przykład aplikacji zaawansowanych modeli i programów komputerowych do oceny stanu dynamicznego maszyn wirnikowych. Wykorzystamy tu przede wszystkim nowoczesne metody wibrodiagnostyki w odniesieniu do nowej i szybko rozwijającej się dyscypliny wiedzy, a mianowicie diagnostyki według modelu.

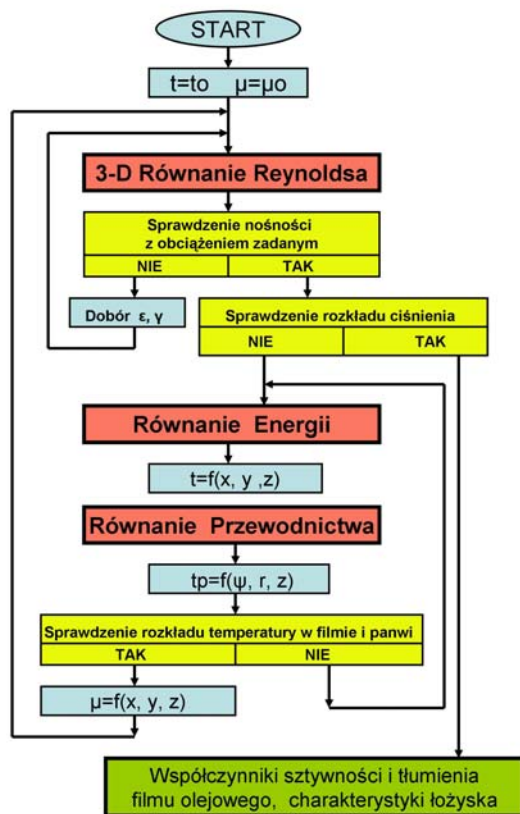
W klasycznej maszynie wirnikowej wyróżnić możemy trzy zasadnicze podukłady:

- linia wirników z dyskami, sprzęgłami i imperfekcjami typu pęknięcia wału lub rozosiowania wałów;
- hydrodynamiczne poprzeczne łożyska ślizgowe i uszczelnienia labiryntowe;
- fundament z podporami i utwierdzeniami zewnętrznymi łożysk.

Szczególnie trudne do zamodelowania teoretycznego są łożyska ślizgowe i uszczelnienia labiryntowe. W IMP PAN w Gdańsku opracowany został 3-D elastodiatermiczny model wymiany ciepła

w łożysku (program DIADEF, a na jego bazie program KINWIR-F, do analizy zagadnień kinetostatyki wirników) obejmujący skojarzone równania Reynoldsa, energii, przewodnictwa a następnie deformacji panwi – rys. 1. Do analizy zagadnień dynamiki układu wirnik-łożyska-fundament w zakresie nieliniowym, a więc w zakresie umożliwiającym analizę stanu obiektu po przekroczeniu granicy stabilności, opracowany został blok programów serii NLDW. Programy serii KINWIR jak i NLDW stanowią element środowiska MESWIR szczegółowo opisanego w monografii [1]. Stąd też nie będą one prezentowane bliżej w niniejszym artykule.

Opracowane narzędzia badawcze umożliwiają m.in. modelowanie utwierdzeń zewnętrznych panwi i zacisków wstępnych w stojakach. Wyłania się tym samym interesująca możliwość badania wpływu np.: propagacji wirów i bicia olejowego przy zmienionym kształcie szczeliny smarnej na skutek wymiany ciepła, deformacji panwi i wpływu utwierdzeń zewnętrznych. Autorowi niniejszego artykułu nie są znane doniesienia literaturowe na ten temat.



Rys. 1. Przyjęty model wymiany ciepła w łożysku

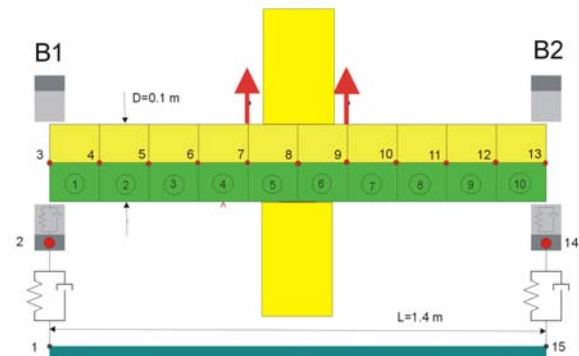
2. WYNIKI ANALIZY

Przyjmijmy do dalszych rozważań dwa przypadki utwardzenia zewnętrznego panwi dla wirnika z rys. 2, a mianowicie:

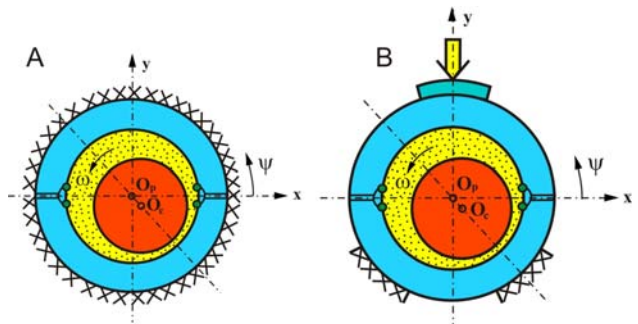
- utwardzenie nieskończenie sztywne na całej powierzchni zewnętrznej panwi (określimy to jako typ A). Oznacza to, że możliwe są jedynie deformacje panwi i szczeliny w kierunku wewnętrznym;
- utwardzenie sztywne częściowe dolnej panwi z jednoczesnym zaciskiem panwi górnej (co określimy jako typ B). Teraz panew może odkształcać się zarówno kierunku zewnętrznym jak i wewnętrznym, to znaczy w kierunku do środka panwi.

Schematy obu przypadków utwardzeń zewnętrznych panwi przedstawia rys. 3.

Wyniki obliczeń rozkładów temperatury i deformacji panwi za pomocą wzajemnie skojarzonych programów KINWIR-F i NLDW-LEW (zagadnienia kinetostatyki i dynamiki) przedstawiają rys. 4 do 6.

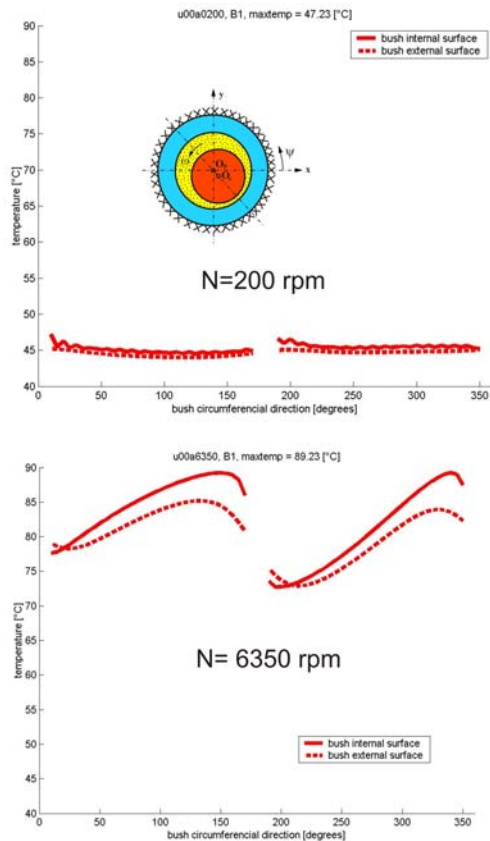


Rys. 2. Przyjęty do analizy diatermiczny model cieplny łożyska (lewa strona: algorytm obliczeń dla programu DIADEF) i przyjęty model obiektu: dwupodporowy wirnik łożyskowy ślizgowo. Obliczenia prowadzone były w oparciu o programy serii KINWIR i NLDW stanowiące elementy środowiska MESWIR [1]

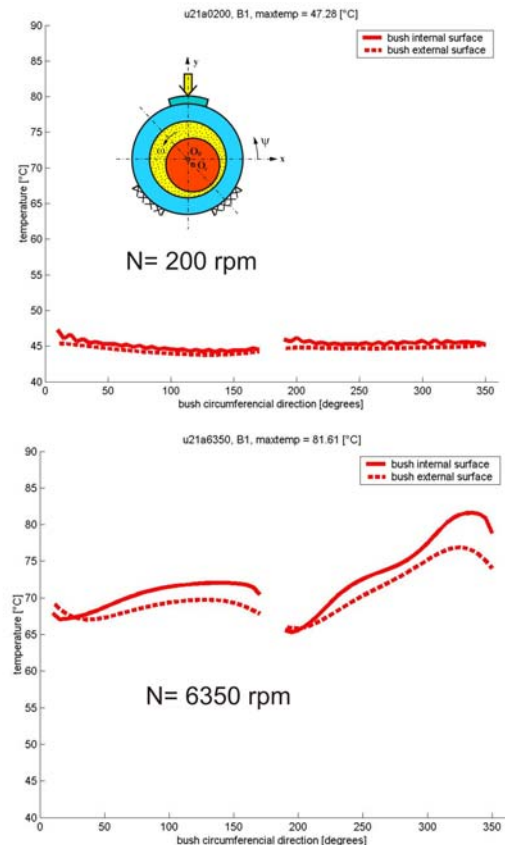


Rys. 3. Analizowane przypadki utwardzeń zewnętrznych panwi. A – utwardzenie sztywne na całej powierzchni zewnętrznej panwi, B – utwardzenie sztywne częściowe panwi dolnej z jednoczesnym zaciskiem panwi górnej

Jak widać z rys. 6 duży wzrost temperatury panwi przy wysokich obrotach wirnika (rys. 4 i rys. 5) spowodował znaczne zmiany kształtu szczeliny smarnej łożyska w stosunku do jej kształtu konstrukcyjnego (w tym przypadku luzu kołowo-cylindrycznego). Zmiany te są szczególnie mocno widoczne w przypadku utwardzenia B (rys. 3). Zacisk panwi górnej spowodował jej niesymetryczne odkształcenie w kierunku zewnętrznym, a częściowe utwardzenie panwi dolnej spowodowały istnienie charakterystycznego „dółka” w części centralnej panwi dolnej – rys. 6. Utwardzenia typu A (rys. 3) spowodowało prawie równomierne zmniejszenie luzu promieniowego o ok. 25 %. Fakt ten zmienił oczywiście przebieg trajektorii wirnika, mimo iż na układ w obu przypadkach działają dokładnie te same siły zewnętrzne (siły niewyważenia resztkowego).



Rys. 4. Rozkłady temperatury na powierzchni zewnętrznej (linia przerywana) i wewnętrznej panwi (linia ciągła) obliczone dla przypadku utwardzenia A z rys. 4 i różnych prędkości obrotowych wirnika z rys. 3



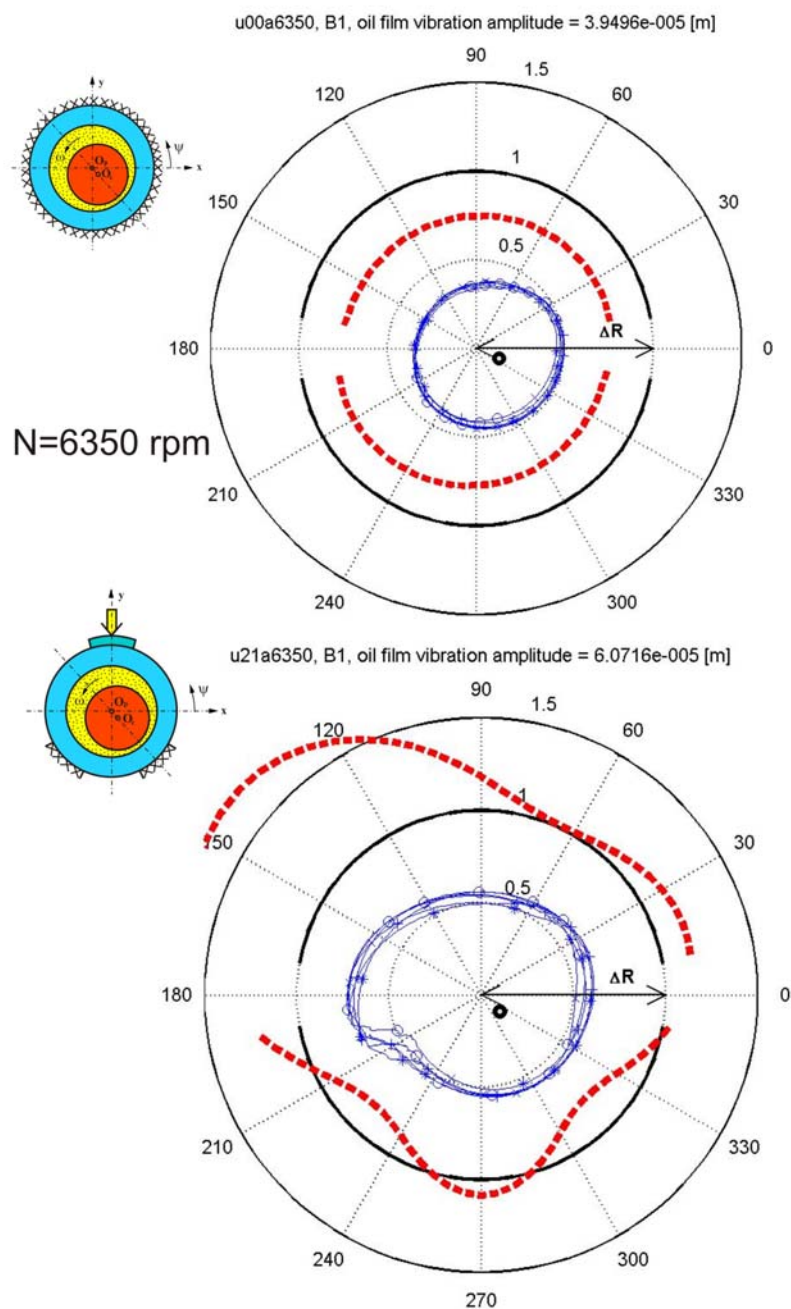
Rys. 5. Rozkłady temperatury na powierzchni zewnętrznej (linia przerywana) i wewnętrznej panwi (linia ciągła) obliczone dla przypadku utwardzenia B z rys. 3 i różnych prędkości obrotowych wirnika z rys. 2

Zobaczmy teraz jak będą rozwijały się wiry i bicia olejowe jeśli zmienimy rodzaj utwardzeń zewnętrznych panwi. Odpowiemy sobie tym samym na pytanie, na ile konieczne jest stosowanie zaawansowanych narzędzi badawczych w ocenie niestabilności hydrodynamicznej układu. Pozostaniemy przy tym samym układzie, a mianowicie wirniku symetrycznym z jedną tarczą posadowionym na dwóch identycznych łożyskach ślizgowych – rys. 2. Wyniki stosownych obliczeń przedstawiają rys. 7 do rys. 12.

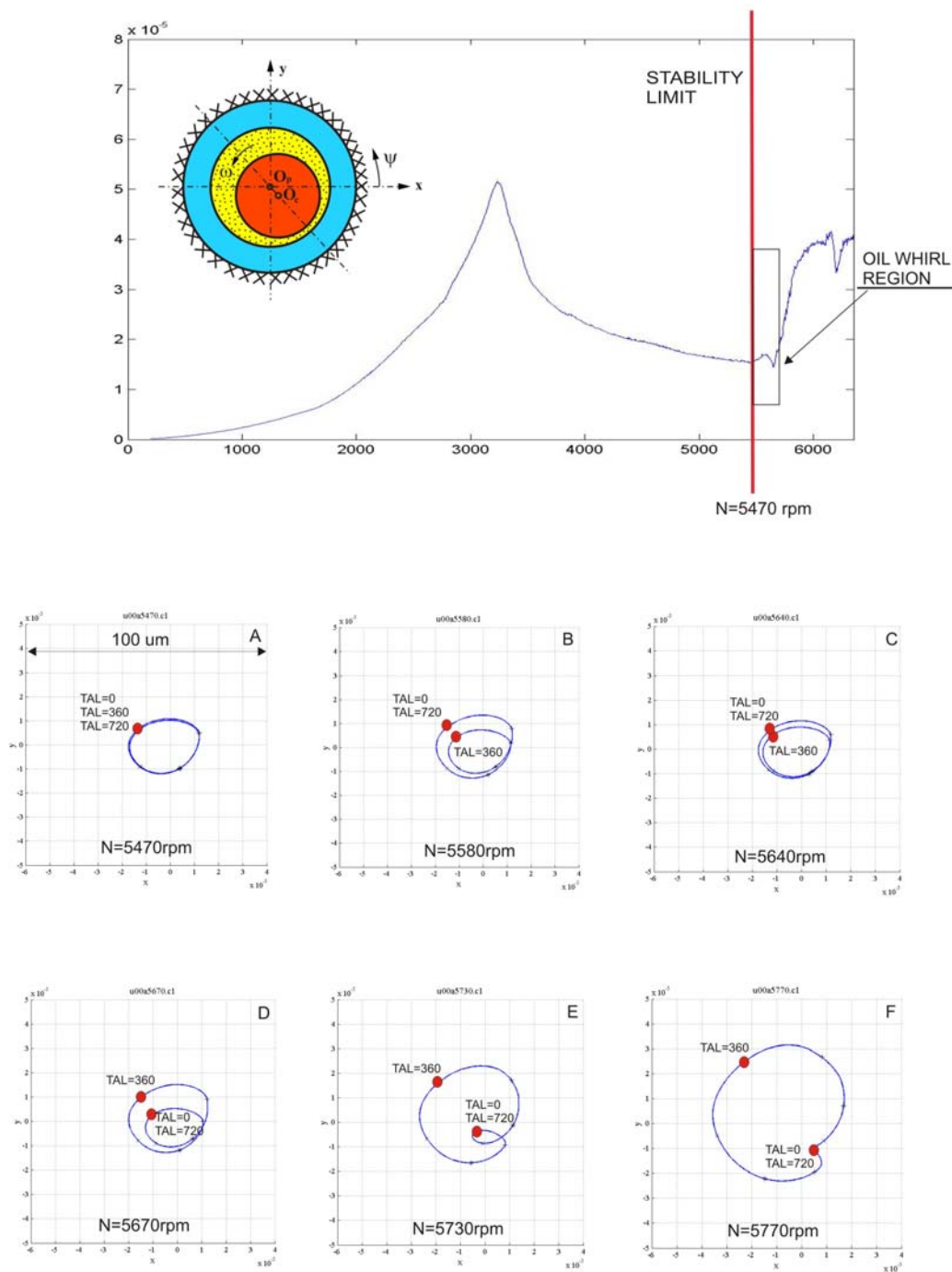
Porównując rys. 7 z rys. 9 a następnie rys. 8 z rys. 10 stwierdzamy, iż utwardzenie typu B okazało się mniej korzystne z punktu widzenia dynamiki układu. Granica stabilności o prawie 400 rpm wypada wcześniej, amplitudy przemieszczeń wirów i bicia są zdecydowanie większe, a ponadto wiry szybciej przechodzą do niebezpiecznej już fazy

bicia olejowego. Są to wyniki dość zaskakujące. Jak wynika bowiem z badań prowadzonych w [2], ten typ utwardzenia zwiększał w efekcie tzw. rezerwę stabilności układu w stabilnym zakresie jego pracy. Liniowy opis dynamiki układu stosowany w [2] nie pozwalał jednakże na prześledzenie warunków pracy po przekroczeniu granicy stabilności. Dopiero opis nieliniowy zastosowany w niniejszej analizie umożliwił wyciągnięcie pełnych i właściwych wniosków.

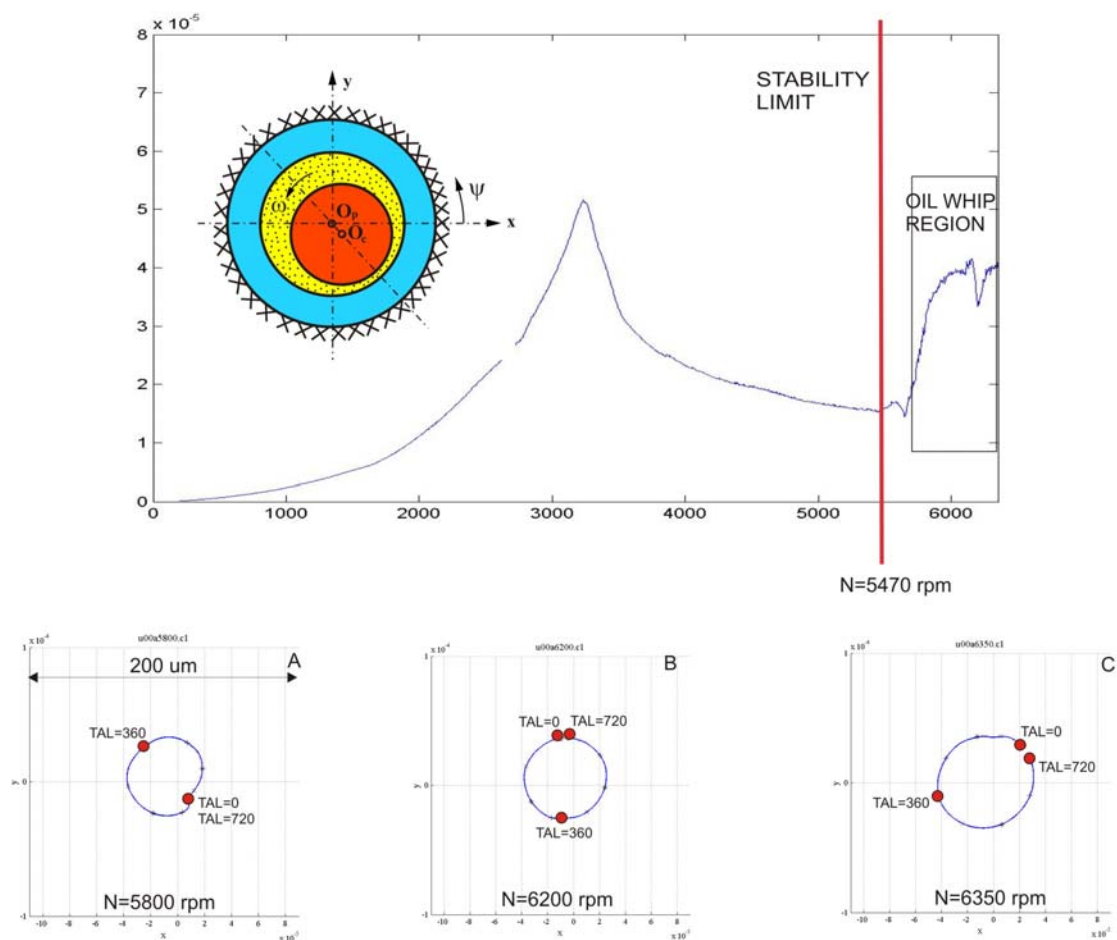
Rys. 11 i rys. 12 pozwalają ujawnić, że dla tych samych warunków pracy i położenia zewnętrznej siły wymuszającej drgania układu „wychwycone w locie” przestrzenne rozkłady ciśnienia obliczone dla przyjętych dwóch typów utwardzenia panwi, wypadają zupełnie inaczej, co uzasadnia wcześniejsze różnice w przebiegu trajektorii.



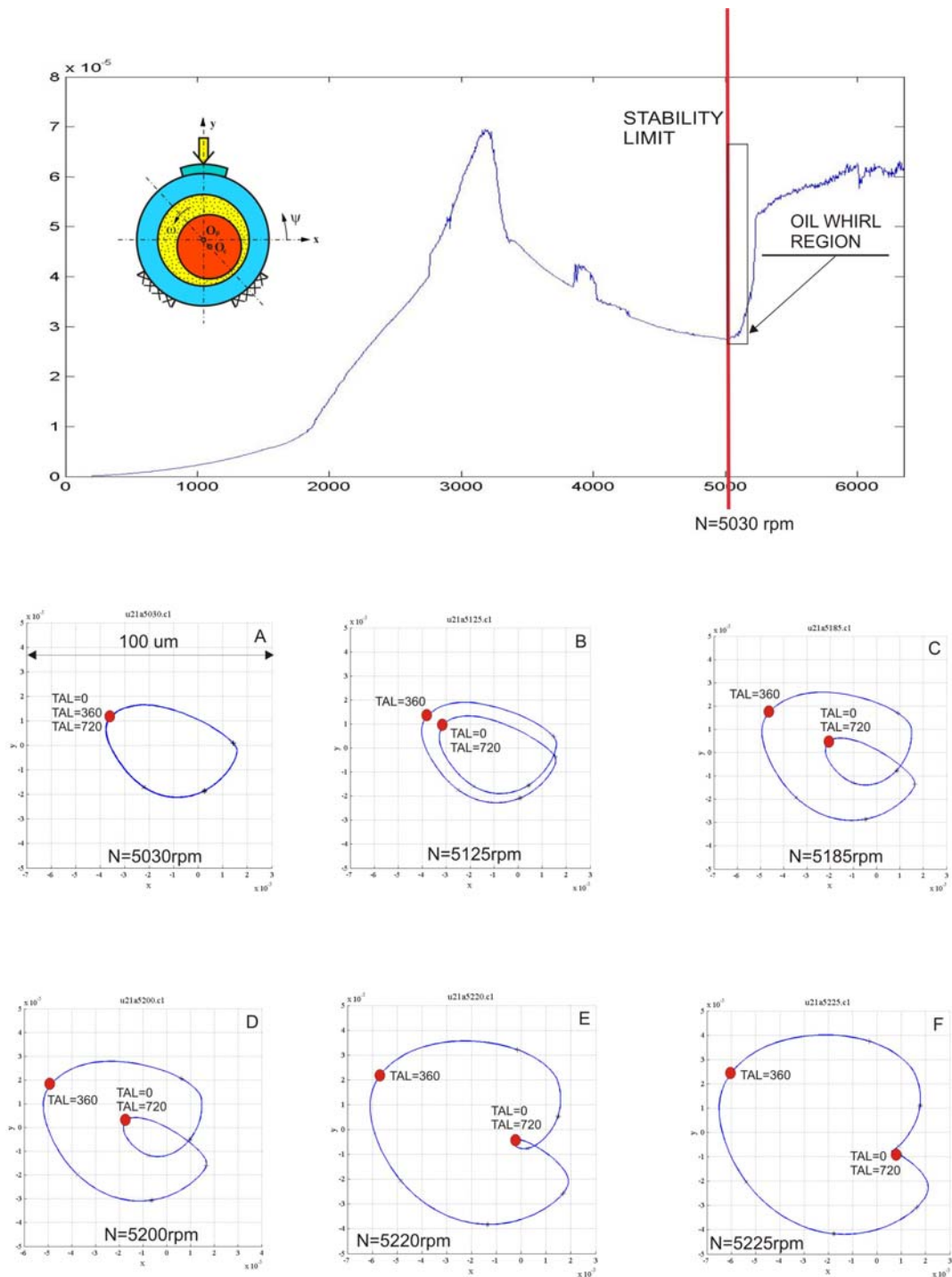
Rys. 6. Kształt szczeliny smarnej z uwzględnieniem deformacji termosprężystych panwi (linia przerywana) na tle kształtu konstrukcyjnego (linia ciągła) i trajektorii przemieszczeń czopa przy prędkości wirnika $N = 6350$ rpm (obszar silnej niestabilności hydrodynamicznej). Obliczenia przeprowadzone zostały dla dwóch przypadków utwardzeń zewnętrznych panwi według rys. 3



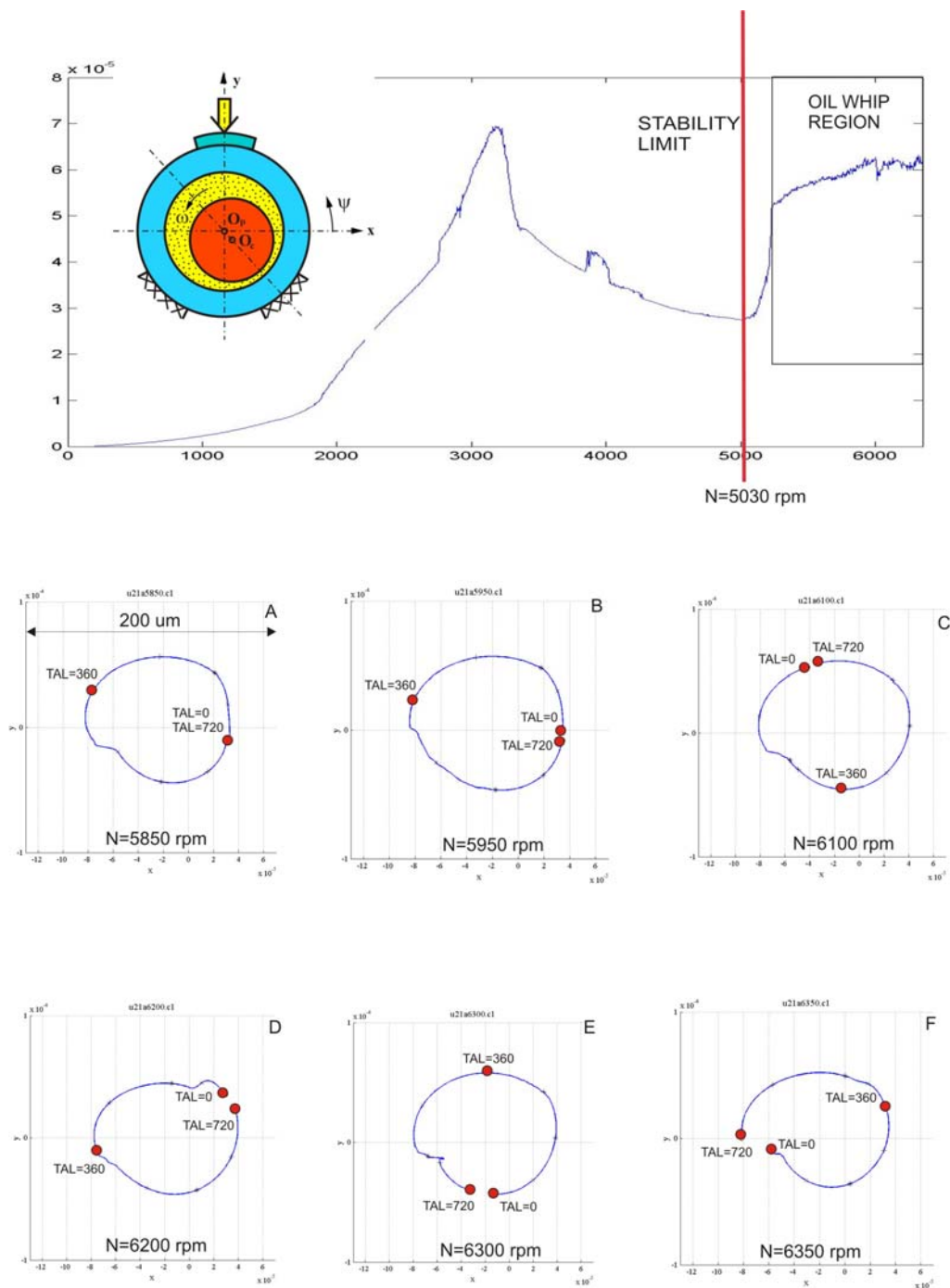
Rys. 7. Rozwój wirów olejowych obliczony dla przypadku utwardzeń zewnętrznych panwi typu A. Określenia TAL=0, 360, 720 oznaczają położenia znacznika na trajektorii przy położeniu wektora sił wymuszających poziomo w prawo (początek przyjętego układu współrzędnych). Część górna wykresu – charakterystyki amplitudowo-częstotliwościowe układu



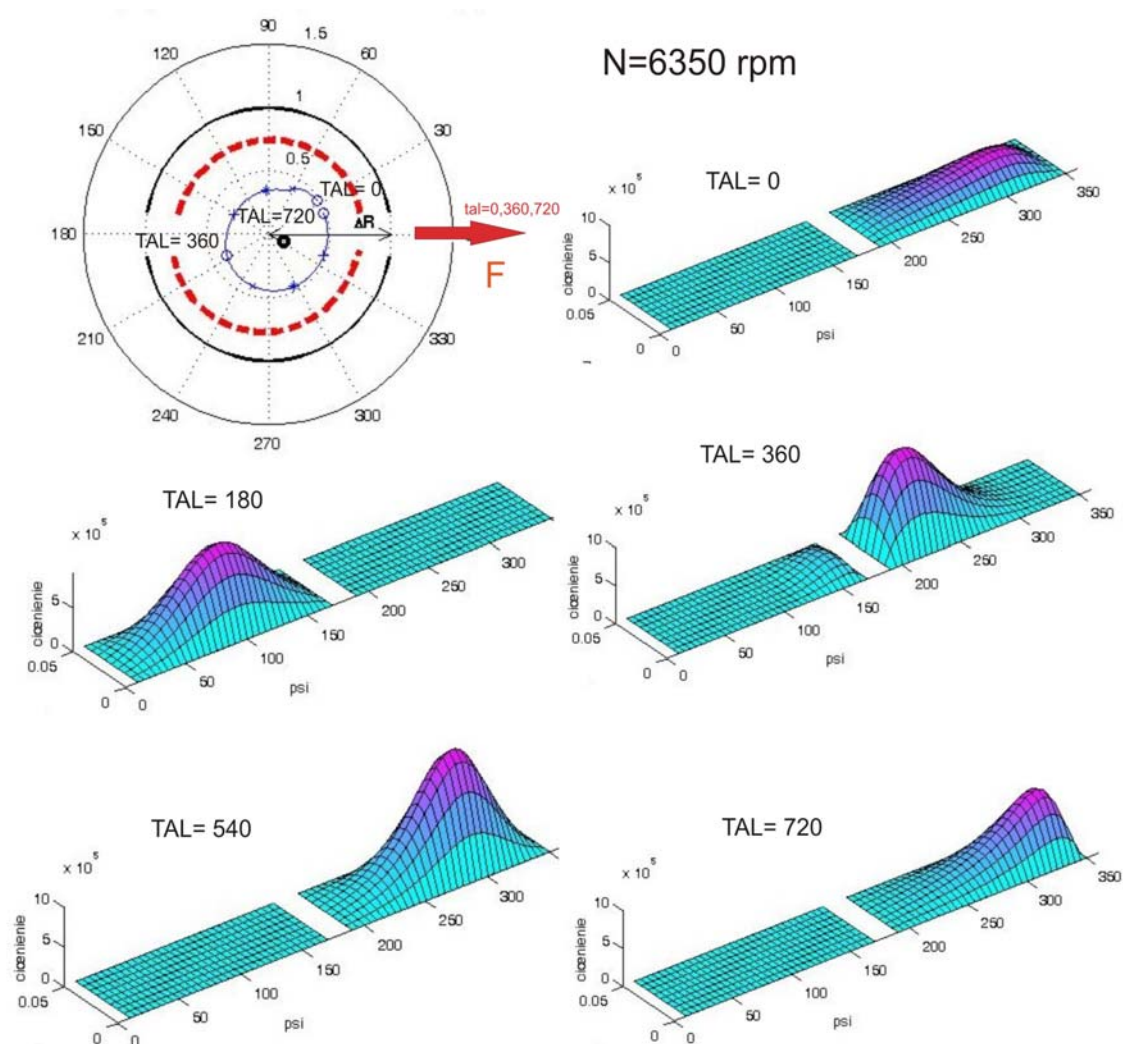
Rys. 8. Rozwój bicia olejowego obliczony dla przypadku utwardzeń zewnętrznych panwi typu A. Określenia $TAL=0, 360, 720$ oznaczają położenia znacznika na trajektorii przy położeniu wektora sił wymuszających poziomo w prawo (początek przyjętego układu współrzędnych). Część górna wykresu – charakterystyki amplitudowo-częstotliwościowe układu



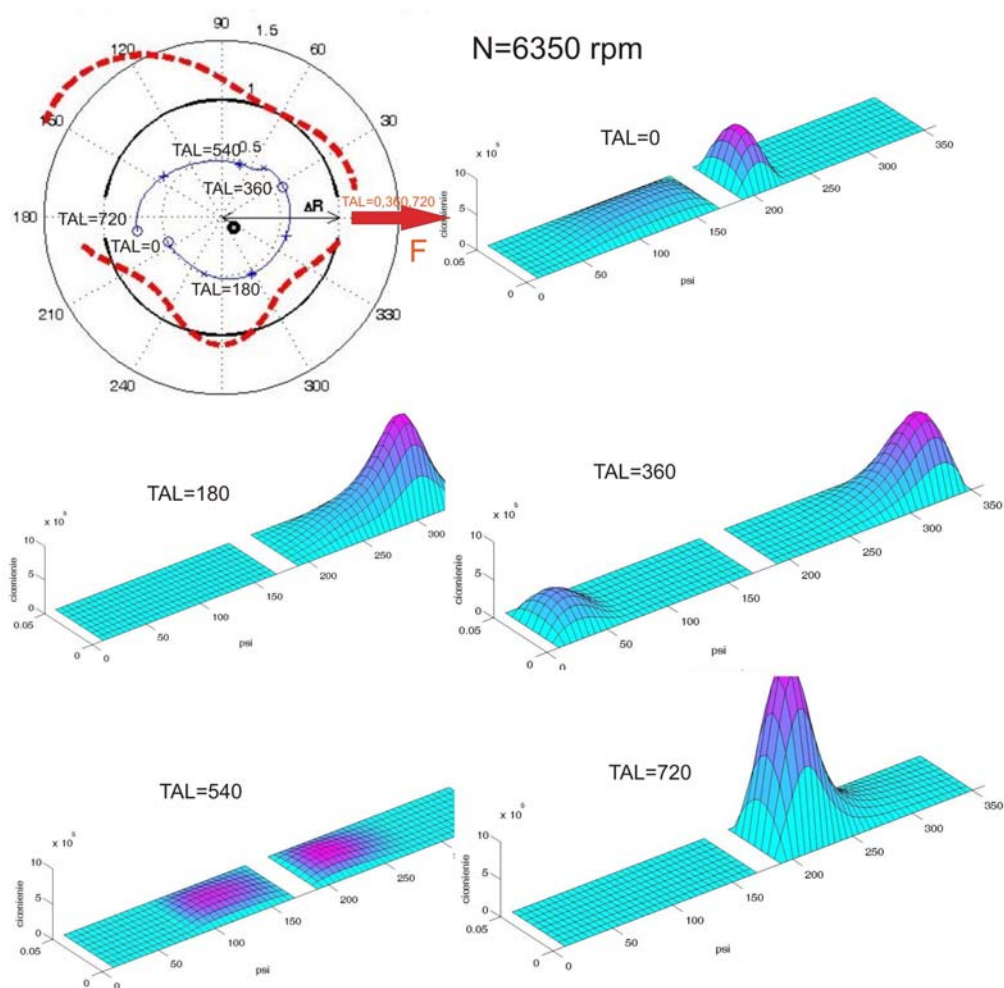
Rys. 9. Rozwój wirów olejowych obliczony dla przypadku utwardzeń zewnętrznych panwi typu B. Określenia TAL=0, 360, 720 oznaczają położenia znacznika na trajektorii przy położeniu wektora sił wymuszających poziomo w prawo (początek przyjętego układu współrzędnych). Część górna wykresu – charakterystyki amplitudowo-częstotliwościowe układu



Rys. 10. Rozwój bicia olejowego obliczony dla przypadku utwardzeń zewnętrznych panwi typu B. Określenia $TAL=0, 360, 720$ oznaczają położenia znacznika na trajektorii przy położeniu wektora sił wymuszających poziomo w prawo (początek przyjętego układu współrzędnych). Część górna wykresu – charakterystyki amplitudowo-częstotliwościowe układu



Rys. 11. Przestrzenny rozkład ciśnienia hydrodynamicznego w szczelinie smarnej łożysk „wychwycony w locie” dla poszczególnych charakterystycznych położenia znacznika TAL na trajektorii. Określenia TAL=0, 360, 720 oznaczają położenia znacznika na trajektorii przy położeniu wektora sił wymuszających poziomo w prawo (początek przyjętego układu współrzędnych). Przypadek A utwardzenia zewnętrznego panwi. Strefa bicia olejowego (N=6350 rpm)



Rys. 12. Przestrzenny rozkład ciśnienia hydrodynamicznego w szczelinie smarnej łożysk „wychwycony w locie” dla poszczególnych charakterystycznych położen znacznika TAL na trajektorii. Określenia TAL=0, 360, 720 oznaczają położenia znacznika na trajektorii przy położeniu wektora sił wymuszających poziomo w prawo (początek przyjętego układu współrzędnych). Przypadek B utwierdzenia zewnętrznego panwi. Strefa bicia olejowego ($N=6350$ rpm)

3. UWAGI KOŃCOWE

Podsumowując nasze rozważania musimy stwierdzić, że odpowiedź na pytanie, czy analiza teoretyczna z zastosowaniem bardzo zaawansowanych narzędzi jest zawsze konieczna, również i w tym przypadku wypada niejednoznacznie. W pewnych sytuacjach (u nas przypadek utwierdzenia A) zastosowanie izotermicznego modelu cieplnego łożysk i sztywnej panwi wydaje się być wystarczające do analizy stanu dynamicznego układu w całym zakresie prędkości obrotowych wirnika. W innych sytuacjach (np.: przypadek utwierdzenia B) zastosowanie bardziej złożonych modeli jest konieczne dla właściwej oceny stanu maszyny. Niestety często trudno jest *a priori* ocenić jaki rodzaj utwierdzenia czy warunków pracy wymagał będzie bardziej zaawansowanych narzędzi.

W przypadku tego rodzaju wątpliwości rozsądnym wydaje się zastosowanie w pierwszej kolejności narzędzi bardziej złożonych, a dopiero później dokonywać ewentualnych uproszczeń.

LITERATURA

- [1] J. Kiciński, *Dynamika wirników i łożysk ślizgowych*, Maszyny Przepływowe t. 28, Wydawnictwo IMP, 2005.
- [2] J. Kiciński, R. Haller, *Computer Optimization of External Fixation and Initial Camping of a Bushing with Respect to Static and Dynamic Properties of Journal Bearings*, Transaction of the ASME, Vol. 116, October 1994.

WYKORZYSTANIE INFORMACJI DIAGNOSTYCZNEJ W ZORIENTOWANEJ NA BEZPIECZEŃSTWO EKSPLOATACJI MASZYN

Stanisław RADKOWSKI

Instytut Podstaw Budowy Maszyn, Politechnika Warszawska
ul. Narbutta 84, 02-524 Warszawa, (022) 660 86 22, ras@simr.pw.edu.pl

Streszczenie

W pracy zaprezentowano możliwości wykorzystania metod utrzymania ruchu ze względu na zapewnienie wymaganego poziomu ryzyka w nowej strategii eksploatacji maszyn. We wprowadzeniu omówiono podstawy teoretyczne, a następnie przedstawiono przykłady detekcji uszkodzeń i ich wpływu na ocenę poziomu ryzyka technicznego. W końcowej części artykułu przedstawiono możliwość wykorzystania metod uczenia statystycznego w obliczaniu i dyskryminacji bezpiecznych i niebezpiecznych wartości parametru.

Słowa kluczowe: niezawodność, diagnostyka, ryzyko techniczne, analiza bispektralna.

PRESUMING OF DIAGNOSTIC INFORMATION IN SAFETY ORIENTED USE OF MACHINERY

Summary

In the paper is explained the use of Risk Based Maintenance methods in new strategy of machinery using. At the beginning is presented a theoretical consideration on some application aspects. Next is presented an example of method failure detection and their impact on evaluation of technical risk. At the end of the paper are discussed the possibility of using the Statistical Learning methods in discrimination of the safety and unsafety values of parameter.

Keywords: reliability, diagnostics, technical risk, bispectral analysis.

1. WPROWADZENIE

Wczesne wykrycie uszkodzenia pozwala uniknąć degradacji wytworu w szczególności awarii systemów technicznych oraz związanych z tym konsekwencji w postaci strat materialnych i pozamaterialnych i stanowić podstawę do optymalnej decyzji w czasie umożliwiającym podjęcie działań korygujących bądź naprawczych. Bazą dla tego typu działań jest analiza ryzyka technicznego, w której do oszacowania ryzyka potrzebujemy zarówno informacji o prawdopodobieństwie wystąpienia zdarzenia niepożądanego jak i wiarygodnej prognozy hipotetycznych konsekwencji tego zdarzenia. W tym ujęciu kształtowanie wymaganego poziomu ryzyka technicznego jest rozumiane jako kluczowy czynnik sukcesów biznesowych i nieodłączny element działań technicznych. Tak rozumiane ryzyko techniczne będzie mieć progresywny i mierzalny wpływ na zmniejszenie liczby wypadków w pracy, chorób zawodowych, incydentów środowiskowych i związanych z wypadkami strat produkcyjnych.

Jednym z problemów, który jest przyczyną szerokiego zainteresowania analizą ryzyka wśród inżynierów i projektantów jest możliwość oceny niepewności procesu projektowania i podejmowania decyzji w warunkach niepewności

odnośnie określonego rozwiązania technicznego. Coraz częściej ocena niepewności jest uwzględniana w przypadkach projektowania złożonych systemów technicznych i antropotechnicznych, podejmowania problemów wymagających zastosowania nowych technologii, niekonwencjonalnych rozwiązań systemowych i nowych materiałów.

Jednym z podstawowych elementów, przyjmowanych w efekcie tak prowadzonych analiz rozwiązań, jest wykorzystanie układów diagnostyki monitoringu stanu technicznego systemów lub odpowiednio układów nadzoru procesu w celu wykrycia i identyfikacji fazy rozwoju uszkodzeń oraz uniemożliwienia wystąpienia niebezpiecznych zaburzeń funkcjonowania krytycznych elementów i zespołów systemu.

2. RYZYKO TECHNICZNE JAKO KRYTERIUM OPTYMALIZACJI INSPEKCJI DIAGNOSTYCZNYCH

Jednym z ważnych czynników kształtujących w strategii eksploatacji jest procedura wydłużania okresu użytkowania maszyn i urządzeń. Oznacza to konieczność opracowania metod diagnozowania i prognozowania dopuszczających wydłużenie resursów systemów technicznych ponad normatywny okres. Oprócz ekonomicznego

rachunku odnośnie użytkowania do wystąpienia zdarzenia niepożądanego, ostatnio coraz częściej jako wielkość kryterialną przyjmuje się bezpieczeństwo techniczne, którego miarą jest poziom ryzyka technicznego. Stąd analiza ryzyka stała się istotnym czynnikiem w podejmowaniu decyzji mających na celu kształtowanie bezpieczeństwa i minimalizację kosztów.

Zgodnie z definicją zawartą w normie [1] analiza ryzyka jest procesem, podczas którego identyfikuje się zarówno prawdopodobieństwo wystąpienia zdarzenia niepożądanego jak i zakres związanych z tym szkód (konsekwencji) spowodowanych zarówno przez działanie, jak i urządzenie lub system.

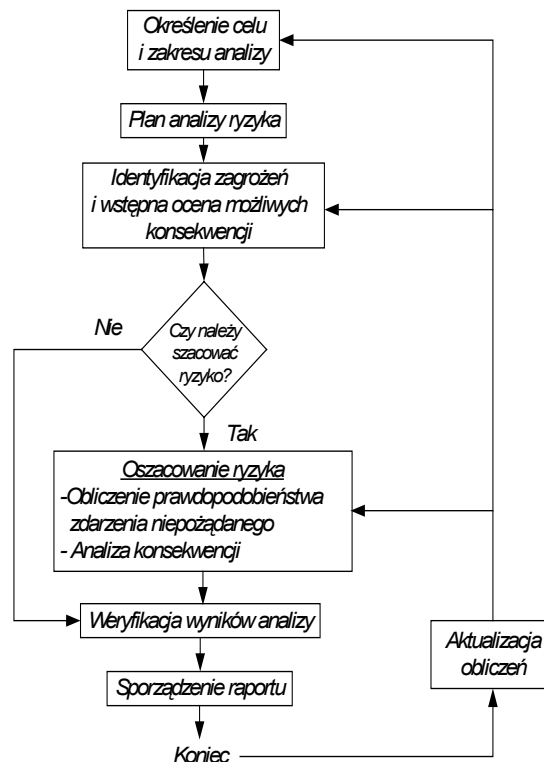
Zgodnie z zaleceniami zawartymi w normach w toku postępowania należy wyszczególnić następujące kroki:

- określenie zakresu analizy;
- identyfikację zagrożeń;
- analizę zagrożeń;
- oszacowanie ryzyka.

Wynikający z zaleceń normowych algorytm analizy ryzyka przedstawiono na rysunku 1. Pełna procedura powinna zawierać weryfikację wyników oraz sprawozdanie, które może kończyć procedurę lub być podstawą do rozpoczęcia kolejnej pętli aktualizującej obliczenia. Na wstępie zakłada się, że zespół przeprowadzający analizę ryzyka, cechuje dogłębna znajomość stosowanych metod i analizowanego systemu. Pozwala to poprawnie określać zakres analizy a szczególnie formułować cele i definiować kryteria przy równoczesnym zdefiniowaniu systemu dla określonych założeń i przyjętych ograniczeń, oraz wymagany rezultat badań umożliwiających wsparcie procesu decyzyjnego. Zaleca się, aby procedura definiowania systemu umożliwiała ogólny opis systemu z uwzględnieniem jego granic i funkcjonalnych sprzężeń, środowiska, rodzajów przepływającej energii, materiałów i informacji oraz warunków użytkowania objętych prowadzoną analizą ryzyka.

Prowadzona w następnym kroku analiza zagrożeń ma na celu określenie źródeł ryzyka, które mogą spowodować wystąpienie niepożądanego zdarzenia i prowadzić do obniżenia poziomu bezpieczeństwa. Norma zaleca określanie znanych zagrożeń z równoczesnym wykorzystaniem metod formalnych pozwalających ustalić zagrożenia wcześniej nierozpoznane, możliwe w specyficznych sytuacjach.

Zidentyfikowane zagrożenia są analizowane w trzecim kroku procedury w celu określenia wielkości ryzyka, jakie z nimi jest związane. Odwołując się do definicji ryzyka, zgodnie z którą wielkość ryzyka jest funkcją prawdopodobieństwa zajścia zdarzenia niepożądanego i skutków jego wystąpienia (konsekwencji), każde zagrożenie jest analizowane pod kątem prawdopodobieństwa wypadku i związanych z nim potencjalnych



Rys. 1. Zalecany przez normę [2] algorytm analizy ryzyka

konsekwencji. Zdarzenia, które doprowadzają do wzrostu zagrożeń często mają charakter lokalny (uszkodzenie jednego z elementów), natomiast konsekwencje zdarzenia przeważnie dotyczą całego systemu i objawiają się na jego granicach, a ponadto przez sprzężenia funkcjonalne mogą zaburzać pracę innych układów. Dlatego jednym z zadań analizy zagrożeń jest ustalenie możliwych związków przyczynowo-skutkowych pomiędzy określonym zagrożeniem a związanymi z nim konsekwencjami, dla dowolnego poziomu analizowanego systemu. Wynika stąd konieczność prowadzenia zarówno analizy częstości zdarzeń jak i analizy konsekwencji. W tym ostatnim przypadku musi być podjęta decyzja odnośnie rodzaju „konsekwencji”, które będą następnie analizowane. Często są to konsekwencje możliwie najgorsze lub konsekwencje, które mogą wystąpić najczęściej. Decyzja dotycząca wyboru rodzaju konsekwencji jest podejmowana w pierwszym kroku analizy ryzyka i ma wpływ na przebieg analizy zagrożeń. W tym miejscu należy zwrócić uwagę na możliwość dominacji subiektywnych ocen, które jeśli nie są wsparte znajomością przedmiotu badań, mogą istotnie wypaczyć wyniki estymacji poziomu ryzyka, dokonywane na tym etapie analizy. To samo dotyczy analizy częstości i analizy niepewności. Analizując obiekt techniczny nie zawsze dysponujemy pełnym opisem zjawiska, co często prowadzi do uproszczeń i pominięcia istotnych czynników procesów zużycia i degradacji. Dlatego zaleca się szczególną rozważność w doborze metody

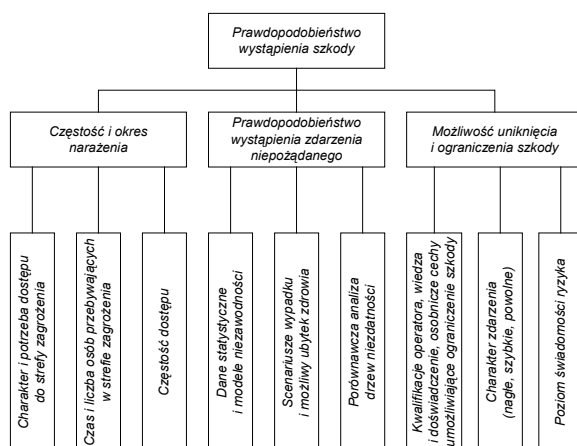
lub metod analizy, które można pogrupować w następujący sposób [3]:

- opracowania statystyczne danych pochodzących z retrospekcji;
- wykorzystanie modeli analitycznych i/lub symulacyjnych do obliczania prawdopodobieństwa wystąpienia zdarzeń;
- wykorzystanie ocen ekspertów.

Sposób zastosowania zależy od oceny stanu rzeczy. Jest możliwe łączne zastosowanie metod, wykorzystanie jednej z nich do przeprowadzenia analizy sprawdzającej, a nawet przeprowadzenie analiz wybranymi metodami jedynie w celu dostarczenia materiału pomocniczego dla eksperta. Ostateczne wyniki identyfikacji i analizy zagrożeń pozwalają przejść do ostatniego etapu analizy ryzyka - szacowania ryzyka.

Zgodnie z istotą definicji ryzyka, w procesie jego szacowania analizowany jest szereg aspektów podstawowych elementów ryzyka. Na przykład szacując stopień ciężkości możliwej szkody, należy przede wszystkim ustalić, czy szkody będą dotyczyć również ludzi, jakiego typu mogą być urazy (odwracalne, nieodwracalne) lub czy mogą być ofiary śmiertelne. Natomiast analizując zakres szkody należy ustalić, czy odnosi się to do wieku osób, czy też należy uwzględnić jedynie operatora. Podobna procedura ustalenia ciężkości szkody powinna być zastosowana w przypadku szacowania mienia lub degradacji środowiska.

Szczególne miejsce zajmuje ocena wpływu czynnika ludzkiego. Jak wynika z rysunku 2 szkolenie, doświadczenie i wiedza, świadomość ryzyka, zapewnienie warunków pracy minimalizujących negatywny wpływ czynników zewnętrznych i wewnętrznych na niezawodność operatora mogą istotnie przyczynić się do zmniejszenia ryzyka.



Rys. 2. Czynniki kształtujące prawdopodobieństwo wystąpienia szkody

Często procedura szacowania ryzyka wymaga wykonania dodatkowych czynności takich jak ustalenie wymaganego poziomu zaufania odnośnie przedstawionych rezultatów lub analiz

porównawczych, możliwych korzyści w stosunku do obliczonego poziomu ryzyka. Zaleca się, aby metody użyte w oszacowaniu ryzyka miały charakter ilościowy, a dopiero w przypadku braku możliwości zastosowania takich metod dopuszcza się wykonanie porównawczego rankingu ryzyka, który również powinien być wynikiem zastosowania metod ilościowych.

Oszacowanie ryzyka powinno umożliwić ustalenie rankingu scenariuszy wypadków, który może mieć charakter jakościowy. Służy temu opracowanie macierzy ryzyka, opisującej poszczególne poziomy ryzyka (tabela 1).

Tabela 1. Macierz ryzyka

Kategorie prawdopodobieństw	Częstość występowania na rok	Kategorie konsekwencji			
		Katastroficzne	Ciężkie	Poważne	Lekkie
Bardzo prawdopodobne	>10-1	H	H	H	I
Prawdopodobne	10-1 ÷ 10-3	H	H	L	L
M mało prawdopodobne	10-3 ÷ 10-5	H	H	L	L
Bardzo mało prawdopodobne	10-5 ÷ 10-7	H	I	L	T
Wyjątkowe	<10-7	I	I	T	T

Przyjęte za normą [1] oznaczenia klas konsekwencji oznaczają ryzyko:

H - wysokie, I - średnie, L - małe, T - pomijalne.

Należy dodać, że kategorie konsekwencji ustalono w następujący sposób:

- *katastroficzne* - wiele ofiar śmiertelnych i praktycznie całkowite zniszczenie analizowanego systemu;
- *ciężkie* - niewiele ofiar śmiertelnych i rozległe szkody materialne w analizowanym systemie;
- *poważne* - brak ofiar śmiertelnych ale są osoby z poważnymi obrażeniami, poważną chorobą zawodową oraz znaczne szkody materialne w analizowanym systemie;
- *lekke* - osoby z lekkimi obrażeniami lub niewielkie szkody materialne w systemie.

Ten sposób postępowania umożliwia skoncentrowanie się na ilościowej ocenie scenariuszy o najwyższym poziomie ryzyka z równoczesnym odrzuceniem tych scenariuszy, w których przewidywany poziom ryzyka jest pomijalnie niski. W celu przedstawienia obliczonego ryzyka zaleca się wybór odpowiedniej charakterystyki. Do najczęściej stosowanych należy zaliczyć:

- ryzyko indywidualne, wyrażone przez przewidywaną częstość zgonów lub zachorowań;

- wykresy $F-N$ - przedstawiające na płaszczyźnie krzywe poziomu ryzyka w funkcji częstości - F i konsekwencji - N , przy czym N przedstawia aproksymowaną liczbę osób, które mogą doznać określonych szkód lub skumulowaną wielkość materialnych czy pozamaterialnych szkód;
- możliwą, hipotetyczną statystyczną wielkość strat, którymi mogą być: liczba ofiar śmiertelnych, wartość strat materialnych lub środowiskowych.

Integralną częścią oszacowania ryzyka jest analiza niepewności zarówno odnośnie danych, jak i metod oraz metod zastosowanych w dotychczas przeprowadzonych etapach analizy ryzyka. Analiza niepewności często jest rozumiana jako analiza wrażliwości. W rzeczywistości, celem analizy wrażliwości jest badanie wpływu zmian poszczególnych parametrów modelu na zmiany odpowiedzi modelu, natomiast celem analizy niepewności jest określenie zmienności i dokładności otrzymanych wyników w zależności od przyjętych założeń odnośnie modelu i zbioru określających go parametrów. Efektem tak prowadzonej analizy powinno być znalezienie odpowiedzi na pytanie jaki wpływ na niepewność otrzymanych rezultatów szacowania ryzyka ma przyjęcie określonego modelu i jakie źródła niepewności mają największy wpływ na końcowy wynik oszacowań.

Zgodnie z propozycjami zawartymi w literaturze [4,5,6] do podstawowych zadań zaawansowanego monitoringu bezpieczeństwa zaliczono: zdefiniowanie zakresu monitoringu oraz sposobu ograniczeń i prezentacji informacji o stanach i wartościach alarmowych; wybór metod i środków umożliwiających prowadzenie monitoringu i wnioskowanie on-line, w sposób umożliwiający wczesne wykrywanie eskalujących zaburzeń i ekstrahowanie z ogólnych sygnałów o anomalii działania cech właściwych powstającym uszkodzeniom; kontrolowanie uszkodzeń i podejmowanie działań korekcyjnych przez operatora w celu minimalizacji, w szczególności uniknięcie wystąpienia zdarzeń niepożądanych a prowadzących do wystąpienia niebezpiecznych konsekwencji; opracowanie prognozy przyszłych zdarzeń na podstawie aktualnych obserwacji i odnotowanych trwałych zmian parametrów, wykrytych za pomocą analizy rezultatów i analiz pomiarów zebranych w bazie danych. Ostatni punkt jest szczególnie ważny w przypadku monitorowania stanu elementów i zespołów mechanicznych oraz pozostałych składowych ulegających procesowi degradacji i zużycia zmęczeniowego, dla których identyfikacja wczesnych faz uszkodzeń może zapobiec wystąpieniu katastroficznej fazy uszkodzenia, w tym zniszczeniu całego systemu.

Współczesne propozycje rozwiązań techniczno-organizacyjne [4] dodatkowo uwzględniają rozwój

i aplikację technologii samoutrzymania przy zachowaniu wysokiej efektywności produkcji i możliwie niskich kosztach eksploatacji.

Stąd analiza ryzyka i zmniejszenie niepewności oszacowań niezawodności stały się krytycznymi metodami w procesie podejmowania decyzji strategicznych, których celem jest zapewnienie bezpieczeństwa technicznego i minimalizacja kosztów.

Zarówno w Europie jak i w Ameryce za najbardziej efektywne metody uznano *RBI* (Risk Based Inspection) – nadzór wsparty analizą ryzyka oraz *RBM* (Risk Based Maintenance) – eksploatacja zorientowana na bezpieczeństwo. W metodach *RBI* i *RBM* odwołujących się do identycznych pojęć i sposobów opisu systemu, podsystemów, podobnej klasyfikacji, funkcjonalnych uszkodzeń i rodzajów uszkodzeń jak w metodzie *RCM* (niezawodnościowo zorientowanej eksploatacji) podejmowane są próby obliczenia ryzyka, w odróżnieniu od metody *RCM*, w której celem jest określenie stopnia (klasy) krytyczności analizowanego obiektu lub elementu.

W metodzie *RCM* dążymy do zakwalifikowania rodzaju uszkodzenia do określonej klasy, przypisując równocześnie wszystkim rodzajom uszkodzeń danej klasy to samo znaczenie i rozległość konsekwencji.

Taki sposób postępowania można przyjąć za satysfakcjonujący jedynie wówczas, gdy wartość strat jest mała i wystarczy jedynie uwzględnić prawdopodobieństwo wystąpienia zdarzenia niepożądanego. Podstawą obliczeniową poziomu ryzyka w metodach *RBI* i *RBM* jest jego definicja, zgodnie z którą jest to oczekiwana wartość prawdopodobnej przyszłej straty określona przez iloczyn prawdopodobieństwa zdarzenia niepożądanego i możliwych konsekwencji.

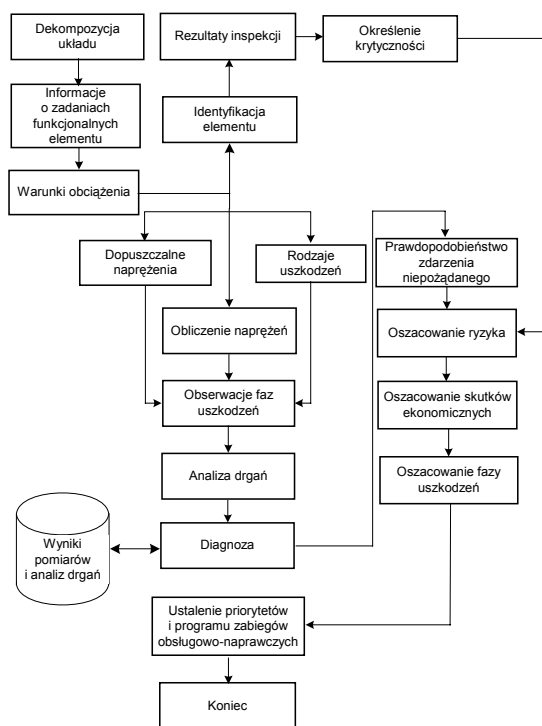
Uwzględniając rozrzut ocen czasu do wystąpienia awarii, głównym celem metod predykcji wspartych analizą ryzyka jest zmniejszenie niepewności prognozy. Jednym z istotnych sposobów rozwiązania tego problemu jest modelowanie i diagnozowanie procesów degradacyjnych i zmęczeniowych, tym samym zmniejszenie wariancji ocen residualnego życia obiektu.

Realizacja tego celu wymaga oceny strukturalnej niezawodności systemu z uwzględnieniem detekcji i analizy procesów degradacji wszystkich składowych w uprzednim i aktualnym okresie użytkowania. Na tej podstawie możliwe jest zmniejszenie niepewności odnośnie oceny okresy czasu do awarii oraz prognozy wystąpienia i rozwoju innych typów uszkodzeń.

Takie ujęcie wymaga dostępu do odpowiedniej bazy danych o możliwych uszkodzeniach składowych systemu oraz wiedzy zgromadzonej na podstawie doświadczeń zdobytych przez odpowiednio przygotowanych pracowników.

Ogólnie w omawianej metodzie wyróżnia się następujące etapy [7]:

1. Dekomponowanie analizowanego systemu na podukłady i elementy a następnie oszacowanie i ich krytyczność w najbardziej niekorzystnych scenariuszach zdarzeń.
2. Określenie czasu i zakresu kolejnej inspekcji.
3. Oszacowanie poziomu ryzyka każdego podukładu i elementu.
4. Wyznaczenie priorytetów utrzymania elementów zgodnie z ustalonym rankingiem poziomu ryzyka technicznego.



Rys. 3. Schemat procedury RBI wspartej diagnostyką wibroakustyczną

Ogólnie, ocena ryzyka powinna uwzględniać każdy rodzaj uszkodzenia, którego wystąpienia można oczekiwać, odnośnie poszczególnych elementów i podzespołów i obejmować zarówno oszacowanie prawdopodobieństwa wystąpienia uszkodzenia jak rozległości i wysokości strat (konsekwencji). Procedury estymacji prawdopodobieństwa wystąpienia zdarzenia niepożądanego powinny uwzględniać rzeczywisty stan rozwoju uszkodzenia, ważniejsze uszkodzenia, rezultaty inspekcji i możliwe przyszłe uszkodzenia.

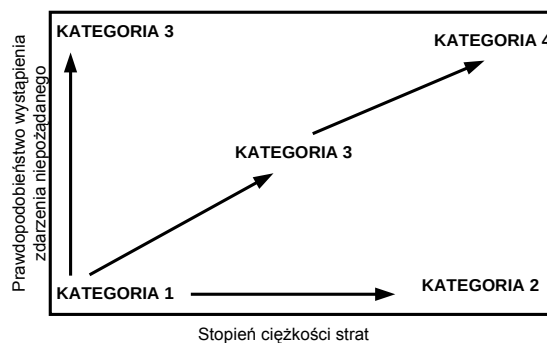
Podobnie oszacowania powinny uwzględniać zarówno aspekty bezpieczeństwa jak i zagadnienia strat ekonomicznych spowodowanych uszkodzeniami wtórnymi, w tym koszty napraw, straty spowodowane przerwą w produkcji i wypłatą rekompensat za utracone zdrowie lub życie.

Otrzymane rezultaty oszacowań ryzyka powinny być zestawione w postaci macierzy ryzyka, która w naturalny sposób pozwala określić kategorie scenariuszy i ustalić priorytety. Schemat przedstawiający możliwość zastosowania metody RBI wspartej diagnostyką wibroakustyczną zaprezentowano na rysunku 3.

Z tego punktu widzenia powinien być analizowany każdy z procesów degradacji i zmęczenia, które mogą spowodować powstanie uszkodzenia w analizowanym węzle układu. Chodzi zarówno o prawdopodobieństwo powstania uszkodzenia jak i związane z tym konsekwencje tak dla analizowanego węzła jak i całego układu.

Przedstawiając tok postępowania należy zwrócić uwagę na znaczenie i różnorodność zadań, jakie mają do spełnienia procedury odpowiednio analizy jakościowej i ilościowej. W pierwszym przypadku, odwołując się do doświadczeń ekspertów z dziedzin projektowania, utrzymania, użytkowania i prac badawczo-rozwojowych ustalane są czynniki, które mogą mieć największy wpływ na poziom bezpieczeństwa technicznego i na podstawie analizy porównawczej ustalana jest ich względna waga i pozycja w macierzy ryzyka.

W szacowaniu prawdopodobieństwa wystąpienia uszkodzenia uwzględnia się wpływ warunków eksploatacji na możliwość wystąpienia uszkodzenia, wpływ wcześniejszych uszkodzeń, jakości i zakresu inspekcji eksploatacyjnych, prawdopodobieństwa wystąpienia uszkodzenia w określonej przyszłości. Podobnie szacowane są konsekwencje, szczególnie ciężkość strat i prawdopodobieństwo wystąpienia najgorszego scenariusza. Utworzona w ten sposób macierz ryzyka stanowi podstawę do ustalenia kategorii ryzyka, priorytetu i zakresu inspekcji oraz o sposobu i kierunku zmian, jakim powinny być poddane procedury utrzymania ruchu. Przykładowy podział na kategorie przedstawiono w pracy [4].



Rys. 4. Kategorie ryzyka wg [4]

Odpowiednio każdej kategorii przypisywany jest następujący zakres działań:

KATEGORIA 1 – (akceptowalna) nadzór eksploatacyjny nie jest niezbędny, oprócz inspekcji przewidzianych odrębnymi regulacjami.

KATEGORIA 2 – (akceptowalny, wymagana kontrola) należy prowadzić właściwe inspekcje z zakresu utrzymania ruchu i nadzór diagnostyczny.

KATEGORIA 3 – (niepożądana) poziom ryzyka powinien być obniżony przy następnej inspekcji; rozważyć: poprawę procedur inspekcyjnych, metod zarządzania operacyjnego, wprowadzenie monitoringu,

środków i metod redukcji możliwości powstawania i rozwoju uszkodzeń.

KATEGORIA 4 – (nieakceptowalna) poziom ryzyka natychmiast obniżyć do KATEGORIA 2 lub KATEGORIA 1.

Zauważmy, że generalnie oszacowanie ryzyka wykonane zostało na podstawie jakościowych ocen ryzyka (bardzo małe, małe, średnie i wysokie prawdopodobieństwo oraz bardzo małe, średnie, poważne i krytyczne straty) bez nadawania rozważanym odpowiednich wartości.

To zadanie realizowane jest w ramach oceny ilościowej, która w inżynierii bezpieczeństwa dokonywana jest na podstawie rozwiązania równania bezpieczeństwa określającego minimalne prawdopodobieństwo wystąpienia uszkodzenia [3]:

$$Z = R - S \quad (1)$$

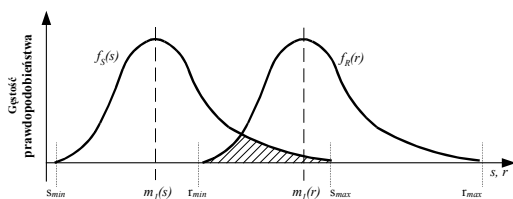
gdzie:

R - obciążalność, np. wytrzymałość,

S - obciążenie, np. naprężenia.

Zakładając, że zarówno obciążenie S , jak i obciążalność R są zmiennymi losowymi można określić ich funkcje gęstości rozkładu prawdopodobieństwa $f_R(x)$, $f_S(x)$.

Sposób wyznaczenia funkcji Z schematycznie przedstawiono na rysunku 5, gdzie obszar zakreskowany oznacza możliwość wystąpienia $Z < 0$, interpretowanego jako prawdopodobieństwa wystąpienia zdarzenia niepożądanego.



Rys. 5. Przykładowe funkcje gęstości prawdopodobieństwa obciążeń i obciążalności

Zauważmy, że prawdopodobieństwo to zależy od następujących parametrów:

- względnego usytuowania obu krzywych, - parametrami, które mogą reprezentować to położenie są wartości oczekiwane i odchylenia standardowe naprężeń i wytrzymałości;
- wpływu skośności i spłaszczenia w funkcji gęstości prawdopodobieństwa dla rozkładów niegausowskich

Jednym z elementów procedury zmniejszenia poziomu ryzyka jest minimalizacja prawdopodobieństwa wystąpienia uszkodzenia. Wyznaczenie prawdopodobieństwa uszkodzenia P_f można przedstawić w różny sposób, na przykład [8]:

$$P_f = \int_{-\infty}^{s_{\max} + \infty} \int_{r_{\min}}^{r_{\max}} f_{R,S}(r, s) dr ds \quad (2)$$

gdzie:

r , s - odpowiednio zmienne.

Przy założeniu, że f_S i f_R są niezależne oraz uwzględnieniu technicznych ograniczeń odnośnie minimalnych obciążeń i maksymalnej obciążalności zależność (2) zapiszemy:

$$P_f = \int_{s_{\min}}^{s_{\max}} \int_{r_{\min}}^{r_{\max}} f_R(r) f_S(s) dr ds \quad (3)$$

Stąd zdarzenie odpowiadające uszkodzeniu wystąpi wówczas, gdy:

$$Z = (R, S) \Rightarrow P_f = P(Z(R, S)) \leq 0 \quad (4)$$

Przy założeniu, że zmienne opisujące obciążalność i obciążenie mają rozkłady normalne, rozważmy możliwość wyznaczenia krzywej rozgraniczającej obszar zagrożony uszkodzeniem od pozostałego:

$$Z = 0 \quad (5)$$

W ogólności granicę rozdzielającą obszar rozwiązań bezpiecznych od obszaru prawdopodobnych uszkodzeń rozwiązuje się aproksymując powszechne odpowiedzi [9]

$$\hat{z}(x) = Q_0 + \sum_{i=1}^d Q_i + \sum_{i=1}^d \sum_{j=1}^d Q_{ij} x_i x_j \quad (6)$$

Zagadnienie wyznaczania wartości zapewniających właściwe dopasowanie do rzeczywistej powierzchni odpowiedzi stało się impulsem do rozwoju wielu nowych metod uczenia statystycznego. Jako przykład sieci neuronowej stosowanej w tego typu zagadnieniach należy wskazać perceptron wielowarstwowy, który implementuje funkcję:

$$\hat{z}(x) = \hat{h} \left(\sum_{k=0}^m w_k h \left(\sum_{i=1}^d w_{ki} x_i \right) \right) \quad (7)$$

gdzie: $h(t)$ najczęściej ma postać funkcji

$$\text{sigmoidalnej } h(t) = \frac{1}{1 + e^{-t}}.$$

Z drugiej strony wykorzystując metodę wektorów podpierających (SVM – Suport Vector Machine) podejmowane są próby separacji obszarów bezpiecznego i niebezpiecznego. Odpowiedni model ma postać:

$$\hat{z}(x) = \sum_{i=1}^m w_i K(x, x_i) - b \quad (8)$$

gdzie:

$K(x, x_i)$ jest jądrem funkcji

b - parametr.

W literaturze wskazuje się wiele różnych metod wyznaczania funkcji $K(x, x_i)$. Dla przybliżeń tego problemu rozważamy funkcję w postaci wielomianu $K(x, y) = \langle x, y \rangle^2$.

Przy założeniu, że rozważane wielkości x , y są wektorami o wymiarze $d=2$ otrzymamy:

$$K(x, y) = x_1^2 y_1^2 + x_2^2 y_2^2 + 2x_1 y_1 x_2 y_2 \quad (9)$$

W praktyce unikając wielu trudności obliczeniowych, wielomianowa funkcja $K(x, y)$ stosowana jest w postaci

$$K(x, y) = (\langle x, y \rangle + Q)^p \quad (10)$$

Rozwiązanie tej funkcji dla $p=2$ będzie miało postać:

$$K(x, y) = \sum_{i=1}^d \sum_{j=1}^d x_i x_j y_i y_j + 2Q \sum_{i=1}^d x_i y_i + Q^2 \quad (11)$$

Dodatkowym problemem, który w tym zadaniu powinien być uwzględniony jest wpływ procesów degradacji na szczegółowe zadania klasyfikacji, ustalania zależności regresyjnych i ewolucji rozkładów prawdopodobieństwa.

3. EWOLUCJA ROZKŁADÓW PRAWDOPODOBIENSTWA OBCIĄŻEŃ I OBCIĄŻALNOŚCI

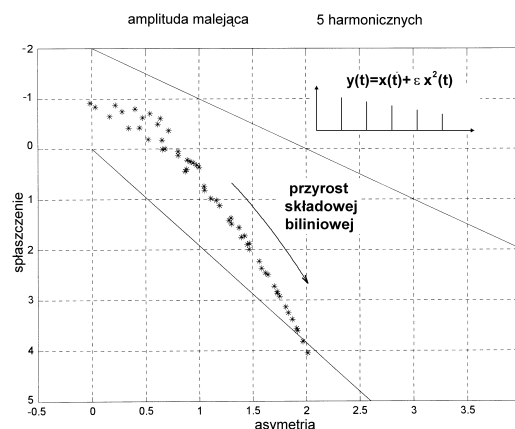
W dotychczas rozważanych modelach określano prawdopodobieństwo wystąpienia uszkodzenia przy założeniu niezmienności badanych rozkładów w czasie eksploatacji. W rzeczywistości, w wyniku zachodzących procesów zużycia i związanych z tym zmian warunków współpracy elementów i węzłów kinematycznych można zaobserwować ewolucję rozkładów prawdopodobieństwa zarówno w sensie ilościowym (zmiana parametrów funkcji gęstości prawdopodobieństwa) jak i jakościowym (zmiana funkcji opisującej rozkład). Dodatkowo, procesy degradacyjne towarzyszące wykonaniu zadań funkcjonalnych mogą wywołać podobną zmienność rozkładów prawdopodobieństwa opisującego obciążalność. W tym przypadku można oczekiwać, że położenie linii rozgraniczającej oraz prawdopodobieństwo wystąpienia uszkodzenia będą zależeć od czasu eksploatacji. W ostatnich latach opracowano szereg metod badania i analizy funkcji niezawodności zależnej od czasu [3].

Postać gęstości prawdopodobieństwa uwzględniająca wystąpienie zaburzeń związanych z występowaniem dodatkowych sprzężeń nieliniowych przedstawiono na rysunku 6, wiąże się to z koniecznością analizy momentów trzeciego i czwartego rzędu, a w dziedzinie częstotliwości odpowiednio bispektrum i trispektrum [11].

Cenność informacji zawartej w bispektrum polega na tym, że pozwala ono badać statystyczne zależności pomiędzy poszczególnymi składowymi widma oraz wykrywać składowe powstałe w wyniku występowania efektów nieliniowych i dodatkowych sprzężeń zwrotnych związanych

z powstającymi uszkodzeniami. Wynika to z faktu, że w odróżnieniu od widma mocy, które jest dodatnie rzeczywiste, funkcja bispektrum jest wielkością zespoloną, zachowującą informację zarówno

o rozkładzie mocy względem poszczególnych składowych widma jak i o zmianach fazy.



Rys. 6. Przykład zmian wartości wyższych momentów wywołanych nieliniowym zaburzeniem

Ogólnie można powiedzieć, że bispektrum może być interpretowane jako miara „ilości” sprzężenia pomiędzy składowymi częstotliwościowymi, wyznaczona na dwuwymiarowej płaszczyźnie o odpowiednich współrzędnych częstotliwościowych.

Przy założeniu, że analizowany przez nas sygnał, został poddany dyskretyzacji, odpowiednie podstawowe estymatory dla wyrazu liniowego i biliniowego przyjmą postać:

$$C_{2x}(k) = E\{x^*(n)x(n+k)\} \quad (13)$$

$$C_{3x}(k, l) = E\{x^*(n)x(n+k)x(n+l)\} \quad (14)$$

Z kolei zależności określające spektrum i bispektrum otrzymamy w postaci:

$$S_{2r}(f) = \sum_{k=-\infty}^{\infty} C_{2x}(k) e^{-j2\pi f k} \quad (15)$$

$$S_{3r}(f_1, f_2) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} C_{3x}(k, l) e^{-j2\pi f_1 k} e^{-j2\pi f_2 l} \quad (16)$$

Zauważmy, że bispektrum pozwala określić zależności pomiędzy istotnymi częstotliwościami badanego układu dynamicznego. Duża wartość bispektrum dla określonych para częstotliwości i kombinacjami ich sum lub różnic, wskazuje na istnienie sprzężenia częstotliwościowego pomiędzy nimi. Może to oznaczać, że rozważane częstotliwości, będące składowymi sum mają wspólny generator, co w obecności nieliniowości wyższych rzędów może prowadzić do syntetyzowania wspomnianych nowych składowych częstotliwościowych.

Jednym z powodów występowania efektów nieliniowych jest kwadratowe sprzężenie fazy (Quadratic Phase Coupling – QPC). QPC może być w ten sposób rozpatrywane jako wskaźnik

zależności pomiędzy dwoma składowymi harmonicznymi procesu nieliniowego. Własność ta może wykryć pewien związek fazowy pomiędzy sumą lub różnicą dwóch częstotliwości (Swami i inni 1995). Jest to różnica pomiędzy analizą bispektralną a typowym podejściem do widma mocy, przy pomocy którego nie jest możliwe wykrycie istnienia QPC.

Inną drogą do poznania nieliniowej struktury częstotliwościowej sygnału wibroakustycznego jest intermodulacja jako podstawowy mechanizm określający strukturę charakterystycznych składowych drgań.

4. PODSUMOWANIE

W każdym ujęciu w centrum uwagi staje problem wydobywania odpowiedniej informacji diagnostycznej. Rezultaty badań potwierdzają fakt, że w wyniku zachodzących procesów eksploatacyjnego zużycia następuje jakościowa zmiana warunków współpracy poszczególnych elementów i węzłów kinematycznych, czego efektem jest ewolucja rozkładów prawdopodobieństwa obciążeń. Spostrzeżenie to dotyczy zarówno zmian wartości parametrów funkcji gęstości jak i klasy samego rozkładu. Podobnie zachodzące procesy degradacji mogą wywoływać zmiany w rozkładach prawdopodobieństwa dopuszczalnej obciążalności.

Oznacza to, możemy zastosować iteracyjne procedury zwiększania dokładności prognozy przez uaktualnianie funkcji gęstości prawdopodobieństwa dopuszczalnych wartości obciążeń i obciążalności. Szczegółowy sposób postępowania zależy od zakresu przyjętej diagnozy jak i typu obserwowanego uszkodzenia. Ogólnie można zauważyć, że największe możliwości poprawnego rozwiązania tak sformułowanego zadania wiążą się z faktem dotarcia do takiej informacji diagnostycznej, dla której daje się zbudować model diagnostyczny odwołujący się w swej istocie do procesu powstawania i rozwoju uszkodzenia.

W tak zarysowanym zadaniu istotnym problemem staje się wybór wielkości progowych, określających wybór przedziałów i klas zmienności, odpowiednio dla parametrów rozkładu i funkcji gęstości prawdopodobieństwa.

Podsumowując można zwrócić uwagę na fakt, że za główne przyczyny występowania niepewności ocen i analiz można uznać:

- procedury wyznaczania wartości parametrów, przykładowo mała liczebność próbek może wpływać na szerokość przedziałów ufności, ekstrapolacja danych z jednego obiektu na drugi, zaburzenia w obserwacji danych;
- procedury modelowania, które skierowane na ustalenie istotnych zmiennych przeważnie w niedostateczny sposób uwzględniają ich wzajemne szczegółowe związki;

- natura zjawiska; stan wiedzy nie zawsze pozwala uwzględnić wszystkie istotne czynniki, w szczególności przebieg procesów uszkodzeń.

„Praca naukowa finansowana ze środków na naukę w latach 2005-2007 jako projekt badawczy”

Autor dziękuje mgr inż. Robertowi Gumińskiemu za pomoc w ostatecznej redakcji publikacji.

LITERATURA

1. PN - IEC 60300-3-9, *Analiza ryzyka w systemach technicznych*, 1999.
2. PN - EN 1050, 1999, *Zasady oceny ryzyka*.
3. Radkowski S., *Podstawy Bezpiecznej Techniki*, Oficyna wydawnicza Politechniki Warszawskiej, Warszawa 2003.
4. Yatomi M. At all: *Application of risk Based Maintenance on Materiale Handling Systems*; IHI Engineering Review vol. 37, No 2 str. 52÷58, (2004).
5. Troy V. Ngugen, Harold W. Nelson: *A system Approach to Machinery Condition Monitoring and Diagnostic*.
6. Drożyner P., Veith E., *Risk Based Inspection Methodology Overview*; Diagnostyka Vol. 27, str. 82÷88.
7. Papadopoulos Y., McDermid J., *Automated Safety Monitoring: A Review and Classification of Methods*, International Journal of COMADEM No. 4 str. 14÷32, 2001.
8. Cruse T. A., *Reliability Based Mechanical Design*, Marcel Dekker Inc., New York 1997.
9. Hurtado J. E., *Structural Reliability*, Springer-verlag, Berlin 2004.
10. Radkowski S.: *Wibroakustyczna diagnostyka uszkodzeń niskoenergetycznych*, Instytut Technologii Eksploatacji, Radom 2002.
11. Dybała J., Mączak J., Radkowski S., *Wykorzystanie sygnału wibroakustycznego w analizie ryzyka technicznego*, ITE, Radom 2006 (w przygotowaniu do druku).



Stanisław RADKOWSKI – profesor Instytutu Podstaw Budowy Maszyn PW, kierownik zespołu Diagnostyki Technicznej i Analizy Ryzyka. Prezes Polskiego Towarzystwa Diagnostyki Technicznej.

W pracy naukowej zajmuje się diagnostyką wibroakustyczną i analizą ryzyka technicznego.

ROZRÓŻNIALNOŚĆ USZKODZEŃ W UKŁADACH LINIOWYCH

Jan Maciej KOŚCIELNY, Bolesław DZIEMBOWSKI

Instytut Automatyki i Robotyki Politechniki Warszawskiej
02-525 Warszawa, ul. Św. Andrzeja Boboli 8, fax (22) 849 03 98, {jmk, b.dziembowski}@mchtr.pw.edu.pl

Streszczenie

W referacie przedstawiono różne metody zapisu związku między uszkodzeniami i wartościami sygnałów diagnostycznych dla obiektów diagnozowanych opisywanych modelami liniowymi. Przeprowadzono analizę porównawczą rozróżnialności uszkodzeń uzyskiwanej z zastosowaniem prezentowanych metod. Omówiono nową metodę zwiększania rozróżnialności uszkodzeń bazującą na sekwencji pojawiających się symptomów. Przedyskutowano zalety i ograniczenia poszczególnych metod.

Słowa kluczowe: lokalizacja uszkodzeń, rozróżnialność uszkodzeń, sekwencja symptomów.

FAULT DISTINGUISHABILITY FOR LINEAR SYSTEMS

Summary

Different methods of describing relations between faults and values of diagnostic signals for systems with linear models are presented in this paper. Realize there comparative analysis of faults distinguishability using presented methods. Talk over new method of fault's distinguishability increase based on sequence of emerge symptoms. Discuss advantages and constraint each presented methods.

Keywords: fault isolation, fault distinguishability, sequence of symptoms.

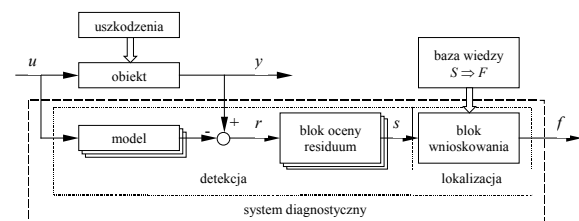
1. WSTĘP

W diagnostyce procesów przemysłowych wyróżniamy dwa podejścia, jedno polegające na analizie sygnałów pomiarowych (*ang. signal based*) oraz drugie wykorzystujące modele procesów (*ang. model based*) [7, 8, 14, 15]. Metody bazujące na modelach zapewniają, w stosunku do metod „sygnałowych”, wcześniejsze wykrywanie oraz rozpoznawanie mniejszych uszkodzeń, dlatego tylko te będą dalej rozważane.

Jednym z głównych wskaźników oceny systemu diagnostycznego jest jego zdolność do rozróżniania uszkodzeń [1, 12]. Dąży się zwykle do uzyskania rozróżnialności wszystkich uszkodzeń. Dokładność uzyskiwanych diagnoz określona jest przez liczbę uszkodzeń wskazywanych w diagnozie [12, 13]. Im liczba ta jest mniejsza tym diagnoza jest bardziej dokładna. Zatem zwiększenie rozróżnialności uszkodzeń umożliwia generowanie diagnoz precyzyjnie wskazujących powstałe uszkodzenia, a w konsekwencji możliwość podjęcia odpowiednich działań zabezpieczających przebieg procesu i dokonania szybkiej wymiany lub naprawy uszkodzonego urządzenia.

Na rys. 1 przedstawiono ogólny schemat diagnozowania na podstawie modeli cząstkowych obiektu. Z wykorzystaniem modeli wyliczane są wartościami sygnałów wyjściowych y . Różnice między wyliczonymi i mierzonymi wartościami

y nazywane są residuami. W stanie bez uszkodzeń i zakłóceń wartości residuów r powinny oscylować w pobliżu zera. Odbiegające od zera wartości residuów są symptomami powstających uszkodzeń f . W celu uproszczenia procesu wnioskowania diagnostycznego w wielu metodach lokalizacji uszkodzeń stosowana jest binarna lub wielowartościowa ocena (kwantyzacja) wartości residuów. W jej wyniku powstają binarne lub wielowartościowe sygnały diagnostyczne s . Na podstawie wartości sygnałów diagnostycznych prowadzona jest lokalizacja uszkodzeń.



Rys. 1. Schemat detekcji i lokalizacji uszkodzeń

2. OPIS UKŁADÓW LINIOWYCH Z UWZGLĘDNIENIEM WPLYWU USZKODZEŃ

Do lokalizacji uszkodzeń niezbędna jest znajomość związku między uszkodzeniami oraz wartościami sygnałów diagnostycznych. Związek

ten może być określony w wyniku modelowania obiektu z uwzględnieniem wpływu uszkodzeń. Uszkodzenia traktowane są jako specyficzne wejścia obiektu. Ogólny opis nieliniowych obiektów dynamicznych przyjmuje postać układu równań stanu i wyjść [7,8]:

$$\dot{\mathbf{x}}(t) = \phi[\mathbf{x}(t), \mathbf{u}(t), \mathbf{f}(t)], \quad (1)$$

$$\mathbf{y}(t) = \psi[\mathbf{x}(t), \mathbf{u}(t), \mathbf{f}(t)]. \quad (2)$$

gdzie: $\mathbf{x}$ – wektor zmiennych stanu obiektu,

$\mathbf{y}$ – wektor wyjść,

$\mathbf{u}$ – wektor wejść,

$\mathbf{f}$ – uszkodzenia.

Układy rzeczywiste zwykle są nieliniowe, ale dla uproszczenia opisu matematycznego przeprowadza się ich linearyzację, co pozwala na sformułowanie przybliżonego opisu liniowego, ważnego w otoczeniu wybranego punktu pracy na charakterystyce statycznej (punkt ten odpowiada najczęściej nominalnym lub uśrednionym warunkom pracy układu). Liniowy układ równań stanu i wyjść ma postać:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{E}\mathbf{f}(t), \quad (3)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) + \mathbf{F}\mathbf{f}(t). \quad (4)$$

Wpływ uszkodzeń na współrzędne stanu oraz wyjścia obiektu odzwierciedlają macierze $\mathbf{E}$ i $\mathbf{F}$.

Inną postacią opisu obiektu liniowego jest model transmitancyjny, zawierający układ równań określających zależność poszczególnych wyjść obiektu od wejść i uszkodzeń [2-8, 12]:

$$\mathbf{y}(s) = \mathbf{G}(s)\mathbf{u}(s) + \mathbf{G}_F(s)\mathbf{f}(s). \quad (5)$$

Poszczególne równania mają postać:

$$y_j(s) = \mathbf{G}_j(s)\mathbf{u}(s) + \mathbf{G}_{Fj}(s)\mathbf{f}(s), \quad (6)$$

przy czym $\mathbf{G}_j(s)$, $j=1, \dots, J$; zawiera transmitancje typu wejście-wyjście:

$$G_{jp}(s) = y_j(s)/u_p(s); \quad p=1, \dots, P, \quad (7)$$

a $\mathbf{G}_{Fj}(s)$, $j=1, \dots, J$; transmitancje dla poszczególnych par uszkodzenie-wyjście:

$$G_{jk}(s) = y_j(s)/f_k(s); \quad k=1, \dots, K. \quad (8)$$

Przy braku uszkodzeń spełniona jest zależność: $y_j(s) - \mathbf{G}_j(s)\mathbf{u}(s) = \mathbf{G}_{Fj}(s)\mathbf{f}(s) = 0$. Residua

wyliczane są z zależności zwanej postacią obliczeniową:

$$r_j(s) = y_j(s) - \mathbf{G}_j(s)\mathbf{u}(s). \quad (9)$$

Równanie (10) (tzw. postać wewnętrzna) odzwierciedla ogólny związek między danym residuum a uszkodzeniami. Dla wszystkich residuów związek ten możemy przedstawić w postaci tab. 1.

$$r_j(s) = \mathbf{G}_{jF}(s)\mathbf{f}(s) = G_{j1}(s)f_1(s) + \dots + G_{jk}(s)f_k(s) + \dots + G_{jK}(s)f_K(s). \quad (10)$$

Tab. 1. Postać wewnętrzna residuów

	f_1	...	f_k	...
r_1	G_{11}		G_{1k}	
...				
r_j	G_{j1}		G_{jk}	
...				

3. METODY ODWZOROWANIA ZWIĄZKU USZKODZENIA - SYGNAŁY DIAGNOSTYCZNE

Modele stosowane do lokalizacji uszkodzeń powinny odwzorowywać przestrzeń residuów lub przestrzeń wartości sygnałów diagnostycznych w dyskretną przestrzeń uszkodzeń. Residua r są ciągłymi sygnałami diagnostycznymi. Odwzorowanie obszarów odpowiadających poszczególnym uszkodzeniom w przestrzeni residuów jest bardzo trudne. Dlatego stosowane są najczęściej uproszczone związki. Zostały one poniżej omówione.

3.1. Residua strukturalne

Postać wewnętrzna residuum $r_j(s) = \mathbf{G}_{Fj}(s)\mathbf{f}(s)$ [5,7,8,12] zawiera informację o uszkodzeniach, które wpływają na wartość residuum. W najprostszym przypadku wrażliwość residuów na poszczególne uszkodzenia można rozpatrywać jakościowo i zapisać w postaci binarnej macierzy diagnostycznej, która jest określona na iloczynie kartezjańskim binarnych (residuów) sygnałów diagnostycznych oraz uszkodzeń. Element tej macierzy przyjmuje wartości:

$$\begin{cases} v_{jk} = 0, & \text{jeśli } G_{jk}(s) = y_j(s)/f_k(s) = 0 \\ v_{jk} = 1, & \text{jeśli } G_{jk}(s) = y_j(s)/f_k(s) \neq 0 \end{cases} \quad (11)$$

Mówimy, że j -te residuum jest wrażliwe na uszkodzenie f_k (wykrywa to uszkodzenie), jeśli $G_{jk}(s) = y_j(s)/f_k(s) \neq 0$. W tab.2 przedstawiono przykład binarnej macierzy diagnostycznej.

Tab. 2. Binarne macierz diagnostyczna

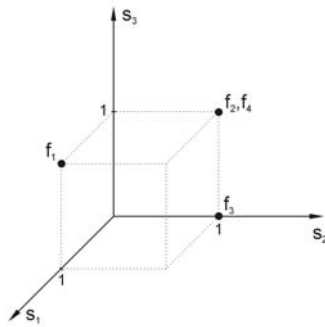
	f_1	f_2	f_3	f_4
s_1	1			
s_2		1	1	1
s_3	1	1		1

Binarne macierz diagnostyczna określa zatem wzorcowe wartości binarnych sygnałów diagnostycznych dla poszczególnych uszkodzeń. Wektor wartości sygnałów diagnostycznych dla k -tego uszkodzenia zapisany w kolumnie macierzy diagnostycznej jest sygnaturą uszkodzenia:

$$V(f_k) = \begin{bmatrix} v_{1k} \\ v_{2k} \\ \dots \\ v_{JK} \end{bmatrix} \quad (12)$$

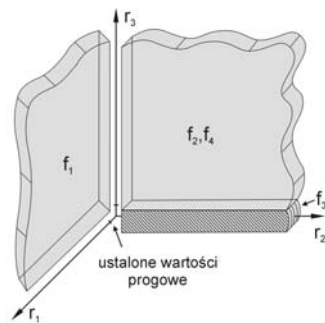
Powyższa postać związku uszkodzenia-sygnały diagnostyczne określana jest także jako residua strukturalne [2, 5, 7, 8, 12] (*ang. structured residuals*).

Graficzna reprezentacja binarnej macierzy diagnostycznej przedstawiona została na rys. 2.



Rys. 2. Graficzna reprezentacja binarnej macierzy diagnostycznej

Opisaną metodę przedstawić można również w przestrzeni residuów (rys. 3).



Rys. 3. Residua strukturalne w przestrzeni residuów zwanej przestrzenią parzystości

Zaletą binarnej macierzy diagnostycznej jest możliwość zaprojektowania nie tylko na podstawie równań obiektu uwzględniających wpływ uszkodzeń, lecz także na podstawie wiedzy eksperckiej.

W przypadku diagnozowania na podstawie residuów strukturalnych przyjmuje się, że wartości poszczególnych residuów oceniane są binarnie (0 – brak symptomu – bliska zera wartość residuum, 1 – symptom – odbiegająca od zera wartość residuum). Binarne sygnały diagnostyczne przyjmują wartość 1, jeśli wystąpi któreś z uszkodzeń, na które jest wrażliwe to residuum.

Diagnoza formułowana jest w wyniku porównania uzyskanych wartości sygnałów diagnostycznych (13) tworzących wektor:

$$C = \begin{bmatrix} c(s_1) \\ c(s_2) \\ \dots \\ c(s_J) \end{bmatrix} \quad (13)$$

z sygnaturą stanu zdatności zawierającą same zera oraz sygnaturami poszczególnych uszkodzeń. Diagnoza wskazuje podzbiór uszkodzeń, których sygnatury są zgodne z aktualnymi wartościami sygnałów diagnostycznych.

$$DGN_B = \{f_k \in F : \forall_{s_j} c(s_j) = v_{jk}\}. \quad (14)$$

3.2. Residua wielowartościowe

Rozszerzeniem binarnej macierzy informacyjnej jest *FIS* (*Fault Isolation System*) wprowadzony w pracach [7-13]. Jest on określony na iloczynie kartezjańskim uszkodzeń i wielowartościowych sygnałów diagnostycznych. Istotne rozszerzenia *FIS* w stosunku do binarnej macierzy diagnostycznej są następujące:

- Dla każdego sygnału diagnostycznego może istnieć indywidualny zbiór jego wartości, przy czym wartość sygnału diagnostycznego $q_j=0$ odpowiada stanowi bez uszkodzeń, a pozostałe wartości są symptomami uszkodzeń.
- Zbiór Q_j wartości j -tego sygnału diagnostycznego zawiera skończoną liczbę wartości (zwykle kilka).
- Dowolny element $q(s_j, f_k) \equiv q_{jk}$ w *FIS* zawierać może zarówno jedną wartość sygnału diagnostycznego jak też ich podzbiór; oznacza to, że w stanie z k -tym uszkodzeniem j -ty sygnał diagnostyczny może przyjmować nie jedną, lecz kilka wartości.

FIS jest zatem tablicą określającą wzorcowe wartości sygnałów diagnostycznych dla poszczególnych uszkodzeń.

Poniżej przedstawiono prosty przykład *FIS*.

Tab. 3. Przykład *FIS*

S/F	f_1	f_2	f_3	f_4	f_5	f_6	V_i
s_1	1	0	1	0	0	1	$\{0,1\}$
s_2	0	-1	0	+1	-1	0	$\{0,+1,-1\}$
s_3	-1	+1	+1,-1	0	+1	+1	$\{0,+1,-1\}$
s_4	+1	0	+1	+1	0	+1,-1	$\{0,+1,-1\}$

Sygnatura k -tego uszkodzenia odpowiada kolumnie *FIS*. Jest ona uogólnieniem sygnatury określonej wzorem (12). Określa ją zależność:

$$Q(f_k) = \begin{bmatrix} q_{1,k} \\ q_{2,k} \\ \dots \\ q_{J,k} \end{bmatrix}. \quad (15)$$

Ponieważ element q_{jk} sygnatury jest w przypadku ogólnym podzbiorem, zatem sygnatura taka jest sygnaturą złożoną, w której wydzielić można sygnatury elementarne jako wektory zawierające różne możliwe kombinacje wartości sygnałów diagnostycznych.

Diagnoza formułowana jest w wyniku porównania uzyskanych wartości sygnałów diagnostycznych tworzących wektor C z sygnaturą stanu zdatności zawierającą same zera oraz sygnaturami poszczególnych uszkodzeń. Diagnoza wskazuje podzbiór uszkodzeń, których sygnatury są zgodne z aktualnymi wartościami sygnałów diagnostycznych.

$$DGN_{FIS} = \{f_k \in F : \forall_{s_j} c(s_j) \in q_{jk}\}. \quad (16)$$

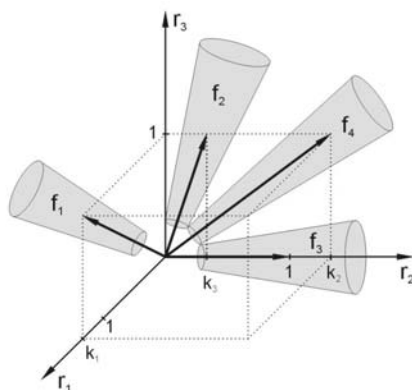
Szczególnym przypadkiem residuów wielowartościowych są residua trójwartościowe, dla których zbiór $Q_j = \{0, -1, +1\}$. W porównaniu z binarną oceną wartości residuów, przy ocenie trójwartościowej brany jest pod uwagę dodatkowo znak residuum. Znak ten ustalany może być na podstawie równań residuów w postaci wewnętrznej (10). Uwzględniany jest znak odpowiedniej transmitancji $G_{jk}(s) = y_j(s)/f_k(s) \neq 0$ oraz znak uszkodzenia f_k . Jeśli uszkodzenie może mieć wartość zarówno dodatnią jak też ujemną, to $q_{j,k} = \{-1, +1\}$. Jeśli natomiast możliwy fizycznie jest tylko jeden znak uszkodzenia (np. wyciek może mieć tylko wartość dodatnią), to $q_{j,k}$ jest jednoelementowy, a znak 1 ustalany jest jako iloczyn znaków transmitancji $G_{jk}(s)$ i uszkodzenia f_k .

3.3. Residua kierunkowe

Mając dany opis wpływu uszkodzeń na wartości residuów w postaci wewnętrznej (10) można określić wrażliwość residuów na poszczególne uszkodzenia nie tylko jakościowo lecz również ilościowo. Wykorzystując tę informację można uzyskać wzmocnienie residuum względem pojawiającego się uszkodzenia:

$$k_{jk} = \lim_{s \rightarrow 0} G_{jk}(s) = \frac{r_j(s)}{f_k(s)}. \quad (17)$$

Wektor wzmocnień residuów odpowiadających danemu uszkodzeniu określa pewien kierunek w przestrzeni residuów (zwanej przestrzenią parzystości), charakterystyczny dla danego uszkodzenia. Każdemu uszkodzeniu odpowiada zatem wektor kierunkowy [2, 5, 7, 8, 12]. Wynika stąd nazwa – residua kierunkowe (*ang. directional residual*).



Rys. 4. Residua kierunkowe w przestrzeni parzystości

W przypadku ogólnym (obiektów nieliniowych) kierunek wektora k -tego uszkodzenia w przestrzeni residuów określony jest przez wektor wartości pochodnych poszczególnych residuów po tym

uszkodzeniu $\partial r_j / \partial f_k$ w stanie ustalonym, tzn. po zaniku przebiegów przejściowych, czyli przy $s \rightarrow 0$. Wpływ uszkodzeń na wartości residuów opisywany jest jak w tabeli 4.

Tab. 4. Wpływ uszkodzeń na wartość residuów

	f_1	f_2	f_3	f_4
r_1	$\partial r_1 / \partial f_1$	$\partial r_1 / \partial f_2$	$\partial r_1 / \partial f_3$	$\partial r_1 / \partial f_4$
r_2	$\partial r_2 / \partial f_1$	$\partial r_2 / \partial f_2$	$\partial r_2 / \partial f_3$	$\partial r_2 / \partial f_4$
r_3	$\partial r_3 / \partial f_1$	$\partial r_3 / \partial f_2$	$\partial r_3 / \partial f_3$	$\partial r_3 / \partial f_4$

Wartości pochodnych w kolumnie odpowiadającej danemu uszkodzeniu określają kierunek tego uszkodzenia w przestrzeni residuów.

Diagnoza formułowana jest w wyniku porównania kierunku wyznaczonego przez bieżące wartości residuów z kierunkami wzorcowymi dla poszczególnych uszkodzeń. Zmienne procesowe dostępne dla systemu diagnostyki obarczone są błędami pomiarowymi. Ponadto modele diagnozowanej instalacji nie są w pełni precyzyjne. Dlatego konieczne jest uwzględnienie obszarów tolerancji przy porównywaniu wektorów. Niewielkie wartości residuów nie powinny powodować wskazania uszkodzenia.

Warunkiem zastosowania opisywanej metody jest dokładna znajomość modelu opisującego zależności pomiędzy uszkodzeniami a residuum.

3.4. Sekwencja symptomów

Wyżej prezentowane metody residuów strukturalnych i kierunkowych nie uwzględniały dynamiki residuów pod wpływem pojawiającego się uszkodzenia. Proponowana w niniejszej pracy metoda bazuje na analizie sekwencji pojawiających się symptomów, której możliwość wykorzystania była sugerowana w pracach [7, 8, 11, 12]. Metoda wykorzystuje równania zgodności (6) w postaci wewnętrznej (10) uwzględniającej wpływ uszkodzeń na residua. Różnica polega na tym, że do dalszej analizy nie jest brana informacja o statycznej wrażliwości danego residuum (wzmocnieniu) (przykład tab. 5), ale informacja o dynamice residuum, zmieniającego się pod wpływem danego uszkodzenia. Różne reakcje residuów na uszkodzenia wpływają na kolejność symptomów, co pozwala rozróżniać te uszkodzenia.

W przypadku równań liniowych dynamikę j -tego residuum pod wpływem wystąpienia k -tego uszkodzenia opisuje pochodna cząstkowa j -tego residuum po k -tym uszkodzeniu $\partial r_j / \partial f_k$. Wpływ uszkodzeń na wartości residuów określony jest jak w tabeli 3. Wartości pochodnych w kolumnie odpowiadającej danemu uszkodzeniu określają dynamikę residuów wrażliwych na to uszkodzenie. Zakładając, że indywidualnie dla każdego uszkodzenia residua zostaną znormalizowane do zakresu $[0, 1]$ (tab. 5) oraz będą podlegać ocenie progowej, można wyznaczyć kolejność pojawiania

się symptomów (sygnałów będących wynikiem binarnej oceny residuów). Normalizacja polega na podzieleniu równania j -tego residuum przez wartość wzmocnienia odpowiadającą k -temu uszkodzeniu, przy założeniu, że pozostałe uszkodzenia przyjmują wartość 0. Operację tę należy przeprowadzić dla wszystkich residuów i uszkodzeń. Wzmocnienie funkcji wrażliwości residuum $\partial r_j / \partial f_k = G_{jk}(s)$ po znormalizowaniu nie ma znaczenia, co było podstawą metody residuów kierunkowych. Kolejność pojawiających się symptomów zależy zatem od wartości stałych czasowych i ewentualnych opóźnień w wyznaczonych funkcjach wrażliwości. Należy jednak podkreślić, że bieżące wartości residuów wyliczane są dla każdego uszkodzenia indywidualnie wg. zależności:

$$r_j(f_k) = r_j / k_{jk}.$$

Założmy, że macierz wrażliwości residuów na uszkodzenia po normalizacji ma postać pokazaną w tab. 4. Poszczególnym uszkodzeniom odpowiadają reakcje residuów z odpowiednią dynamiką.

Tab. 5. Zastępcze znormalizowane transmitancje residuów do uszkodzeń

	f_1	f_2	f_3	f_4
r_1	$\frac{1}{T_1 s + 1}$			
r_2		1	1	$\frac{1}{T_5 s + 1} e^{-T_0 s}$
r_3	1	$\frac{1}{T_2 s + 1}$		1

Zakładając jednakowy próg oceny wartości znormalizowanych residuów, dla których generowane są symptomy danego uszkodzenia, nie ma on wpływu na wyznaczenie kolejności pojawiających się symptomów (tab. 6). Na podstawie znajomości stałych czasowych poszczególnych transmitancji występujących w sygnaturach uszkodzeń można określić kolejność n_{jk} pojawiania się symptomów dla poszczególnych uszkodzeń f_k . Im krótsza stała czasowa znormalizowanej transmitancji $G_{jk}(s)/k_{jk}$, tym szybciej pojawi się symptom.

Kolejność pojawiania się symptomów wynikająca z sygnatur uszkodzeń podanych w tab. 5 została pokazana w tab. 6.

Tab. 6. Macierz sekwencji pojawiających się symptomów uszkodzeń

	f_1	f_2	f_3	f_4
s_1	2			
s_2		1	1	2
s_3	1	2		1

Czasy powstawania symptomów można wyznaczyć analitycznie na podstawie znormalizowanych sygnatur przy założeniu wartości

progowej residuum oraz charakteru zmienności uszkodzenia (np. skokowe). Jednak ze względu na brak znajomości funkcji powstawania uszkodzenia w czasie działanie takie nie jest celowe. Istotne jest to, że niezależnie od charakteru zmienności w czasie uszkodzenia, kolejność powstawania symptomów dla poszczególnych uszkodzeń jest zawsze taka sama.

Diagnozowanie polega na porównywaniu wzorcowych sygnatur sekwencji symptomów będących kolumnami macierzy jak w tab. 6 z kolejnością pojawiających się symptomów w momencie pojawienia się uszkodzenia.

4. ROZRÓŻNIALNOŚĆ USZKODZEŃ

4.1. Rozróżnialność uszkodzeń w binarnej macierzy diagnostycznej

Definicja 1. Dwa uszkodzenia $f_k, f_m \in F$ są nierozróżnialne na podstawie binarnej macierzy diagnostycznej wtedy i tylko wtedy, gdy ich sygnatury (kolumny macierzy) są jednakowe:

$$\forall_{s_j \in S} v_{kj} = v_{mj}. \quad (18)$$

Definicja 2. Rozróżnialność wszystkich uszkodzeń zachodzi wtedy, gdy wszystkie kolumny binarnej macierzy diagnostycznej są różne:

$$\forall_{\substack{f_k, f_m \in F \\ k \neq m}} \exists_{s_j} v_{kj} \neq v_{mj}. \quad (19)$$

Taka rozróżnialność określana jest przez Gertlera [3-5] jako słaba (*weakly isolating*) w odróżnieniu do mocnej rozróżnialności uszkodzeń, która występuje w przypadku sygnatur różniących się wartościami, co najmniej na dwóch pozycjach. Mocna rozróżnialność nie będzie analizowana w niniejszej pracy.

Maksymalna liczba uszkodzeń, które mogą być rozróżnione na podstawie wyników zbioru binarnych sygnałów diagnostycznych $S = \{s_j : j = 1, 2, \dots, J\}$ wynosi $2^J - 1$. Zerowe wartości wszystkich sygnałów diagnostycznych odpowiadają stanowi pełnej zdadności procesu. W praktyce liczbie uszkodzeń rozróżnianych na podstawie takiego zbioru jest zwykle mniejsza.

4.2. Rozróżnialność uszkodzeń w FIS

Zdefiniujmy pojęcia bezwarunkowej i warunkowej nierozróżnialności uszkodzeń w FIS.

Definicja 3. Uszkodzenia $f_k, f_m \in F$ są nierozróżnialne (bezwzględnie nierozróżnialne) w FIS (dla danego zbioru sygnałów diagnostycznych S) wtedy i tylko wtedy, gdy sygnatury tych uszkodzeń są jednakowe:

$$\forall_{s_j \in S} q_{jk} = q_{jm}. \quad (20)$$

Definicja 4. Uszkodzenia $f_k, f_m \in F$ są warunkowo nierozróżnialne w FIS ze względu na wszystkie sygnały diagnostyczne $s_j \in S$ wtedy i tylko wtedy, gdy dla każdego sygnału podzbiory jego wartości odpowiadające uszkodzeniom f_k i f_m mają część wspólną i uszkodzenia te nie są bezwarunkowo nierozróżnialne:

$$\forall_{s_j \in S} q_{jk} \cap q_{jm} \neq \emptyset \wedge \exists_{s_j \in S} q_{jk} \neq q_{jm}. \quad (21)$$

Warunkowa nierozróżnialność uszkodzeń oznacza, że mogą wystąpić wartości sygnałów diagnostycznych spełniające warunek:

$$\forall_{s_j \in S} v_j \in q_{jk} \cap q_{jm}, \quad (22)$$

przy których dane dwa uszkodzenia są nierozróżnialne. Jednak możliwe są inne wartości sygnałów diagnostycznych, przy których te same uszkodzenia są rozróżnialne. Spełniony wtedy jest warunek następujący:

$$\exists_{s_j \in S} [v_j \in q_{jk} \wedge v_j \notin q_{jm}] \vee [v_j \notin q_{jk} \wedge v_j \in q_{jm}]. \quad (23)$$

Definicja 5. Dwa dowolne uszkodzenia są w FIS bezwarunkowo rozróżnialne, jeśli istnieje sygnał diagnostyczny, dla którego podzbiory wartości odpowiadające tym uszkodzeniom są rozłączne:

$$\exists_{s_j \in S} q_{jk} \cap q_{jm} = \emptyset. \quad (24)$$

W FIS pokazanym w tab. 4 uszkodzenia: f_2 i f_5 są bezwarunkowo nierozróżnialne, f_1 i f_3 oraz f_3 i f_6 są warunkowo nierozróżnialne, podczas gdy uszkodzenia f_1 i f_6 są bezwarunkowo rozróżnialne. Uszkodzenie f_4 jest bezwarunkowo rozróżnialne od pozostałych uszkodzeń.

W przypadku wielowartościowej klasyfikacji sygnałów diagnostycznych nie można zatem określić rozróżnialności w sposób jednoznaczny. Zależy ona od kombinacji wartości tych sygnałów. W przypadku niektórych kombinacji uzyskujemy rozróżnialność większą, a w przypadku innych mniejszą.

Maksymalna liczba uszkodzeń, które mogą być rozróżnione na podstawie wyników zbioru trójwartościowych sygnałów diagnostycznych $S = \{s_j : j = 1, 2, \dots, J\}$ wynosi $3^J - 1$. W przypadku n wartościowej oceny wzrasta do wartości $n^J - 1$. Zatem wielowartościowa ocena wartości residuów może, lecz nie musi prowadzić do zwiększenia rozróżnialności uszkodzeń w stosunku do oceny binarnej. W praktyce liczbie uszkodzeń rozróżnianych jest zwykle znacznie mniejsza niż maksymalna. Wynika ona z fizycznych zależności w procesie. Określenie FIS w przypadku wielowartościowych sygnałów diagnostycznych może być dosyć trudne. Dlatego zwykle stosowana jest tylko trójwartościowa ocena residuów (-1, 0, +1), w której uwzględniany jest znak residuum.

Umożliwia to rozróżnienie uszkodzeń, które powodują odchylenie residuum w przeciwnych kierunkach (wartości dodatnie i ujemne).

4.3. Rozróżnialność uszkodzeń w przypadku residuów kierunkowych

W przypadku residuów kierunkowych rozróżnialność uszkodzeń zależy od kierunków wektora residuów charakterystycznych dla poszczególnych uszkodzeń.

Definicja 5. Dwa uszkodzenia $f_k, f_m \in F$ są nierozróżnialne wtedy i tylko wtedy, gdy ich sygnatury są jednakowe lub wartości w poszczególnych kolumnach są proporcjonalne (co odpowiada jednakowym kierunkom wektora residuów dla uszkodzeń nierozróżnialnych) sygnatury (kolumny macierzy) są jednakowe:

$$\forall_{s_j \in S} (k_{kj} = a k_{mj}), \quad (25)$$

gdzie a jest dowolną, skończoną liczbą rzeczywistą.

Założmy, że macierz wrażliwości na uszkodzenia residuów ma postać pokazaną w tab. 7. Poszczególnym uszkodzeniom odpowiadają kierunki w przestrzeni residuów pokazane na rys. 4.

Tab. 7. Macierz wrażliwość residuów na poszczególne uszkodzenia

	f_1	f_2	f_3	f_4
r_1	k_1			
r_2		k_3	1	k_2
r_3	1	1		1

Zastosowanie koncepcji residuów kierunkowych umożliwiło rozróżnienie uszkodzeń f_2 i f_4 , które w przypadku lokalizacji bazującej na residuach strukturalnych byłyby nierozróżnialne.

4.4. Rozróżnialność uszkodzeń w przypadku residuów sekwencyjnych

W przypadku residuów sekwencyjnych rozróżnialność uszkodzeń zależy od kolejności powstawania symptomów dla poszczególnych uszkodzeń.

Definicja 6. Dwa uszkodzenia $f_k, f_m \in F$ są nierozróżnialne wtedy i tylko wtedy, gdy kolejność powstawania symptomów tych uszkodzeń jest taka sama (co odpowiada jednakowym sygnaturom - kolumnom macierzy sekwencji symptomów):

$$\forall_{s_j \in S} (n_{kj} = n_{mj}). \quad (26)$$

Uszkodzenia są rozróżnialne, jeżeli sekwencje symptomów dla poszczególnych uszkodzeń są różne. Jak widać na przykładzie tab. 6 zastosowanie koncepcji sekwencji pojawiających się symptomów pozwoliło uzyskać rozróżnialność uszkodzeń f_2 i f_4 , które dla residuów strukturalnych nie były rozróżnialne.

W wyżej wymienionym przykładzie było prosto wyodrębnić kolejność pojawiających się symptomów, gdyż tylko część residuów posiadała inercję względem występującego uszkodzenia, pozostałe residua dla danego uszkodzenia albo nie były wrażliwe, albo reagowały w sposób proporcjonalny do uszkodzenia. W przypadku, gdy każde residuum posiada pewną inercję, to zakładając, że pojawiające uszkodzenie ma charakter funkcji skoku jednostkowego i dla poziomu generacji symptomów równego 0,63 (zakres zmian residuum jest znormalizowany do zakresu [0,1]) symptomy będą pojawiać się po czasie równym stałej czasowej, tab. 5: $\partial r_1 / \partial f_1 - T_1$, $\partial r_3 / \partial f_2 - T_2$ lub po czasie równym sumie zastępczej stałej czasowej i zastępczego opóźnienia (tab. 5: $\partial r_2 / \partial f_4 - T_5 + T_6$).

Rozważając prosty przykład założmy, że mamy dostępne następujące residua w postaci wewnętrznej z informacją o dynamice w postaci stałej czasowej członu inercyjnego pierwszego rzędu danego residuum względem tego uszkodzenia:

$$r_1 = 20f_1|_{T_{11}} + f_2|_{T_{12}} + f_3|_{T_{13}},$$

$$r_2 = 5f_1|_{T_{21}} + 15f_2|_{T_{22}} + f_3|_{T_{23}}.$$

Jeżeli zachowane są przykładowe relacje pomiędzy stałymi czasowymi: $T_{11} > T_{21}$, $T_{12} < T_{22}$ i $T_{13} > T_{23}$, to czasy pojawiania się symptomów będą wyglądały następująco:

$$t(s_1(r_1))|_{f_1} > t(s_2(r_2))|_{f_1},$$

$$t(s_1(r_1))|_{f_2} < t(s_2(r_2))|_{f_2},$$

$$t(s_1(r_1))|_{f_3} > t(s_2(r_2))|_{f_3}.$$

W związku z tym możemy uzyskać następujące sekwencje pojawiających się symptomów s_1 i s_2 dla uszkodzeń f_1, f_2 i f_3 :

Tab. 8. Sekwencja pojawiających się symptomów

	f_1	f_2	f_3
s_1	2	1	2
s_2	1	2	1

Na tej podstawie możemy stwierdzić, że uszkodzenia f_1 i f_2 oraz f_2 i f_3 nierozróżnialne dla binarnej macierzy diagnostycznej stają się rozróżnialne w przypadku wykorzystania wiedzy o kolejności występowania ich symptomów. Uszkodzenia f_1 i f_3 nadal są nierozróżnialne.

5. PODSUMOWANIE

W pracy przedstawiono nową metodę wnioskowania diagnostycznego dla liniowych obiektów diagnozowania, nazwaną metodą residuów sekwencyjnych. Metoda ta została porównana ze znanymi podejściami: tzw. residuami strukturalnymi, trójwartościowymi oraz kierunkowymi w aspekcie uzyskiwanej rozróżnialności uszkodzeń. Pokazano, że residua wielowartościowe, kierunkowe i sekwencyjne

zapewniają większą rozróżnialność uszkodzeń w stosunku do residuów strukturalnych. Zaletą residuów strukturalnych, a także wielowartościowych jest możliwość określenia związku między uszkodzeniami a wartościami sygnałów diagnostycznych zarówno na podstawie równań zgodności w postaci wewnętrznej, jak też na podstawie wiedzy eksperckiej. Ten drugi sposób jest znacznie prostszy i dlatego jest powszechnie stosowany w praktyce. Uzyskanie równań zgodności, uwzględniających wpływ uszkodzeń na wyjścia obiektu, jest trudne i bardzo pracochłonne, szczególnie w przypadku obiektów złożonych. Jest ono jednak niezbędne do zaprojektowania residuów kierunkowych i sekwencyjnych.

Metoda residuów trójwartościowych może być interpretowana jako uogólnienie residuów strukturalnych i szczególnie przypadek residuów kierunkowych. Poszczególne sygnatyry elementarne wyznaczają bowiem w przestrzeni residuów kierunki, charakterystyczne dla poszczególnych uszkodzeń. Kierunki te leżą tylko na hiperpłaszczyznach wyznaczonych przez osie przestrzeni residuów. Uwzględnienie znaku residuum pozwala w wielu przypadkach uzyskać większą rozróżnialność uszkodzeń niż przy binarnej ocenie wartości residuów.

Pozostałe podejścia nie mogą być ze sobą porównane w sposób ogólny, wykorzystują bowiem inną wiedzę o diagnozowanym obiekcie. Residua kierunkowe bazują na wartościach wzmocnienia transmitancji dla poszczególnych uszkodzeń, a residua sekwencyjne na wartościach stałych czasowych i ewentualnie opóźnień występujących w tych transmitancjach. Wzmocnienia określają kierunek (w stanie ustalonym) charakterystyczny dla danego uszkodzenia w przestrzeni residuów, natomiast stałe czasowe pozwalają wyznaczyć sekwencje powstawania symptomów dla poszczególnych uszkodzeń. Uzyskanie rozróżnialności jedną metodą nie gwarantuje rozróżnialności w przypadku drugiej metody. Zależy ona od własności obiektu diagnozowania. Obie metody uzupełniają się zatem. Zastosowanie obu tych podejść równocześnie podczas wnioskowania diagnostycznego zapewnia uzyskanie największej rozróżnialności uszkodzeń, przy danym zbiorze residuów.

Uzyskiwana rozróżnialność uszkodzeń zależy nie tylko od zastosowanej metody zapisu związków między uszkodzeniami i wartościami sygnałów diagnostycznych, ale także od zbioru sygnałów pomiarowych wykorzystywanych do diagnozowania obiektu i w konsekwencji od zbioru generowanych residuów. Im więcej pomiarów, tym więcej może być wyliczanych residuów, co zwiększa rozróżnialność uszkodzeń [7, 12]. Rozszerzenie zbioru generowanych residuów możliwe jest także bez wprowadzania dodatkowych pomiarów, przez tzw. strukturyzację residuów [5, 7, 8, 12, 15]. Polega

ona na projektowaniu residuów wtórnych, na podstawie zbioru residuów pierwotnych, w taki sposób aby nowe residua były wrażliwe na inne podzbiory uszkodzeń niż residua pierwotne.

LITERATURA

- [1] Basseville M. (2001). *On fault detectability and isolability*, European Journal of Control, Vol 7, No 6., 625-637.
- [2] Chen J., Patton R.J. (1999). *Robust model based fault diagnosis for dynamic systems*. Kluwer Academic Publishers, Boston.
- [3] Gertler J., Singer D. (1990). *A new structural framework for parity equation based failure detection and isolation*. Automatica, Vol. 26, no. 2, 381-388.
- [4] Gertler J. (1991). *Analytical redundancy methods in fault detection and isolation*. IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes - SAFEPROCESS'91, Baden-Baden, Vol. 1, 9-21.
- [5] Gertler J. (1998). *Fault Detection and Diagnosis in Engineering Systems*. Marcel Dekker, Inc. New York - Basel - Hong Kong.
- [6] Frank P. M. (1990). *Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy*. Automatica 26, 459-474.
- [7] Korbicz, J., J.M. Kościelny, Z. Kowalczyk and W. Cholewa (2004). *Fault Diagnosis: Models, artificial intelligence methods, applications*. Springer.
- [8] Korbicz J., Kościelny J.M., Kowalczyk Z., Cholewa W. (2002). *Diagnostyka procesów. Modele. Metody sztucznej inteligencji. Zastosowania*. WNT, 2004.
- [9] Kościelny J.M. (1999). *Application of Fuzzy Logic Fault Isolation in a Three-Tank System*. 14th World Congress IFAC, Beijing, Chiny, P-7e-05-1, 73-78.
- [10] Kościelny J.M., Bartyś M. (2000). *Application of Information System Theory for Actuator Diagnosis*. 4th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Process – SAFEPROCESS' 2000, Budapest, Hungary, June 14-16, Vol.2, 949-954.
- [11] Kościelny J.M., Zakroczymski K. (2001). *Fault Isolation Algorithm that Takes Dynamics of Symptoms Appearances Into Account*. Bulletin of the Polish Academy of Sciences. Technical Sciences. Vol.49, No 2, 323-336.
- [12] Kościelny J.M. (2001). *Diagnostyka zautomatyzowanych procesów przemysłowych*. Akademicka Oficyna Wydawnicza Exit, Warszawa.
- [13] Kościelny J.M., Bartyś M., Rzepiejewski P., Sá da Costa J.: *Actuator fault distinguishability study*. Control Engineering Practice (Elsevier) 14, 2006, 645-652.
- [14] Patton R., Frank P., Clark R. (Eds.) (1989). *Fault diagnosis in dynamic systems. Theory and Applications*. Prentice Hall, Englewood Cliffs, New York.
- [15] Patton R., Frank P., Clark R. (Eds.) (2000). *Issues of fault diagnosis for dynamic systems*. Springer.



Prof. dr hab. inż. **Jan Maciej KOŚCIELNY** jest zatrudniony w Instytucie Automatyki i Robotyki Politechniki Warszawskiej. Zajmuje się problemami diagnostyki procesów przemysłowych, a szczególnie metodami detekcji i lokalizacji uszkodzeń przeznaczonymi dla złożonych instalacji technologi-

cznych oraz systemami diagnostycznymi. Autor ponad 180 publikacji naukowych. Kierował pracami, w wyniku których opracowano i wdrożono w przemyśle systemy monitorowania i diagnostyki procesów. Jest członkiem International Editorial Board czasopisma naukowego *International Journal of Applied Mathematics and Computer Science* oraz redaktorem działowym czasopisma *Pomiary-Automatyka-Kontrola*.



Mgr inż. **Bolesław DZIEMBOWSKI** jest asystentem w Instytucie Automatyki i Robotyki Politechniki Warszawskiej i jednocześnie uczestnikiem Studium Doktoranckiego „Mechatronika”.

INTELIĞENTNY POJAZD MECHANICZNY JAKO MECHATRONICZNY OBIEKT DIAGNOZOWANIA

Stanisław NIZIŃSKI, Sławomir WIERZBICKI

Katedra Budowy, Eksploatacji Pojazdów i Maszyn, Wydział Nauk Technicznych,
UWM Olsztyn, nizinski@uwm.edu.pl

Streszczenie

W pracy zdefiniowano inteligentny pojazd mechaniczny jako system mechatroniczny automatycznej regulacji. Rozpatrzono elementy metodyki diagnostyki takich urządzeń. Przedstawiono koncepcje zintegrowanego informatycznego systemu sterowania, scalającego i koordynującego pracę wszystkich elementów, będącego podstawowym narzędziem inteligentnego pojazdu mechanicznego.

Słowa kluczowe: sztuczna inteligencja, pojazd mechaniczny, mechatronika, diagnostyka.

INTELLIGENT MOTOR VEHICLE AS A MECHATRONIC OBJECT OF DIAGNOSIS

Summary

The paper defines the motor vehicle as a mechatronic system of automatic regulation. Elements of diagnostic methodology for such devices were reviewed. The concept of integrated IT supported control system integrating and coordinating operation of all components and representing the basic tool of an intelligent motor vehicle was presented.

Keywords: artificial intelligence, motor vehicle, mechatronic engineering, diagnostics.

WPROWADZENIE

Obecnie obserwuje się ogromny rozwój jakościowy systemów technicznych, w tym pojazdów mechanicznych. Można powiedzieć, że w chwili obecnej nie ma typowych obiektów technicznych, a raczej bytu mają tylko systemy mechatroniczne, będące tworam mechatroniki, czyli działu nauki i techniki zajmującej się doskonaleniem systemów technicznych z wykorzystaniem metod i urządzeń: mechaniki, elektrotechniki, elektroniki, informatyki, cybernetyki, teorii systemów i innych nauk.

Dąży się do budowy inteligentnych obiektów technicznych tzn. takich, których funkcjonowanie naśladuje inteligentnego człowieka.

Praca [1] definiuje „inteligencję” następująco: jest to właściwość psychiki człowieka polegająca na zdolności do uczenia się na podstawie doświadczenia, zdolności przystosowania się do nowych sytuacji, zdolności rozumienia i operowania pojęciami abstrakcyjnymi, używania wiedzy do manipulowania środowiskiem.

Inteligencja człowieka jest podstawą do tworzenia sztucznej inteligencji, której definicję przedstawiono poniżej.

Inteligencja sztuczna – ang. Artificial Intelligence (AI) [5], dział informatyki, którego przedmiotem jest badanie reguł rządzących tzw. inteligentnymi zachowaniami człowieka, tworzenie modeli formalnych tych zachowań i w rezultacie programów komputerowych symulujących te

zachowania. Do zachowań inteligentnych należy postrzeganie, rozpoznawanie, posługiwanie się językiem, rozwiązywanie problemów i inne.

Można powiedzieć, że przez sztuczną inteligencję rozumie się właściwości specjalistycznych urządzeń cybernetycznych, których istotę stanowi naśladownictwo niektórych funkcji psychicznych człowieka, oraz zdolność do sterowania skomplikowanymi procesami.

Według [7] zasadne byłoby nadawanie miana inteligentnego takiemu tworowi materialnemu, samoistnemu lub stanowiącego część jakiegoś systemu, który ma przynajmniej niektóre z cech inteligentnego człowieka.

Według autora system inteligentny powinien mieć dwie cechy:

- zdolność do samoczynnego wykonywania zadań wynikających z jego przeznaczenia;
- korygowanie własnego działania na skutek zmieniających się warunków zewnętrznych.

Wynikiem rozwoju sztucznej inteligencji i automatyki są roboty I, II i III generacji [1].

Roboty I generacji są wyposażone w układy sterowania śledzącego (nadażania), są zdolne do samodzielnego wykorzystywania prostych zaprogramowanych czynności manipulacyjnych.

Roboty II generacji – dzięki połączonym z komputerem sensorom reagują na dotyk i sygnały dźwiękowe, rozróżniają w pewnym stopniu kształty, barwy, co w następstwie pozwala na identyfikację elementów.

Roboty III generacji, tzw. inteligentne są wyposażone w wielosensorowe układy pełniące funkcje zmysłów; wzroku, dotyku, słuchu, a także umożliwiające komunikację głosową.

Dorobek wymienionych nauk istotnie wpływa na rozwój inteligentnych mechatronicznych pojazdów mechanicznych, które jako urządzenia techniczne nowej generacji wymagają innego podejścia w zakresie: diagnozowania, prognozowania i genezowania, a także sterowania. Do chwili obecnej bardzo mało jest informacji na temat inteligentnych pojazdów mechatronicznych a nawet nie ma ich wyczerpującej definicji za wyjątkiem definicji podanej w pracy [6].

1. POJAZD MECHANICZNY JAKO SYSTEM MECHATRONICZNY AUTOMATYCZNEJ REGULACJI

Pojazdy mechaniczne ze względu na właściwości można podzielić następująco: nieinteligentne, częściowo inteligentne i inteligentne.

Nieinteligentny pojazd mechaniczny jest to urządzenie, w którym podstawowe funkcje są realizowane z udziałem operatora (załogi). Częściowo inteligentny pojazd mechanicznych to maszyna, w której niektóre podstawowe funkcje są realizowane bez udziału operatora. W inteligentnym pojeździe mechanicznym wszystkie zasadnicze funkcje są wykonywane bez udziału operatora. Zasadnicze funkcje pojazdu mechanicznego obejmują: wykrywanie zagrożeń, bezpieczeństwo, ochrona środowiska, komfort jazdy, dynamika, diagnostyka, logistyka.

Inteligentny pojazd mechaniczny jest to obiekt techniczny o następujących właściwościach [3, 4]:

- samoczynnie utrzymuje określony wysoki poziom bezpieczeństwa pasażerów;
- ogranicza wpływ szkodliwych czynników (spaliny, hałas) na środowisko naturalne;
- samoczynnie dopasowuje się do istniejących oporów ruchu;

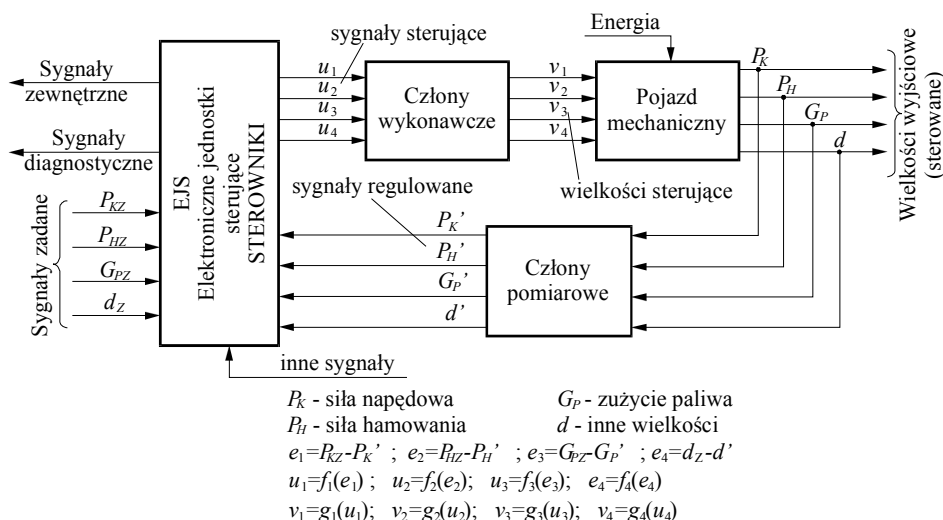
- zabezpiecza określony wysoki poziom komfortu jazdy;
- zabezpiecza wysoki poziom usług logistycznych, eksploatacyjnych i diagnostycznych.

Uzyskanie wymienionych właściwości jest możliwe, jeżeli potraktuje się pojazd mechaniczny jako system mechatroniczny automatycznej regulacji. Każdy system automatycznej regulacji zawiera co najmniej trzy elementy podstawowe: pomiarowy, sterujący i wykonawczy. Elementy te tworzą człony operacyjne regulatora, a łącznie z obiektem regulacji stanowią układ regulacji automatycznej, ze sprzężeniem zwrotnym. W ogólnym przypadku przez sterowanie rozumiemy taką zmianę sygnałów wejściowych, która zapewnia osiągnięcie celu rozpatrywanego systemu.

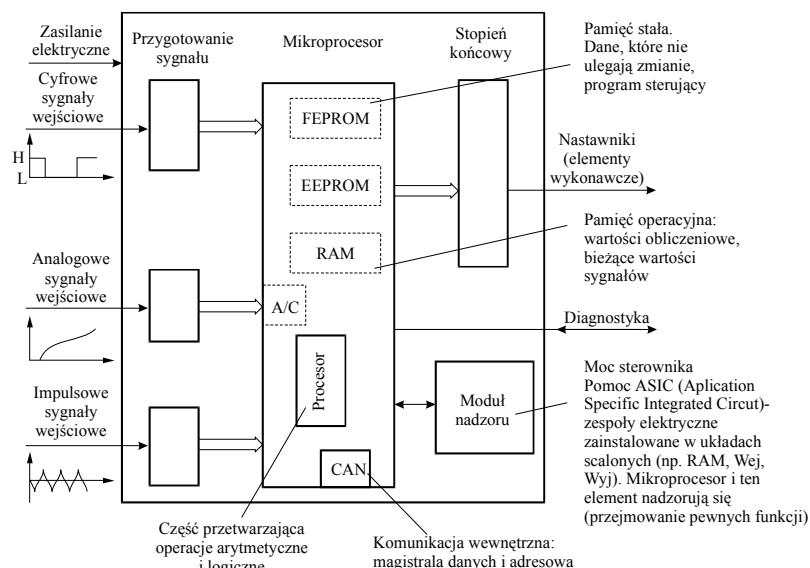
Obserwuje się wprowadzenie do pojazdu wielu nowych układów takich jak: bezpieczeństwa, elektroniczno-informatycznych, specjalnych i innych. Ilustrację graficzną pojazdu jako systemu mechatronicznego automatycznej regulacji przedstawiono na rys. 1. Człony pomiarowe systemu mierzą takie wielkości jak: siła napędowa P_K , siła hamowania P_H , zużycie paliwa G_P , inne d . Mierzone mogą być inne wielkości pośrednie. Zmierzone wartości P_K' , P_H' , G_P' , d' tych wielkości są przekazywane do elektronicznych jednostek sterujących (sterowników), gdzie porównywane są z wartościami wielkości zadanych P_{KZ} , P_{HZ} , G_{PZ} , d_Z .

Wynikiem tej operacji jest ustalenie sygnałów uchybu e_1 , e_2 , e_3 i e_4 . Sygnały uchybu są przetwarzane w sterownikach na sygnały sterujące u_1 , u_2 , u_3 , u_4 , które oddziałują na człony wykonawcze, a te z kolei generują wielkości v_1 , v_2 , v_3 , v_4 sterujące, bezpośrednio oddziałujące na obiekt regulowany, z reguły mechaniczny.

Sercem pojazdu jako automatycznego systemu diagnostycznej regulacji są sterowniki (rys. 2), czyli zmodyfikowany system informatyczny (komputer pokładowy).

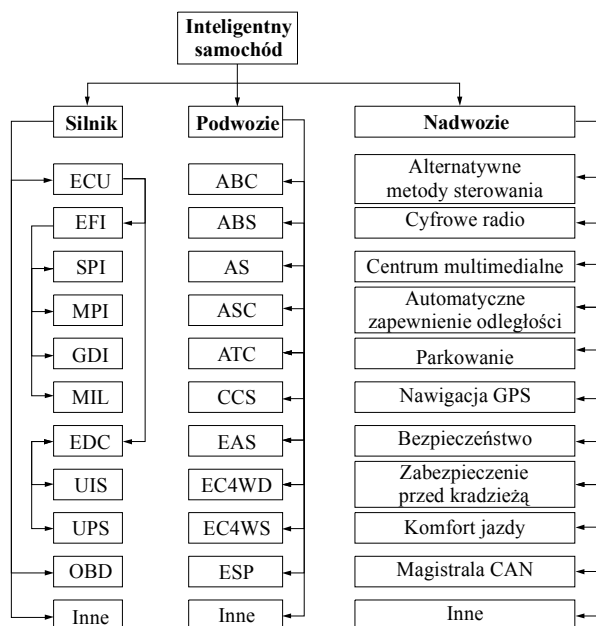


Rys. 1. Pojazd jako mechatroniczny system automatycznej regulacji



Rys. 2. Schemat sterownika stosowanego w układach samochodu

Obecnie pojazd jest wyposażony w wiele układów specjalnych, które zbliżają go do obiektu częściowo inteligentnego (rys. 3). Podstawą tych układów są elementy informatyki i automatyki.



Rys. 3. Niektóre układy specjalne inteligentnego samochodu: ECU – Engine Control Unit – moduł (jednostka) sterowania silnikiem; EFI – Electronic Fuel Injection – elektroniczny wtrysk paliwa, SPI – Single Point Injection – jednopunktowy wtrysk paliwa do kolektora dolotowego, MPI – Multi Point Injection – wielopunktowy wtrysk paliwa do kolektora dolotowego, GDI – Gasoline Direct Injection – bezpośredni wtrysk paliwa do komory spalania (zapłon iskrowy), MIL – Malfunction Indicator Light – sygnalizator (lampka) uszkodzenia układu kontroli czystości spalin, EDC – Electronic Diesel Control – elektroniczny wtrysk paliwa w silnikach o zapłonie samoczynnym, UIS – Unit Injector System – pompowtryskiwacz, UPS – Unit Pump system –

pompa – przewód – wtryskiwacz na każdy cylinder, ABS – Anti Blok System – układ przeciwblokujący hamulców, AS – Active Suspension – zawieszenie aktywne, ASC – Antil Slip Control – układ regulacji przeciwoślizgowej, ATC – Automatic Transmission Control – układ automatycznej zmiany przełożeń, CCS – Cruise Control System – układ regulacji prędkości jazdy, EAS – Electronically Assisted Steering – elektroniczne wspomaganie układu kierowniczego, EC4WD – Electronically Controlled 4 Wheel Drive – elektroniczne sterowanie napędem czterech kół, EC4WS – Electronically Controlled 4 Wheel Steering – elektroniczne sterowanie skrzem czterech kół, ESP – Electronic System Position – układ stabilizacji toru jazdy

2. METODYKA DIAGNOSTYKI INTELIGENTNEGO MECHATRO- NICZNEGO POJAZDU MECHANICZNEGO

Metodyka diagnostyki inteligentnego mechatronicznego pojazdu mechanicznego obejmuje następujące zagadnienia [2]:

- 1) analizę diagnostyczną obiektu;
- 2) dekompozycję obiektu;
- 3) modele diagnostyczne;
- 4) diagnostyczne badania eksperymentalne;
- 5) algorytmy diagnozowania;
- 6) metody i urządzenia diagnostyczne;
- 7) podatność diagnostyczną pojazdu.

Analiza diagnostyczna pojazdu obejmuje następujące zasadnicze elementy:

- przeznaczenie, budowę i zasadę działania;
- warunki funkcjonowania obiektu i jego elementów;
- zakres zmian wartości parametrów stanu i możliwości ich regulacji;
- zbiór parametrów sygnałów diagnostycznych;
- wartości graniczne parametrów stanu i parametrów sygnałów diagnostycznych;

- dane dotyczące niezawodności, w tym trwałości, nieuszkodzalności, naprawialności i przechowywalności;
- wykorzystane metody i urządzenia diagnostyczne.

Wynikiem analizy diagnostycznej pojazdu powinno być dokonanie jego dekompozycji. Do celów praktycznych wystarczy podzielić pojazd na następujące poziomy diagnostyczne:

- a) poziom pierwszy – pojazd;
- b) poziom drugi – układy np.: przetwarzania i przenoszenia energii, nośny, prowadzenia, instalacja elektryczna, elektroniczno-informatyczne, bezpieczeństwa i specjalne;
- c) poziom trzeci – zespoły np.: silnik, skrzynia biegów, sterownik;
- d) poziom czwarty – podzespoły np.: układ korbowo-tłokowy, pamięć RAM, ROM;
- e) poziom piąty – pary kinematyczne np.: łożyska główne, przekładnia kierownicza.

W praktyce ustalenie stanów wszystkich elementów obiektu jest możliwe, lecz nieuzasadnione. Istnieje potrzeba wyboru tych elementów, które mają być diagnozowane, wybierając je za pomocą kryteriów: bezpieczeństwa, słabych ogniw, prawdopodobieństwa uszkodzeń

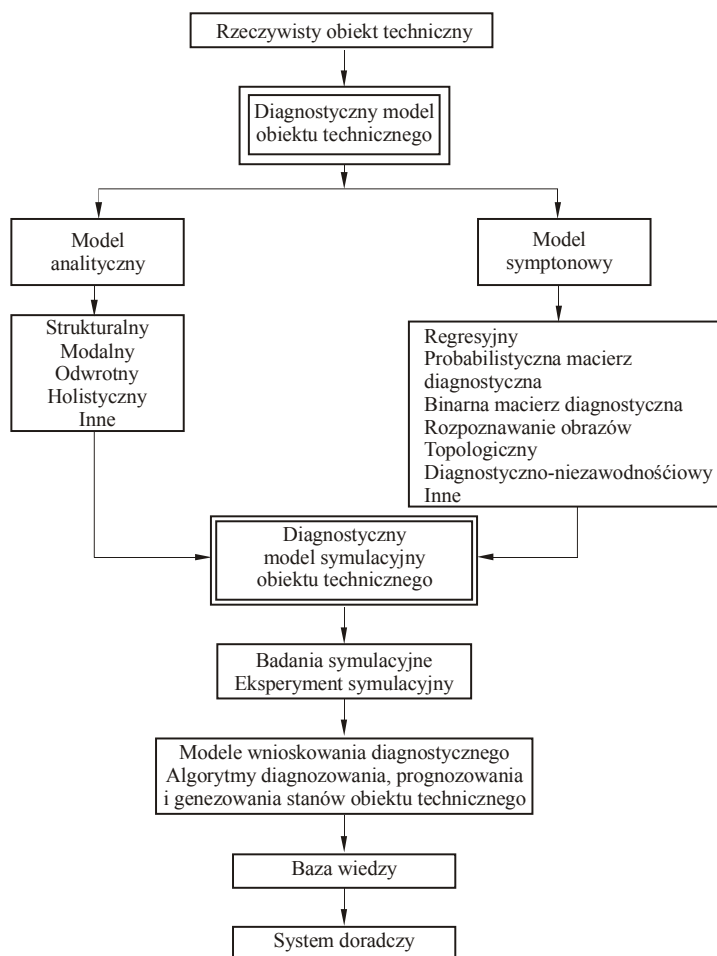
elementów, kosztów uszkodzeń, niezawodnościowo-ekonomicznych.

Następnym etapem postępowania jest budowa modelu diagnostycznego pojazdu mechanicznego. Model diagnostyczny obiektu zawiera podstawowe informacje o jego stanie i jest podstawą: realizacji potrzebnych, a nie przypadkowych badań diagnostycznych, algorytmów diagnozowania, wyboru koniecznych metod i urządzeń diagnostycznych, a także podatności diagnostycznej.

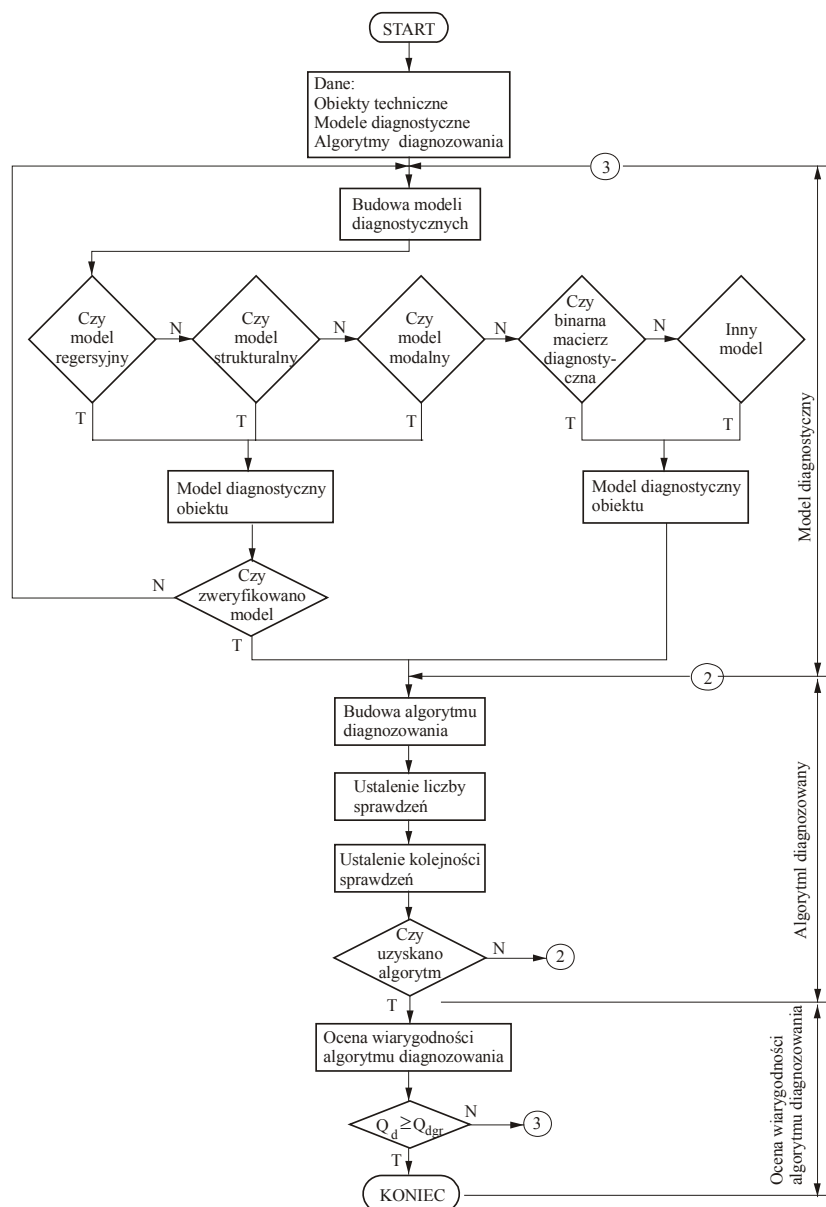
Do wyboru mamy dużą grupę modeli (rys. 4):

- 1) modele symptomowe: regresyjne, probabilistyczną macierz diagnostyczną, binarną macierz diagnostyczną, topologiczne, diagnostyczno-niezawodnościowy;
- 2) modele analityczne: strukturalne, modalny, holistyczny, odwrotny;
- 3) modele symulacyjne.

Ustalenie stanu technicznego ma miejsce w wyniku przeprowadzonych jego badań diagnostycznych. Proces badań diagnostycznych polega na wykonaniu określonego zbioru sprawdzeń i analizy uzyskanych wyników. Uporządkowany nominalny zbiór sprawdzeń diagnostycznych – to algorytm diagnozowania, którego istotę wyboru i weryfikacji przedstawiono na rys. 5.



Rys. 4. Ilustracja graficzna diagnostycznych modeli jako narzędzi badań i oceny stanów pojazdów mechanicznych



Rys. 5. Proces weryfikacji modelu diagnostycznego i algorytmu diagnozowania pojazdu mechanicznego

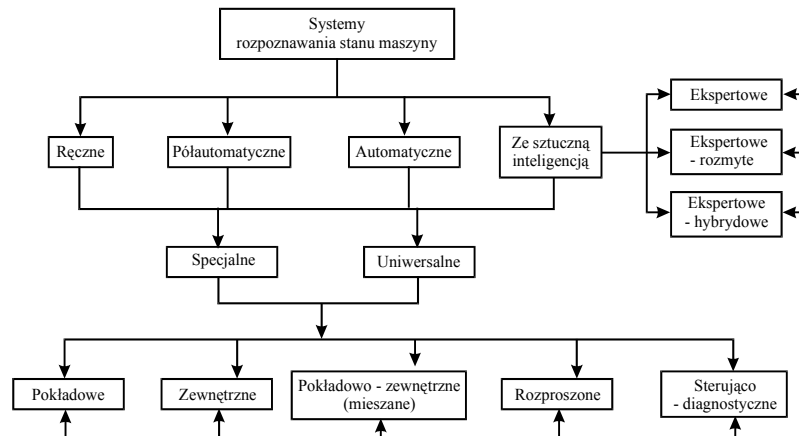
Zbudowany, efektywny algorytm diagnozowania pojazdu mechanicznego wymusza i wskazuje wybór potrzebnych a nie przypadkowych metod i systemów rozpoznawania stanów.

System rozpoznawania stanów pojazdów mechanicznych (rys. 6) jest zbiorem elementów pomiarowych przeznaczonych do: wydobywania, zbierania, gromadzenia, przetwarzania i przedstawiania informacji o stanie obiektu technicznego. Synonimami systemu rozpoznawania stanów pojazdów są: system diagnostyczny, urządzenia diagnostyczne, środek diagnozy, środek diagnozowania.

Mając rozwiązane rozpatrzone problemy dotyczące modeli diagnostycznych, algorytmów, metod i urządzeń diagnostycznych można ustalić podatność eksploatacyjną, w tym diagnostyczną pojazdów mechanicznych, tzn. ich przystosowanie

do realizacji procesów diagnozowania, prognozowania i genezowania.

Na dzień dzisiejszy zastosowanie do diagnostyki pojazdów mechanicznych urządzeń (systemów): pokładowych, zewnętrznych lub mieszanych nie jest wystarczające. Chodzi o to, aby wynik badań oceny stanu, tzn. diagnoza była podstawą celowego oddziaływania na wejścia obiektu tak, aby zapewnić uzyskanie zakładanych w miarę możliwości optymalnych jego wyjść. Inaczej mówiąc chodzi o celowe oddziaływanie na obiekt, czyli o sterowanie obiektem. Zaznaczyć należy, że w literaturze nazwa zintegrowany system informatyczny sterowania pojazdów nie jest stosowana, a najbardziej przybliżone dla niej są pojęcia: system informatyczny, zintegrowany system kierowania i informacji, vectronix system, system sterująco-diagnostyczny.



Rys. 6. Klasyfikacja systemów rozpoznawania stanu pojazdów mechanicznych

Zintegrowane systemy informatyczne sterowania pojazdów są budowane na bazie techniki mikroprocesorowej, w tym: z wykorzystaniem metod sztucznej inteligencji, tzn. systemów ekspertowych, sieci neuronowych i logiki rozmytej.

W chwili obecnej, w pojeździe mechanicznym istnieje wiele podsystemów informatycznych (np. sterowników) nie całkowicie ze sobą zintegrowanych, w aspekcie właściwego rozdziału i wykorzystania informacji do sterowania układami pojazdu. Nie ma jeszcze kompleksowych systemów diagnostycznych opartych na modelach diagnostycznych i optymalnych algorytmach kontroli stanu oraz lokalizacji uszkodzeń. Wydaje się, że te cechy ujemne można zminimalizować stosując zintegrowany informatyczny system sterowania, (rys. 7), przybliżając pojazd do urządzenia inteligentnego [4].

Zintegrowany informatyczny system sterowania pojazdu samochodowego służy do:

- zbierania, gromadzenia, przetwarzania, przedstawiania i przekazywania informacji dotyczących otoczenia i funkcjonowania pojazdu;
- przygotowania syntetycznych danych lub propozycji decyzji, które może podjąć operator (użytkownik, kierowca) pojazdu w aspekcie jego efektywnego wykorzystania;
- przygotowania syntetycznych danych lub/i propozycji decyzji, które może podjąć operator pojazdu w zakresie utrzymania go w stanie gotowości działania, w tym stanie gotowości technicznej;
- integracji i koordynacji funkcjonowania wszystkich elementów funkcjonalnych pojazdu w aspekcie: utrzymania określonego poziomu bezpieczeństwa i wysokiego poziomu usług logistycznych.

Na rys. 7 przedstawiono schemat pojazdu samochodowego jako celowego antropotechnicznego systemu działania S_D , w którym można wyróżnić:

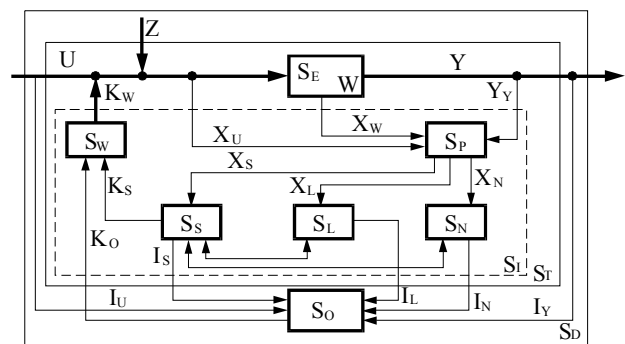
- pojazd jako podsystem techniczny – S_T ;

– operatora (kierowcę, użytkownika) – S_O .

$$S_D = \langle S_T, S_O, R_{TO} \rangle \quad (1)$$

gdzie:

R_{TO} – relacje.



Rys. 7. Ilustracja graficzna pojazdu mechanicznego z wyodrębnionym zintegrowanym podsystemem informatycznym sterowania S_T : S_D – pojazd jako celowy system działania, S_T – podsystem pomiarowy, S_S – informatyczny podsystem sterowników, S_L – informatyczny podsystem logistyki, eksploatacji i diagnostyki, S_N – informatyczny podsystem nawigacji, S_W – podsystem wykonawczy, S_E – podsystem elementów funkcjonalnych pojazdu, U, Y, W, Z – zbiór wejść, wyjść, stanów i zakłóceń, X_U, X_W, X_Y – zbiory parametrów mierzonych sygnałów, X_S, X_L, X_N – zbiory wartości zmierzonych parametrów sygnałów, K_O – zbiór sygnałów sterujących operatorem, K_S – zbiór sygnałów sterujących ze sterowników układu pojazdów, K_W – zbiór wielkości sterujących, I_S, I_L, I_N, I_U, I_Y – informacje dla operatora

Pojazd jako podsystem techniczny opisuje wyrażenie:

$$S_T = \langle S_E, S_L, R_{EL} \rangle \quad (2)$$

gdzie:

$S_E = \{s_n\}; n=1, \overline{N}$ – zbiór elementów funkcjonalnych pojazdu (np. silnik, układ napędowy, układ kierowniczy itp.);

S_I – zintegrowany informatyczny podsystem sterowania;

R_{EI} – relacje.

Zintegrowany informatyczny podsystem sterowania pojazdem umieszczony w sprzężeniu zwrotnym zdefiniowano następująco:

$$S_I = \langle S_P, S_S, S_L, S_N, S_W, R_{PW} \rangle \quad (3)$$

gdzie:

S_P – podsystem pomiarowy;

S_S – informatyczny podsystem sterowników;

S_L – informatyczny podsystem logistyki, eksploatacji i diagnostyki;

S_N – informatyczny podsystem nawigacji;

S_W – podsystem nastawczy (np. hamulców, układu kierowniczego);

R_{PW} – relacje.

Zasadniczym zadaniem podsystemu pomiarowego S_P jest dostarczenie informacji o aktualnych wartościach zbioru X_U, X_W, X_Y parametrów wejść U , stanów W i wyjść Y , charakteryzujących proces funkcjonowania pojazdu. Podsystem pomiarowy obejmuje zbiór przetworników mechatronicznych i układów dopasowujących. Przekazuje on zbiory X_S, X_L i X_W wartości zmierzonych parametrów do podsystemów: S_S, S_L i S_N .

Podsystem S_S sterowników obejmuje zbiór sterowników różnych układów samochodu, na przykład silnika, AB i innych. Są to specjalizowane systemy informatyczne. Każdy ze sterowników ma za zadanie wypracowanie zbioru K_S sygnałów sterujących, oddziałujących na odpowiedni podsystem wykonawczy S_W samochodu.

Informatyczny podsystem S_L logistyki, eksploatacji i diagnostyki realizuje następujące zadania:

- ustala stan pojazdu w chwili t , $t+\Delta t$ oraz $t-\Delta t$;
- lokalizuje niezdatne elementy obiektu w przypadku jego niezdatności;
- przewiduje termin następnego diagnozowania i obsługiwanego samochodu;
- określa zapasy materiałów pędnych i smarów oraz płynów eksploatacyjnych;
- przewiduje możliwości wykonania zadań operacyjnych przez system.

Podkreślić należy, że zasadniczym elementem systemu S_L jest podsystem diagnostyczny.

Zadaniem informatycznego podsystemu S_N nawigacji jest:

- określenie położenia pojazdu w terenie;
- określanie warunków atmosferycznych;
- przekazywanie decyzji i meldunków w układzie: przełożony – operator – przełożony;
- wyznaczenie optymalnej trasy ruchu pojazdu;
- podwyższenie bezpieczeństwa operatora pojazdu i ładunku.

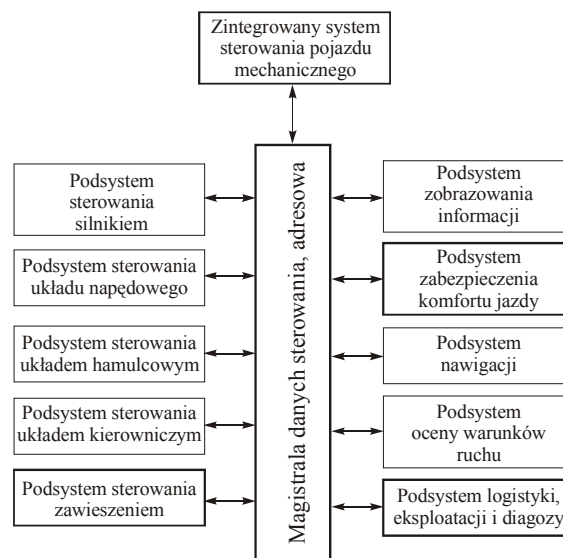
System S_W nastawczy obejmuje podsystemy bezpośrednio oddziałujące na elementy funkcjonalne samochodu, za pomocą zbioru wielkości K_W sterujących. Zasadniczymi elementami systemu S_W nastawczego są podsystemy: silnika,

układów: napędowego, hamulcowego, kierowniczego, zawieszenia, bezpieczeństwa czynnego i biernego.

W ogólnym ujęciu, istotę funkcjonowania zintegrowanego informatycznego systemu sterowania pojazdem można przedstawić następująco.

Podsystem S_P pomiarowy odbiera sygnały z wejść U , stanów W i wyjść Y . Rejestruje zbiory X_S, X_L i X_N wartości zmierzonych parametrów, które z kolei są przekazywane do podsystemów S_S, S_L i S_N . W tych podsystemach wyniki pomiarów wartości parametrów są transformowane w diagnozy chwilowe czyli decyzje, które w postaci zbioru sygnałów K_S sterujących bezpośrednio oddziałują na podsystem S_W wykonawczy, zatem na elementy funkcjonalne pojazdu, z pominięciem operatora S_O . Tak funkcjonujący pojazd mechaniczny można określić jako: system mechatroniczny automatycznej regulacji, system automatyczny lub robot.

W rzeczywistości podsystem S_S, S_L i S_N przekazuje informacje: I_S, I_L, I_N – operatorowi S_O . Na bazie tych informacji i informacji o stanie I_U wejść i stanie I_Y wyjść, operator za pomocą sygnałów K_O sterujących może również oddziaływać na podsystem S_W wykonawczy pojazdu. Należy przypuszczać, że nastąpi pełna integracja budowy i funkcjonowania pojazdu w aspekcie sterowania (rys. 8), zatem sztucznej inteligencji.



Rys. 8. Przewidywana koncepcja rozwoju sterowania w aspekcie inteligentnego pojazdu mechanicznego

PODSUMOWANIE

Reasumując zagadnienia dotyczące inteligentnego pojazdu mechanicznego jako obiektu diagnozowania można stwierdzić, co następuje:

- 1) można przyjąć, że inteligentny pojazd mechaniczny to taki twór fizyczny, który ma

- przynajmniej niektóre z cech inteligentnego człowieka;
- 2) uzyskanie określonych właściwości inteligencji przez pojazd mechaniczny jest możliwe jeżeli się potraktuje ten obiekt jako system mechatroniczny automatycznej regulacji, mający trzy elementy podstawowe: pomiarowy, sterujący i wykonawczy.
 - 3) podstawą wszelkich racjonalnych działań diagnostycznych jest model diagnostyczny i algorytm diagnozowania pojazdu mechanicznego;
 - 4) podsystem diagnostyczny jest elementem zintegrowanego systemu informatycznego sterowania w zakresie ustalenia diagnozy, prognozy i genezy pojazdu;
 - 5) w chwili obecnej samochodu wyposaża się w różnego rodzaju specjalistyczne systemy informatyczne, które nie są całkowicie zintegrowane w zakresie gromadzenia, przetwarzania, przedstawiania przekazywanych informacji;
 - 6) zintegrowany informatyczny system sterowania scala i koordynuje pracę wszystkich elementów funkcjonalnych pojazdu;
 - 7) na podstawie zebranych informacji o stanie pojazdu i stanie otoczenia zintegrowany informatyczny system sterowania wypracowuje i przygotowuje syntetyczne dane potrzebne do podjęcia decyzji;
 - 8) zintegrowany informatyczny system sterowania jest podstawowym narzędziem operatora (załogi) utrzymania pojazdu w stanie zdolności funkcjonalnej i stanie zdolności zadaniowej;
 - 9) zintegrowany informatyczny system sterowania jest podstawowym narzędziem operatora (załogi) w zakresie efektywnego wykorzystania pojazdu do zaspokojenia potrzeb człowieka;
 - 10) tak ujęte: zadania, budowa i organizacja zintegrowanego informatycznego systemu sterowania są nową jakością zarówno w sensie poznawczym jak i utylitarnym w stosunku do obecnie istniejących systemów informatycznych pojazdów mechanicznych;
 - 11) racjonalnie i kompleksowo rozwiązany informatyczny system sterowania jest podstawowym elementem i narzędziem inteligentnego pojazdu mechanicznego.

5. Nowa Encyklopedia Powszechna. PWN, Warszawa 2005.
6. Szczepaniak C.: *Podstawy modelowania systemu człowiek – pojazd – otoczenie*. PWN, Warszawa 1999.
7. Świątnicki W., Świątnicki Z.: *Bronie inteligentne*. Bellona, Warszawa 1992.



Prof. dr hab. inż. **Stanisław NIZIŃSKI** jest pracownikiem naukowym Katedry Budowy Eksploatacji Pojazdów i Maszyn, UWM w Olsztynie oraz WITPS w Sulejówku. Jest wieloletnim członkiem Sekcji Podstaw Eksploatacji KBM PAN oraz PTDT. Jego

zainteresowania naukowe obejmują zagadnienia dotyczące logistyki w systemach działania, eksploatacji i diagnostyki obiektów technicznych, szczególnie pojazdów mechanicznych. Jest autorem i współautorem wielu prac naukowych i dydaktycznych.



Dr inż. **Sławomir WIERZBICKI** pracuje obecnie na stanowisku adiunkta w Katedrze Budowy, Eksploatacji Pojazdów i Maszyn, UWM w Olsztynie.

W pracy zajmuje się zagadnieniami eksploatacji i diagnostyki pojazdów i maszyn. Jest członkiem PTDT, od 2004 roku pełni funkcję sekretarza czasopisma „Diagnostyka”.

LITERATURA

1. *Encyklopedia Britannica*. Kurpicz, Poznań 2000.
2. Niziński S., Michalski R.: *Diagnostyka obiektów technicznych*. ITE, Radom 2002.
3. Niziński S., Wierzbicki S.: *Inteligentny pojazd mechaniczny a diagnostyka*. V Sympozjum nt. „Diagnostyka systemów i urządzeń technicznych”, WAT, PTDT, Ustronie 2003.
4. Niziński S., Wierzbicki S.: *Zintegrowany system informatyczny sterowania pojazdów*. Diagnostyka, nr 30 t.2/2004. s. 47–52.

DIAGNOZOWANIE Z DWUPOZIOMĄ KOMPARACJĄ NIEPEWNYCH SYMPTOMÓW I SYNDROMU STANU OBIEKTU

Lesław BĘDKOWSKI

Instytut Systemów Elektronicznych, Wydział Elektroniki
Wojskowa Akademia Techniczna, ul. S. Kaliskiego 2, 00-908 Warszawa,
fax (22) 6839125, e-mail: lbedkowski@wel.wat.edu.pl

Streszczenie

W artykule rozważa się zagadnienie wiarygodności diagnoz formułowanych w oparciu o niepewne (np. z powodu zakłóceń procesu diagnozowania) symptomy i syndromy stanu obiektu. Proponuje się metodę diagnozowania umożliwiającą otrzymywanie dostatecznie wiarygodnych diagnoz poprzez powtarzanie testowania oraz komparację uzyskanych symptomów i syndromów. Opisano 6 reguł wyboru diagnozy oraz omówiono ograniczenia ich zastosowań i wskaźniki ocenowe.

Słowa kluczowe: diagnozowanie, wiarygodność diagnozy, niepewność symptomu i syndromu.

DIAGNOSING WITH DOUBLE-LEVEL COMPARISON OF UNCERTAIN SYMPTOMS AND SYNDROME OF AN OBJECT STATE

Summary

A question of authenticity of diagnosis formulated on the basis of uncertain symptoms and syndromes of an object state is discussed. A diagnosing method that allows receiving the sufficiently authentic diagnoses through repetition of testing and comparison of obtained symptoms and syndromes is proposed. The 6 rules of diagnostic inference are described. Additionally application limits and assessment indicators are discussed.

Key words: diagnosing, authenticity of diagnosis, uncertainty of symptom and syndrome.

1. WPROWADZENIE

W [1] opisano przypadek diagnozowania z jednopoziomą, dwustopniową komparacją syndromową. Komparacja ta składa się:

– z **komparacji segregującej** otrzymane syndromy na podzbiory o jednakowej postaci z jednoczesnym określeniem ich liczności;

– z **komparacji wartościującej** polegającej na wyznaczeniu wartości syndromu¹ i wyborze syndromu o największej wartości.

Metoda taka jest przydatna przede wszystkim wtedy, gdy (z różnych powodów) wskazane jest wnioskowanie diagnostyczne na poziomie **kompletnych** syndromów.

W niniejszym artykule przedstawiona jest jedna z najprostszych metod diagnozowania z komparacją dwupoziomą – to jest na poziomie komparacji symptomów i syndromów. Metoda taka jest wskazana wtedy, gdy liczba możliwych stanów obiektu jest duża i gdy występują trudności ze ścisłym zdefiniowaniem a priori wszystkich

możliwych stanów obiektu. Diagnozowanie dwupoziomowe nie wymaga takiego definiowania. Ponadto metodę tę można stosować wtedy, gdy występują zróżnicowane wymagania dotyczące granicznych wartości symptomów².

Algorytm diagnozowania stanu obiektu technicznego w oparciu o komparację dwupoziomą (w przypadku niepewnych symptomów) przedstawia rys.1.

2. KOMPARACJA NA POZIOMIE SYMPTOMU

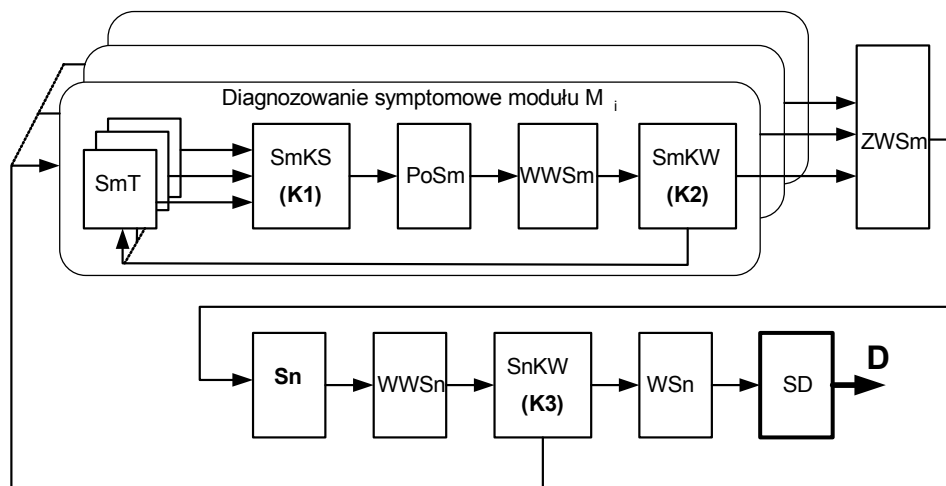
Planowanie procedury diagnozowania należy rozpocząć od syntezy struktury diagnostycznej obiektu³.

Rozpatrzmy najprostszy przypadek, dla którego przyjęto:

¹ Wartością syndromu nazywamy wartość prawdopodobieństwa prawdziwości syndromu.

² Wartością symptomu nazywamy wartość prawdopodobieństwa prawdziwości symptomu.

³ Strukturę diagnostyczną interpretuje się tu jako zbiór modułów diagnostycznych; modulem diagnostycznym nazywany jest fragment obiektu, o definiowalnym stanie.



Rys. 1. Algorytm komparacyjnej procedury diagnostycznej

Oznaczenia: SmT – symptomowe testowanie modułu; SmKS – symptomowa komparacja segregująca; PoSm – podzbiory symptomów; WWSm – wyznaczanie wartości symptomów; SmKW – symptomowa komparacja wartościująca; ZWSm – zbiór wybranych symptomów; Sn – syndrom stanu; WWSn – wyznaczanie wartości syndromu; SnKW – syndromowa komparacja wartościująca; WSn – wybrany syndrom; SD – synteza diagnozy; D – diagnoza

– wszystkie moduły diagnostyczne są rozłączne (nie mają wspólnych elementów), a zarazem stanowią tory symptomów⁴; na rys. 2 pokazano przykładową strukturę diagnostyczną o modułach rozłącznych;

– każdy moduł jest podobieństwem jednosymptomowym tzn., że jeden symptom wystarcza do zdefiniowania stanu modułu;

– każdy symptom jest jednowymiarowy; tzn., że tworzony jest na podstawie wyniku jednego pomiaru (lub innej pojedynczej informacji diagnostycznej);

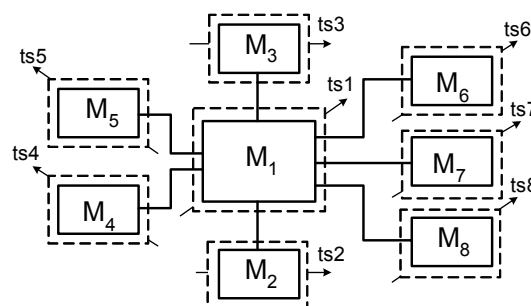
– każdemu symptomowi przyporządkowuje się symboliczną wartość logicznego zera (wynik negatywny), oznaczającego niezdatność modułu lub symboliczną wartość logicznej jedynki (wynik pozytywny), oznaczającej zdatność modułu; zatem wynikiem każdej realizacji testu jest liczba „0” lub „1”.

– **symptomowa komparacja segregująca (SmKS)** polega na komparacji i segregacji symptomów otrzymanych w wyniku testowania modułu na podzbiory symptomów pozytywnych i negatywnych oraz na określeniu liczności tych podzbiorów; operację taką przeprowadza się dla każdego modułu diagnostycznego;

– **symptomowa komparacja wartościująca (SmKW)** polega na określaniu wartości symptomów negatywnych i pozytywnych oraz na porównywaniu otrzymanych wartości z indywidualnymi wymaganiami (sformułowanymi wcześniej);

– jeżeli wartości wszystkich symptomów spełniają indywidualne wymagania, należy

utworzyć syndrom, wyznaczyć jego wartość i przeprowadzić **syndromową komparację wartościującą (SnKW)** ze względu na spełnienie wymagań syndromowych.



Rys. 2. Przykład obiektu o rozłącznych modułach diagnostycznych

Oznaczenia: M₁, ..., M₈ – moduły diagnostyczne; ts₁, ..., ts₈ – tory symptomów

3. MODEL WYZNACZANIA WARTOŚCI SYMPTOMU

Prawdopodobieństwo zdarzenia polegającego na L_i^0 -krotnym wystąpieniu **prawdziwego** symptomu negatywnego s_i^0 pod warunkiem występowania (w czasie diagnozowania) stanu niezdatności $\mathcal{E}_i^0$ modułu M_i – wyznacza wyrażenie:

$$p_i(0|0) = [P(s_i^0 | \mathcal{E}_i^0)]^{L_i^0} = [R_{zi}(0|0)]^{L_i^0}$$

(1)

gdzie: – i – numer modułu diagnostycznego;

⁴ Torem symptomu nazywamy taki podzbiór modułów obiektu, od których stanu zależy postać symptomu.

– L_i^0 – liczność podzbioru symptomów negatywnych;
 – $R_{zi}(0|0)$ – prawdopodobieństwo jednokrotnego wystąpienia prawdziwego symptomu negatywnego.

Prawdopodobieństwo zdarzenia polegającego na L_i^0 -krotnym wystąpieniu **falszywego** symptomu negatywnego $\bar{S}_i^0$ pod warunkiem występowania stanu zdadności $\mathcal{E}_i^1$ modułu M_i – wyznacza wyrażenie:

$$p_i(0|1) = \left[P(\bar{S}_i^0 | \mathcal{E}_i^1) \right]^{L_i^0} = [R_{zi}(0|1)]^{L_i^0} \quad (2)$$

gdzie: – $R_{zi}(0|1)$ – prawdopodobieństwo jednokrotnego wystąpienia falszywego symptomu negatywnego.

Prawdopodobieństwo zdarzenia polegającego na L_i^1 -krotnym wystąpieniu **prawdziwego** symptomu pozytywnego S_i^1 pod warunkiem występowania stanu zdadności $\mathcal{E}_i^1$ modułu M_i – wyznacza wyrażenie:

$$p_i(1|1) = \left[P(S_i^1 | \mathcal{E}_i^1) \right]^{L_i^1} = [R_{zi}(1|1)]^{L_i^1} \quad (3)$$

gdzie: – $R_{zi}(1|1)$ – prawdopodobieństwo jednokrotnego wystąpienia prawdziwego symptomu pozytywnego.

Prawdopodobieństwo zdarzenia polegającego na L_i^1 -krotnym wystąpieniu **falszywego** symptomu pozytywnego $\bar{S}_i^1$ pod warunkiem występowania stanu niezdatności $\mathcal{E}_i^0$ modułu M_i – wyznacza wyrażenie:

$$p_i(1|0) = \left[P(\bar{S}_i^1 | \mathcal{E}_i^0) \right]^{L_i^1} = [R_{zi}(1|0)]^{L_i^1} \quad (4)$$

gdzie: – $R_{zi}(1|0)$ – prawdopodobieństwo jednokrotnego wystąpienia falszywego symptomu pozytywnego.

Liczba testowań L_{Ti} modułu M_i jest równa sumie liczności podzbioru symptomów negatywnych $L_i(S_i^0)$ i liczności podzbioru symptomów pozytywnych $L_i(S_i^1)$:

$$L_{Ti} = L_i(S_i^0) + L_i(S_i^1) \quad (5)$$

W czasie wykonywania sesji testowej modułu, znajduje się on (dla przyjętych założeń) w jednym z dwu stanów: w stanie niezdatności albo w stanie zdadności. Ponadto przyjęto tu, że **stan modułu nie zmienia się w czasie trwania sesji**.

Zatem może zachodzić jedno z dwu zdarzeń:

Zdarzenie Z_i^0 . W czasie sesji testowej modułu M_i jest w stanie **niezdatności** $\mathcal{E}_i^0$ oraz wystąpił L_i^0 -krotnie prawdziwy symptom negatywny $\bar{S}_i^0$ i wystąpił L_i^1 -krotnie falszywy symptom pozytywny $\bar{S}_i^1$ stanowi to iloczyn trzech zdarzeń.

Zatem prawdopodobieństwo zdarzenia Z_i^0 czyli prawdopodobieństwo a’posteriori niezdatności modułu M_i jest iloczynem trzech prawdopodobieństw:

$$P(Z_i^0) \equiv P(\mathcal{E}_i^0) = (1 - R_{ei}) p_i(0|0) p_i(1|0) \quad (6)$$

gdzie: – R_{ei} – prawdopodobieństwo a’priori stanu zdadności modułu M_i .

Zdarzenie Z_i^1 . W czasie sesji testowej modułu M_i jest w stanie **zdadności** $\mathcal{E}_i^1$ oraz wystąpił L_i^0 -krotnie falszywy symptom negatywny $\bar{S}_i^0$ i wystąpił L_i^1 -krotnie prawdziwy symptom pozytywny S_i^1 ; stanowi to iloczyn trzech zdarzeń.

Zatem prawdopodobieństwo zdarzenia Z_i^1 czyli prawdopodobieństwo a’posteriori zdadności modułu M_i jest iloczynem trzech prawdopodobieństw:

$$P(Z_i^1) \equiv P(\mathcal{E}_i^1) = R_{ei} p_i(0|1) p_i(1|1) \quad (7)$$

Stan zdadności i stan niezdatności stanowią zupełną sumę zdarzeń. Zatem:

– **warunkowe** prawdopodobieństwo zdarzenia polegającego na tym, że moduł M_i jest w stanie niezdatności przedstawia wyrażenie:

$$P_{wi}(\mathcal{E}_i^0) = \frac{P(\mathcal{E}_i^0)}{P(\mathcal{E}_i^0) + P(\mathcal{E}_i^1)} \quad (8)$$

Wyrażenie (8) stanowi warunkową **wartość symptomu negatywnego** $\bar{S}_i^0$, wynikającą z testowania symptomowego modułu M_i ;

– analogicznie, **warunkowe** prawdopodobieństwo zdarzenia polegającego na tym, że moduł M_i jest w stanie zdadności przedstawia wyrażenie:

$$P_{wi}(\mathcal{E}_i^1) = \frac{P(\mathcal{E}_i^1)}{P(\mathcal{E}_i^0) + P(\mathcal{E}_i^1)} \quad (9)$$

Wyrażenie (9) stanowi warunkową **wartość symptomu pozytywnego** S_i^1 , wynikającą z testowania symptomowego modułu M_i .

Wartości symptomów (8) i (9) stanowią podstawę diagnoz symptomowych – a w konsekwencji także diagnozy syndromowej.

4. WYBÓR SYNDROMU

W niektórych przypadkach wystarczające są diagnozy symptomowe dla **wybranych** modułów diagnostycznych. Najczęściej wymagane są diagnozy symptomowe dla **wszystkich** modułów oraz diagnoza syndromowa.

Syndromem stanu obiektu jest zbiór wybranych w postępowaniu komparacyjnym symptomów:

$$S = [s_1, s_2, \dots, s_N] \quad (10)$$

Wartość syndromu jest iloczynem wartości symptomów:

$$P_S = \prod_{i=1}^N P_{wi}^* \quad (11)$$

gdzie: P_{wi}^* – **warunkowa** wartość **wybranego** symptomu.

Syndromowa komparacja wartościująca polega na zbadaniu czy otrzymany syndrom spełnia wymagania, tzn. na przykład czy wartość syndromu jest niemniejsza od wartości wymaganej (jeśli takie wymaganie istnieje).

Do sformułowania ostatecznej diagnozy można zastosować jedną z reguł wyboru diagnozy.

5. REGUŁY WYBORU DIAGNOZY

5.1. SYMPTOMOWA REGUŁA WIĘKSZOŚCIOWA (SmW)

Reguła ta polega na realizowaniu testów symptomowych każdego modułu. W tym przypadku należy przyjąć pewną minimalną, nieparzystą liczbę testów np. $L_{i \min} = 3$. Po zrealizowaniu tej liczby testów wyznacza się wartości symptomów negatywnych i pozytywnych oraz sprawdza się spełnienie warunku (12):

$$L_{i \min} : P_{wi}(\mathcal{E}_i^0) \neq P_{wi}(\mathcal{E}_i^1); \quad i = 1, 2, \dots, N \quad (12)$$

Jeśli warunek (12) nie jest spełniony to realizuje się kolejny test, aż do uzyskania takiej liczby testów L_{Ti}^* , dla której warunek ten jest spełniony.

Dla tej liczby testów i dla każdego modułu znajduje się większą wartość symptomu:

$$P_{wi}^* = \max[P_{wi}(\mathcal{E}_i^0), P_{wi}(\mathcal{E}_i^1)]; \quad i = 1, 2, \dots, N \quad (13)$$

Na tej podstawie można wskazać domniemane stany modułów zgodnie z relacjami (14).

$$\left\{ \begin{aligned} [P_{wi}^* = P_{wi}(\mathcal{E}_i^0)] &\Rightarrow s_i^{0*} \Rightarrow \mathcal{E}_i^0 \\ [P_{wi}^* = P_{wi}(\mathcal{E}_i^1)] &\Rightarrow s_i^{1*} \Rightarrow \mathcal{E}_i^1 \end{aligned} \right\}; \quad i = 1, 2, \dots, N \quad (14)$$

Z wyznaczonych w ten sposób symptomów i ich wartości oraz domniemanych stanów modułów można utworzyć zapis tabelaryczny (tab.1).

Tabela 1

Tabelaryczny zapis symptomów i ich wartości oraz domniemanych stanów

Moduł o numerze „i”	M ₁	M ₂		M _N
Wartość wybranego symptomu	P _{w1} [*]	P _{w2} [*]		P _{wN} [*]
Wybrany symptom	s ₁	s ₂		s _N
Stan modułu	ℰ ₁	ℰ ₂		ℰ _N

5.2. SYMPTOMOWA REGUŁA PROGOWA (SmP)

Reguła ta polega na realizowaniu testów symptomowych każdego modułu. Po zrealizowaniu kolejnego testu oblicza się wartości symptomów negatywnych i pozytywnych oraz sprawdza się spełnienie pierwszego warunku (12) (jak w regule SmW) i wyznacza wartość większą:

$$P_{wi}^* = \max[P_{wi}(\mathcal{E}_i^0), P_{wi}(\mathcal{E}_i^1)]; \quad i = 1, 2, \dots, N \quad (15)$$

Następnie sprawdza się drugi warunek:

$$P_{wi}^* \geq P_{igr}; \quad i = 1, 2, \dots, N \quad (16)$$

gdzie: P_{igr} – graniczna wartość symptomu dla i-tego modułu.

Jeśli warunek (16) jest spełniony to testowanie modułu można uznać za zakończone i zarejestrować liczbę wykonanych testów.

Dalsze postępowanie jest analogicznie jak dla reguły SmW.

Reguła SmP wymaga zazwyczaj większej długości sesji lecz wartości diagnoz są znacznie wyższe (liczba błędnych diagnoz jest mniejsza).

5.3. SYMPTOMOWA REGUŁA WIĘKSZOŚCIOWO-PROGOWA (SmWP)

Jest to reguła mieszana.

Praktyka eksploatacyjna wskazuje na to, że w niektórych przypadkach należy różnicować wymagania dotyczące wartości symptomów. Można zatem podzielić zbiór modułów obiektu na podzbiory modułów o różnych wymaganiach. Np. jeden podzbiór może zawierać moduły o szczególnym znaczeniu ze względów ekonomicznych, bezpieczeństwa itp. Do diagnozowania modułów tego podzbioru należy stosować regułę SmP. Do drugiego podzbioru wystarczy zastosować regułę SmW. Otrzymuje się w ten sposób regułę o pośrednich właściwościach.

5.4. SYMPTOMOWA REGUŁA WIELOPROGOWA (SmWiP)

W praktyce eksploatacyjnej występują sytuacje, w których konieczne jest różnicowanie progów symptomowych dla różnych modułów

i różnych symptomów (w znaczeniu: negatywnych i pozytywnych) w zależności od skutków otrzymywanych na tej podstawie diagnoz (wyższe wartości progów dla ważnych modułów i stanów). Sprzyja to zmniejszaniu liczby błędów szczególnie ważnych diagnoz, przy równoczesnym skracaniu długości sesji diagnostycznych.

5.5. SYNDROMOWA REGUŁA PROGOWA (SnP)

W wielu przypadkach w procesach eksploatacyjnych ważna jest wartość syndromu stanu obiektu bez szczegółowej analizy wartości symptomów. W takich przypadkach przyjmuje się taką samą liczbę testów dla wszystkich modułów. Po zrealizowaniu kolejnych testów wyznacza się wartość syndromu. Następnie sprawdza się spełnienie syndromowego warunku progowego:

$$P_S(L) \geq P_{S_{gr}} \quad (17)$$

gdzie: $P_S(L)$ – wartość syndromu dla zbioru testów o liczności L ;

$P_{S_{gr}}$ – graniczna wartość syndromu.

Jeśli warunek (17) jest spełniony, to sesję diagnostyczną można zakończyć. Na podstawie zbioru wartości symptomów tworzy się diagnozę zupełną, zapisując ją w postaci tabeli analogicznej jak tab.1. Zauważmy, że zgodnie ze wzorem (11) wartość żadnego symptomu nie może być mniejsza od wartości syndromu. Natomiast wartości symptomów mogą być znacznie zróżnicowane w przedziale:

$$P_{S_{gr}} \geq P_{wi}^* \geq 1 \quad (18)$$

5.6. SYMPTOMOWO-SYNDROMOWA REGUŁA WIELOPROGOWA (SmnWiP)

Zastosowanie reguły SnP może w niektórych przypadkach prowadzić do błędnych diagnoz symptomowych, co może być istotne, szczególnie w odniesieniu do ważnych modułów. W takich przypadkach można wprowadzić zbiór progów symptomowych i wymagać spełnienia nie tylko syndromowego warunku progowego lecz również spełnienia symptomowych warunków progowych (19):

$$\left. \begin{array}{l} P_S(L) \geq P_{S_{gr}} \\ \bigwedge_{i=1,2,\dots,N} (P_{wi}^* \geq P_{igr}) \end{array} \right\} \quad (19)$$

6. WSKAŹNIKI PORÓWNAWCZE REGUŁ WYBORU DIAGNOZ

Porównanie reguł można przeprowadzać w oparciu o następujące wskaźniki:

1) suma liczb niezbędnych testów symptomowych w jednej sesji diagnostycznej:

$$L_{TS}^* = \sum_{i=1}^N L_{Ti}^* \quad (20)$$

gdzie: L_{Ti}^* – liczba niezbędnych testów symptomowych dla i -tego modułu; długość sesji symptomowej, charakteryzuje niezbędną długość jednej sesji syndromowej;

2) w przypadku zmiennej liczby testów (ustalanej w każdej sesji w zależności od spełnienia wymagań symptomowych i syndromowych) statystycznym wskaźnikiem charakteryzującym regułę wyboru może być wartość oczekiwana długości sesji $E[L_T^*]$;

3) w przypadku stałej liczby testów w każdej sesji – statystycznym wskaźnikiem charakteryzującym regułę wyboru może być wartość oczekiwana względnej liczby diagnoz spełniających wymagania symptomowe i syndromowe (czyli diagnoz przyjmowanych przez diagnostę) $E[\eta_D]$; estymatorem tego wskaźnika może być stosunek liczby diagnoz przyjętych L_{DP} do liczby zrealizowanych sesji diagnostycznych L_{SD} ; stosunek tych liczb charakteryzuje również **sprawność diagnozowania**:

$$\eta_D = \frac{L_{DP}}{L_{SD}} \quad (21)$$

4) istotnym wskaźnikiem charakteryzującym regułę wyboru diagnoz ze względu na **niezawodność przyjętych diagnoz** może być – wyznaczony na podstawie wielu sesji diagnostycznych – stosunek liczby diagnoz bezbłędnych L_{DC} (tzn. rzeczywiście prawdziwych) do liczby diagnoz przyjętych:

$$R_D = \frac{L_{DC}}{L_{DP}} \quad (22)$$

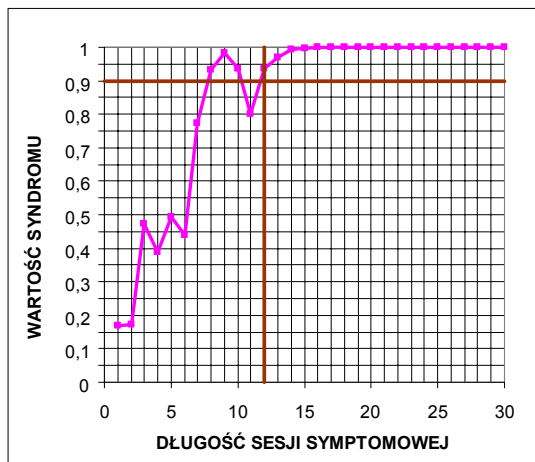
7. PRZYKŁADY ILUSTRACYJNE

1) Na rys.3 – dla reguły SmP – pokazano przykład przebiegu wartości syndromu w zależności od liczby zrealizowanych testów w jednej sesji diagnostycznej.

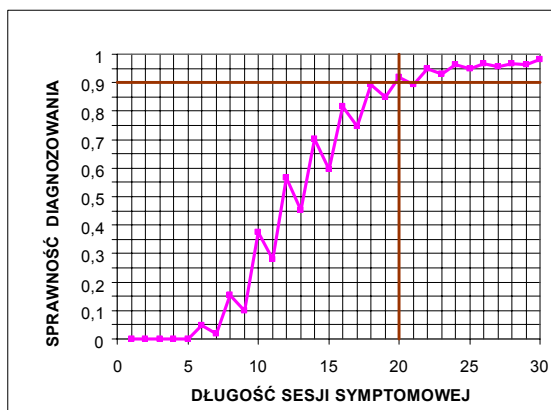
Jak wynika z wykresu – w rozpatrywanej sesji (dla wartości progowej syndromu wynoszącej 0,9) diagnosta przyjmie diagnozę po zrealizowaniu 12 testów.

2) Na rys.4 – dla reguły SmP – pokazano przykład przebiegu sprawności diagnozowania jako funkcji liczby zrealizowanych testów dla każdego modułu na podstawie obserwacji 200 symulacyjnych sesji syndromowych.

Jak wynika z wykresu – przy wartości progowej syndromu równej 0,9 – względną liczbę przyjętych diagnoz na poziomie ok. 0,9 uzyskuje się dla liczby 20 testów.



Rys. 3. Przykładowy przebieg wartości syndromu jako funkcja długości sesji symptomowych (dla jednej sesji syndromowej) przy zastosowaniu symptomowej reguły progowej (SmP)



Rys. 4. Przykładowy przebieg sprawności diagnozowania jako funkcja długości sesji symptomowych przy zastosowaniu symptomowej reguły progowej (SmP)

8. WNIOSKI

Opisaną metodę „diagnozowania z komparacją dwupoziomą” poddano sprawdzeniu w oparciu o modele symulacyjne [2].

Wynika z tego, że:

- metoda z odpowiednio dobraną regułą nadaje się do diagnozowania obiektów podlegających silnym zakłóceniom i przy wykorzystaniu układów diagnozujących też podlegających zakłóceniom – pod warunkiem spełnienia następujących nierówności:

$$\left. \begin{aligned} R_{zi}(1|1) &> R_{zi}(0|1) \\ R_{zi}(0|0) &> R_{zi}(1|0) \end{aligned} \right\}; \quad i = 1, 2, \dots, N \quad (23)$$

przy tym **nie jest konieczna** dokładna znajomość wartości tych prawdopodobieństw;

- metoda może być stosowana nawet jeśli nie są znane prawdopodobieństwa a’priori stanów modułów;
- ze względu na konieczność wykonywania w trakcie diagnozowania znacznej liczby obliczeń niezbędne jest wykorzystanie odpowiednich programów komputerowych;
- możliwe jest wykorzystanie komputera stanowiącego jeden z modułów obiektu lub komputera stanowiącego część układu diagnozującego.

LITERATURA

- [1] Będkowski L., Dąbrowski T.: *Diagnostowanie na podstawie niepewnych syndromów stanu obiektu*. Diagnostyka, PTDT, Vol. 37, 2006, ss. 6.
- [2] Dąbrowski T.: *Badanie symulacyjne skuteczności diagnozowania komparacyjnego na przykładzie systemu alarmowego*. Materiały Konferencji DIAG’2006, ss. 6.
- [3] Grabski F., Jaźwiński J.: *Metody bayesowskie w niezawodności i diagnostyce*. WKiŁ, Warszawa 2001, ss. 320.



Prof. dr hab. inż. **Lesław BĘDKOWSKI** jest nauczycielem akademickim w Wydziale Elektroniki Wojskowej Akademii Technicznej w Warszawie. Zainteresowania naukowe Profesora skupiają się głównie wokół następujących problemów: teoria diagnostyki technicznej, optymalizacja

procedur diagnostycznych, diagnostyka systemów antropotechnicznych, diagnostyka w ujęciu potencjałowo-efektywnym, teoria użytkowania w ujęciu wieloprotocowym, wielopoziomowe systemy dozoru i terapeutyczne, diagnostyka bezpieczeństwa.

BADANIE SYMULACYJNE SKUTECZNOŚCI DIAGNOZOWANIA KOMPARACYJNEGO NA PRZYKŁADZIE SYSTEMU ALARMOWEGO

Tadeusz DĄBROWSKI

Instytut Systemów Elektronicznych, Wydział Elektroniki
Wojskowa Akademia Techniczna, ul. S. Kaliskiego 2, 00-908 Warszawa,
fax (22) 683 91 25, e-mail: tdabrowski@wel.wat.edu.pl

Streszczenie

Jednym z podstawowych wymagań stawianych procesom diagnozowania jest żądanie możliwie wysokiej wiarygodności formułowanych diagnoz. Wymaganie to ma szczególne znaczenie w przypadku diagnozowania urządzeń i systemów mających wpływ na bezpieczeństwo ludzi i/lub mienia o dużej wartości. Artykuł poświęcony jest prezentacji wyników symulacyjnych badań skuteczności diagnozowania (charakteryzowanej dwoma wskaźnikami: sprawnością diagnozowania i niezawodnością diagnoz) systemu alarmowego przy zastosowaniu tzw. diagnozowania komparacyjnego. Metoda ta, dzięki wielokrotnemu powtarzaniu testowania oraz wielopoziomowej komparacji symptomów i syndromu stanu, zapewnia – mimo wysokiego prawdopodobieństwa fałszywych wyników poszczególnych testów – wiarygodne diagnozy.

Słowa kluczowe: diagnozowanie, wiarygodność diagnozy, niepewność symptomu i syndromu.

EFFICIENCY OF COMPARATIVE DIAGNOSING FOR AN ALARM SYSTEM – SIMULATION RESULTS

Summary

The key requirement for diagnosing processes is the demand of high authenticity of formulated diagnoses. This requirement is especially important in the case of devices and systems having an influence on man and possessions safety. The paper presents simulation results of diagnosing efficiency of an alarm system characterized by two coefficients: efficiency of diagnosing and reliability of diagnoses, obtained by the use of comparative diagnosing. Despite high probability of false results for individual tests this method provides reliable diagnoses accomplished by repetitive testing and multi-level comparison of symptoms and syndrome of state.

Key words: diagnosing, authenticity of diagnosis, uncertainty of symptom and syndrome.

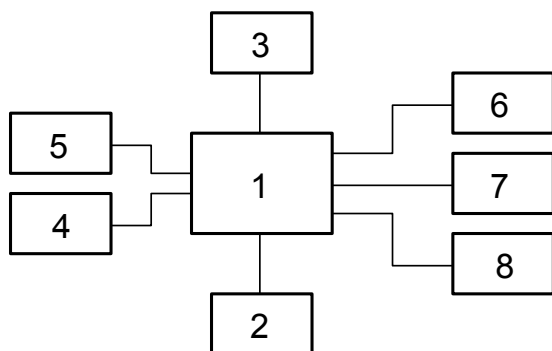
1. CHARAKTERYSTYKA OBIEKTU BADAŃ

System alarmowy należy do kategorii systemów nazywanych zazwyczaj systemami bezpieczeństwa. Pod pojęciem systemu bezpieczeństwa rozumie się na ogół system wykrywania zagrożeń i przeciwdziałania destrukcyjnym procesom. Oddziaływanie na system bezpieczeństwa czynników inicjujących procesy destrukcyjne powinno wywołać jego określoną, pożądaną reakcję. Prawidłowe działanie systemu bezpieczeństwa zależy od wielu zmiennych (np. oddziaływań elektromagnetycznych, klimatycznych, mechanicznych, zasileniowych – a także od stanu technicznego poszczególnych elementów wchodzących w skład systemu). Sygnał wyjściowy systemu bezpieczeństwa powinien być dwustanowy: 0 – brak zagrożeń i system bezpieczeństwa jest zdalny, 1 – występują zagrożenia lub system jest niezdalny. Dlatego bardzo ważnym zagadnieniem w systemach bezpieczeństwa jest cykliczne diagnozowanie stanu

poszczególnych części składowych systemu (np. czujek, kamer, central alarmowych, układów zasilania) i dostarczanie informacji o ich stanie operatorowi systemu – w celu umożliwienia podjęcia, z możliwie małą zwłoką, niezbędnych decyzji użytkowych lub obsługowych.

W niniejszym artykule rozpatrywane jest zagadnienie sprawności diagnozowania i niezawodności diagnoz uzyskiwanych w wyniku testowania poszczególnych modułów systemu alarmowego (np. Systemu Sygnalizacji Włamania i Napadu – SSWiN).

System taki może posiadać różne struktury funkcjonalne i niezawodnościowe – w zależności m. in. od kategorii chronionego obiektu. Dla skonkretyzowania dalszych rozważań przyjmijmy modelową strukturę funkcjonalną SSWiN w postaci pokazanej na rys. 1.



Rys. 1. Przykładowa struktura funkcjonalna systemu alarmowego

Oznaczenia: 1 – centrala alarmowa; 2 – układ zasilania; 3 – manipulator szyfrowy; 4 – monitor (komputer) dozorowy; 5 – układ alarmowania (lokalnego i/lub zdalnego); 6 – czujka ruchu; 7 – czujka rozwarcia; 8 – czujka stłuczenia szyby

Przyjmijmy ponadto, że każdy z modułów funkcjonalnych stanowi zarazem moduł diagnostyczny – gdyż może być diagnozowany (testowany) indywidualnie, niezależnie od pozostałych modułów wchodzących w skład systemu.

Jak już wspomniano, wyniki diagnozowania poszczególnych modułów systemu zależą nie tylko od ich stanu technicznego ale także od różnorodnych oddziaływań środowiskowych (elektrycznych, elektromagnetycznych, temperaturowych, mechanicznych itp.) oraz od celowego działania obsługowego lub destrukcyjnego.

W tej sytuacji należy liczyć się zarówno ze stabilnymi jak i niestabilnymi symptomami stanu. Niezdatności techniczne modułów generują na ogół stabilne zmiany w sygnałach diagnostycznych, natomiast np. zakłócenia komunikacji pomiędzy poszczególnymi modułami systemu wywołane czynnikami atmosferycznymi mają z zasady charakter chwilowy, zanikający.

Niezawodne działanie systemu alarmowego polega m. in. na reagowaniu tylko na takie zmiany stanu systemu i jego otoczenia, które zagrażają bezpieczeństwu chronionego obiektu oraz na ignorowaniu chwilowych zakłóceń i przypadkowych destrukcyjnych oddziaływań.

Zadania realizowane przez systemy bezpieczeństwa i wymagane właściwości użytkowe tych systemów wymuszają stosowanie specyficznych metod pozyskiwania informacji o ich stanie. Pożądane jest testowanie poszczególnych modułów podczas ich funkcjonowania. Dopuszczalne są jedynie krótkie (np. milisekundowe) przerwy na realizację zadań diagnostycznych. Sygnały funkcjonalne i sygnały diagnostyczne nie powinny się wzajemnie zakłócać. Zalecane jest wykorzystywanie w celach diagnostycznych sygnałów funkcjonalnych.

Preferowane są metody cyklicznego diagnozowania (testowania).

Niniejsze opracowanie zawiera opis wyników badań metody diagnozowania z komparacją niepewnych symptomów i syndromów stanu [2] aplikowanej do modelowego systemu alarmowego.

2. WYMAGANIA METODY DIAGNOZOWANIA

Metoda diagnozowania z komparacją niepewnych symptomów i syndromów może być stosowana, gdy zachowane są m. in. następujące warunki:

1. Stany modułów są wzajemnie niezależne.
2. Stan obiektu jest stabilny tzn. nie zmienia się w czasie testowania oraz w czasie ewentualnego powtarzania testowania.
3. W procedurze diagnozowania testowane są wszystkie moduły, sygnał diagnostyczny każdego modułu mierzony jest indywidualnie.
4. Wynikiem testowania jest symptom stanu modułu, tor symptomu (tor testu) obejmuje tylko jeden moduł, tory różnych symptomów są rozłączne.
5. Symptom może przyjmować logiczną wartość 0 lub 1; symptom negatywny 0 jest symptomem charakterystycznym stanu niezdatności modułu, symptom pozytywny 1 jest symptomem charakterystycznym stanu zdatności modułu.
6. Zbiór symptomów stanowi syndrom stanu obiektu.
7. Fałszywy symptom determinuje fałszywy syndrom, a w konsekwencji fałszywą diagnozę.
8. Wyniki komparacji symptomów oraz wyniki komparacji syndromów są pewne.
9. Dokładna znajomość wartości prawdopodobieństwa a priori uzyskania prawdziwego wyniku testu nie jest wymagana – wystarczająca jest wiedza szacunkowa pozwalająca stwierdzić czy większe jest prawdopodobieństwo uzyskania prawdziwych czy też fałszywych symptomów.
10. Prawdopodobieństwa a priori zdatności poszczególnych modułów obiektu nie muszą być znane.

3. CHARAKTERYSTYKA REGUŁ WYBORU DIAGNOZ

W procesie diagnozowania z komparacją niepewnych symptomów i syndromu mogą być zastosowane różne reguły postępowania [2]. W omawianym przykładzie badania skuteczności diagnozowania systemu alarmowego zastosowano trzy reguły: SmW, SmP, SnP.

Reguła **SmW** to **symptomowa reguła większościowa**. Polega ona na realizowaniu nieparzystej liczby testów symptomowych każdego modułu i na wyznaczaniu wartości symptomów negatywnych oraz pozytywnych – aż do spełnienia warunku:

$$L_{i \min} : P_{wi}(\mathcal{E}_i^0) \neq P_{wi}(\mathcal{E}_i^1); \quad i = 1, 2, \dots, N \quad (1)$$

gdzie: P_{wi} – wartość symptomu – odpowiednio: negatywnego i pozytywnego;

$L_{i \min}$ – minimalna, niezbędna liczba testów i-tego modułu.

Dla tej liczby testów, dla każdego modułu, znajduje się większe wartości symptomów:

$$P_{wi}^* = \max[P_{wi}(\mathcal{E}_i^0), P_{wi}(\mathcal{E}_i^1)]; \quad i = 1, 2, \dots, N \quad (2)$$

i na ich podstawie wnioskuje się o stanie poszczególnych modułów.

Reguła **SmP** to **symptomowa reguła progowa**. Polega ona na realizowaniu testów symptomowych każdego modułu. Po zrealizowaniu każdego kolejnego testu oblicza się wartości symptomów negatywnych i pozytywnych oraz sprawdza czy spełniony jest warunek większościowy (1). Jeśli tak to wyznacza się wartość większą zgodnie z (2).

Następnie sprawdza się drugi warunek (progowy):

$$P_{wi}^* \geq P_{igr}; \quad i = 1, 2, \dots, N \quad (3)$$

gdzie: P_{igr} – graniczna wartość symptomu dla i-tego modułu.

Jeśli warunek (3) jest spełniony to testowanie określonego modułu uznaje się za zakończone a na podstawie uzyskanych wartości symptomów wnioskuje się o stanie tego modułu.

Reguła **SnP** to **syndromowa reguła progowa**. Polega ona na postępowaniu analogicznym jak w progowej regule symptomowej z tą jednak różnicą, że podstawą wnioskowania o stanie obiektu jest progowa wartość syndromu:

$$P_S(L) \geq P_{Sgr} \quad (4)$$

gdzie: $P_S(L)$ – wartość syndromu dla zbioru testów o licznosci L ;

P_{Sgr} – graniczna wartość syndromu.

4. CHARAKTERYSTYKA BADAŃ I UZYSKANYCH WYNIKÓW

Badaniom poddano system alarmowy o strukturze przedstawionej na Rys. 1. Podstawowym celem tych badań było wyznaczenie wartości następujących wskaźników [2]:

- **sprawności diagnozowania η_D** ;
- **niezawodności przyjętych diagnoz R_D** ;
- **niezawodności sesji diagnostycznych R_{SD}** .

Wartości tych wskaźników zależą od długości sesji symptomowych (tzn. od liczby testów L_{Ti} realizowanych dla każdego modułu).

Przypomnijmy, że zgodnie z [2], **sprawnością diagnozowania** nazywamy prawdopodobieństwo tego, że uzyskana diagnoza spełnia

wiarygodnościowe wymagania użytkownika diagnozy. Oszacowaniem tego prawdopodobieństwa jest stosunek liczby diagnoz przyjętych (tzn. uznanych za prawdziwe) do liczby obserwowanych syndromowych sesji diagnostycznych:

$$\eta_D^* = \frac{L_{DP}}{L_{SD}} \quad (5)$$

gdzie: L_{DP} – liczba diagnoz przyjętych;
 L_{SD} – liczba syndromowych sesji diagnostycznych.

Podobnie, jak podano w [2], **niezawodnością przyjętych diagnoz** nazywamy prawdopodobieństwo tego, że przyjęta diagnoza okaże się rzeczywiście prawdziwa. Oszacowaniem tego prawdopodobieństwa jest stosunek liczby diagnoz, które w zbiorze diagnoz przyjętych okazały się rzeczywiście prawdziwe (L_{DC}), do liczby diagnoz przyjętych (L_{DP}):

$$R_D^* = \frac{L_{DC}}{L_{DP}} \quad (6)$$

Zauważmy, że niezawodność przyjętych diagnoz jest zarazem prawdopodobieństwem tego, że obiekt uznany (na podstawie diagnozy) za zdatny lub niezdatny okaże się rzeczywiście zdatny lub niezdatny.

Oszacowaniem **niezawodności sesji diagnostycznych** jest stosunek liczby diagnoz prawdziwych (L_{DC}) do liczby obserwowanych sesji syndromowych (L_{SD}):

$$R_{SD}^* = \frac{L_{DC}}{L_{SD}} \quad (7)$$

Dla uproszczenia badań i opisu wyników przyjęto, że długość sesji symptomowej (tzn. liczba testów modułu) jest taka sama dla każdego modułu, czyli: $L_T = L_{Ti}$, $i = 1, \dots, N$.

Oszacowaniem **sprawności diagnozowania**, dla badanych reguł wyboru diagnozy, są następujące wyrażenia:

$$\eta_D^*(L_T; SmW) = \frac{L_{DP}(L_T; SmW)}{L_{SD}} 100\% = W_{1-1} \quad (8)$$

$$\eta_D^*(L_T; SmP) = \frac{L_{DP}(L_T; SmP)}{L_{SD}} 100\% = W_{2-1} \quad (9)$$

$$\eta_D^*(L_T; SnP) = \frac{L_{DP}(L_T; SnP)}{L_{SD}} 100\% = W_{3-1} \quad (10)$$

gdzie: $L_{DP}(L_T; SmW)$; $L_{DP}(L_T; SmP)$;

$L_{DP}(L_T; SnP)$ – liczba diagnoz przyjętych – odpowiednio dla reguł SmW, SmP, SnP.

Oszacowaniem **niezawodności przyjętych diagnoz**, dla rozpatrywanych reguł, są – z kolei – następujące wzory:

$$R_D^*(L_T; SmW) = \frac{L_{DC}(L_T; SmW)}{L_{DP}(L_T; SmW)} 100\% = W_{1-3} \quad (11)$$

$$R_D^*(L_T; SmP) = \frac{L_{DC}(L_T; SmP)}{L_{DP}(L_T; SmP)} 100\% = W_{2-3} \quad (12)$$

$$R_D^*(L_T; SnP) = \frac{L_{DC}(L_T; SnP)}{L_{DP}(L_T; SnP)} 100\% = W_{3-3} \quad (13)$$

gdzie: $L_{DC}(L_T; SmW)$; $L_{DC}(L_T; SmP)$;

$L_{DC}(L_T; SnP)$ – liczba diagnoz, które w zbiorze diagnoz przyjętych okazały się rzeczywiście prawdziwe – odpowiednio dla reguł SmW, SmP, SnP.

Wartość wskaźnika **niezawodności sesji diagnostycznych** wyznaczają następujące wyrażenia:

$$R_{SD}^*(L_T; SmW) = \frac{L_{DC}(L_T; SmW)}{L_{SD}} 100\% = W_{1-2} \quad (14)$$

$$R_{SD}^*(L_T; SmP) = \frac{L_{DC}(L_T; SmP)}{L_{SD}} 100\% = W_{2-2} \quad (15)$$

$$R_{SD}^*(L_T; SnP) = \frac{L_{DC}(L_T; SnP)}{L_{SD}} 100\% = W_{3-2} \quad (16)$$

W celu skonkretyzowania postępowania przyjęto, że minimalna, wymagana sprawność i niezawodność przyjętych diagnoz, dla badanych reguł wyboru diagnozy, wynosi:

$$\eta_{Dmin} = 92\%; \quad R_{Dmin} = 92\%$$

Obserwacjom poddano 200 syndromowych sesji diagnostycznych. Badania przeprowadzono na trzech diagnostycznie i niezawodnościowo różnych modelach systemu alarmowego (jako obiektu diagnozowania) o strukturze jak na rys. 1.

A. Pierwszy z badanych modeli charakteryzował się następującymi parametrami:

– prawdopodobieństwem a priori zdolności każdego z modułów:

$$R_{ei} = 0,5 \quad (i = 1, \dots, 8)$$

– prawdopodobieństwem jednokrotnego wystąpienia prawdziwego symptomu negatywnego:

$$R_{zi}(0|0) = 0,8 \quad (i = 1, \dots, 8)$$

– prawdopodobieństwem jednokrotnego wystąpienia prawdziwego symptomu pozytywnego:

$$R_{zi}(1|1) = 0,8 \quad (i = 1, \dots, 8)$$

– stanem obiektu opisanym w tab. 1.

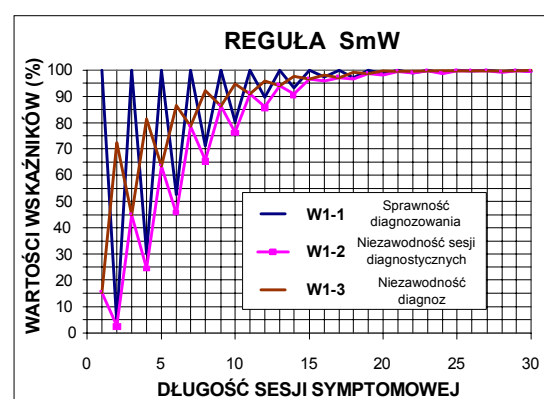
Tabela 1

Moduł	M ₁	M ₂	M ₃	M ₄	M ₅	M ₆	M ₇	M ₈
Stan modułu	$\mathcal{E}_1^1$	$\mathcal{E}_2^1$	$\mathcal{E}_3^1$	$\mathcal{E}_4^0$	$\mathcal{E}_5^1$	$\mathcal{E}_6^1$	$\mathcal{E}_7^1$	$\mathcal{E}_8^1$

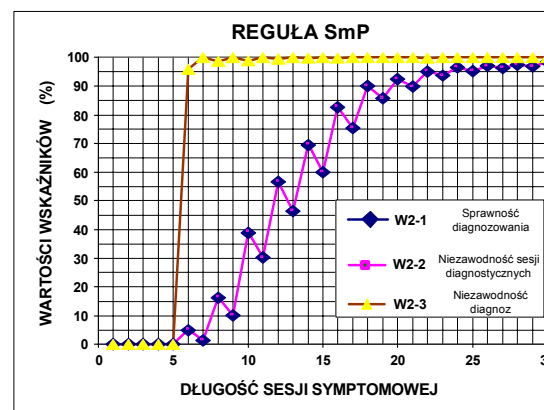
Jak wynika z porównania efektu badań (rys. 2÷4) wymagane wartości wskaźników sprawnościowych i niezawodnościowych uzyskano dla minimalnych długości sesji symptomowych (modułowych) podanych w tab.2.

Tabela 2

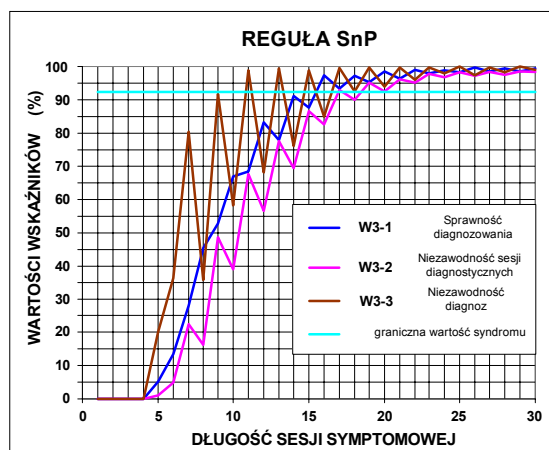
	SmW	SmP	SnP
dla η_{Dmin}	14	22	17
dla R_{Dmin}	12	6	18



Rys. 2. Wyniki badania reguły SmW (model A)



Rys. 3. Wyniki badania reguły SmP (model A)



Rys. 4. Wyniki badania reguły SnP (model A)

B. Drugi z badanych modeli obiektu charakteryzował się następującymi parametrami:

– prawdopodobieństwem a priori zdatości każdego z modułów:

$$R_{ei} = 0,5 \quad (i = 1, \dots, 8)$$

– prawdopodobieństwem jednokrotnego wystąpienia prawdziwego symptomu negatywnego:

$$R_{zi}(0|0) = 0,8 \quad (i = 1, \dots, 8)$$

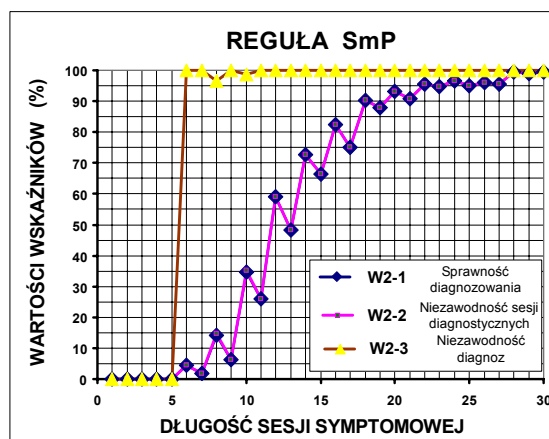
– prawdopodobieństwem jednokrotnego wystąpienia prawdziwego symptomu pozytywnego:

$$R_{zi}(1|1) = 0,8 \quad (i = 1, \dots, 8)$$

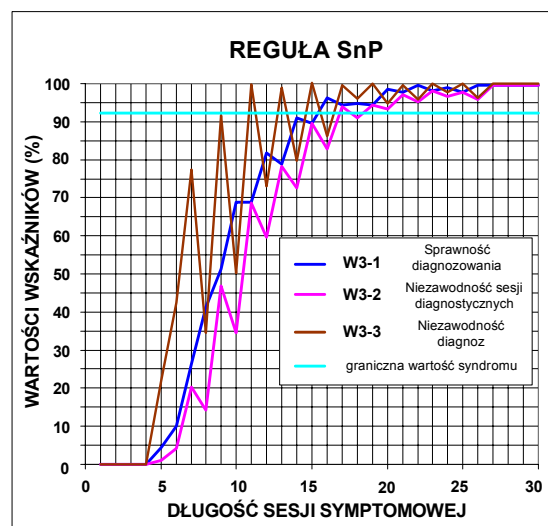
– stanem obiektu opisanym w tab.3.

Tabela 3

Moduł	M ₁	M ₂	M ₃	M ₄	M ₅	M ₆	M ₇	M ₈
Stan modułu	$\mathcal{E}_1^1$	$\mathcal{E}_2^1$	$\mathcal{E}_3^1$	$\mathcal{E}_4^1$	$\mathcal{E}_5^1$	$\mathcal{E}_6^1$	$\mathcal{E}_7^1$	$\mathcal{E}_8^1$



Rys. 5. Wyniki badania reguły SmP (model B)



Rys. 6. Wyniki badania reguły SnP (model B)

Jak wynika z badań wymagane wartości wskaźników sprawnościowych i niezawodnościowych uzyskano dla minimalnych długości sesji symptomowych podanych w tab.4.

Tabela 4

	SmW	SmP	SnP
dla η_{Dmin}	13	22	16
dla R_{Dmin}	14	6	19

C. Trzeci z badanych modeli obiektu charakteryzował się następującymi parametrami:

– prawdopodobieństwem a priori zdatości każdego z modułów:

$$R_{ei} = 0,9 \quad (i = 1, \dots, 8)$$

– prawdopodobieństwem jednokrotnego wystąpienia prawdziwego symptomu negatywnego:

$$R_{zi}(0|0) = 0,8 \quad (i = 1, \dots, 8)$$

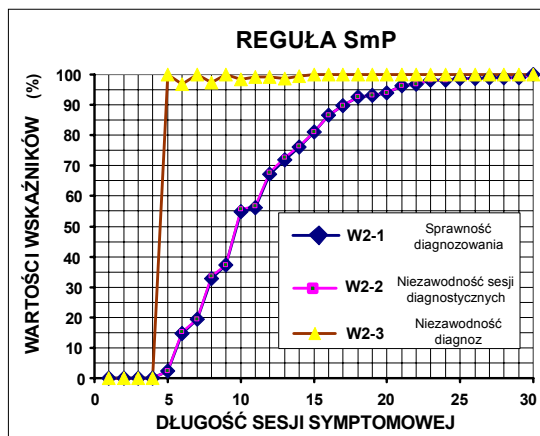
– prawdopodobieństwem jednokrotnego wystąpienia prawdziwego symptomu pozytywnego:

$$R_{zi}(1|1) = 0,8 \quad (i = 1, \dots, 8)$$

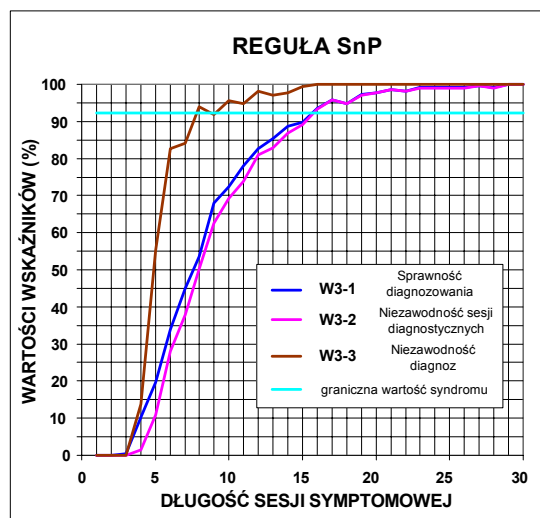
– stanem obiektu opisanym w tab. 5.

Tabela 5

Moduł	M ₁	M ₂	M ₃	M ₄	M ₅	M ₆	M ₇	M ₈
Stan modułu	$\mathcal{E}_1^1$	$\mathcal{E}_2^1$	$\mathcal{E}_3^1$	$\mathcal{E}_4^0$	$\mathcal{E}_5^1$	$\mathcal{E}_6^1$	$\mathcal{E}_7^1$	$\mathcal{E}_8^1$



Rys. 7. Wyniki badania reguły SmP (model C)



Rys. 8. Wyniki badania reguły SnP (model C)

Uzyskane wyniki pokazują, że wymagane wartości wskaźników sprawnościowych i niezawadnościowych osiągnięto dla minimalnych długości sesji symptomowych podanych w tab.6.

Tabela 6

	SmW	SmP	SnP
dla η_{Dmin}	1	18	16
dla R_{Dmin}	10	5	10

5. WNIOSKI

Zauważmy, że dla pierwszego modelu (tzn. modelu A) obiektu:

- przy regule SmW spełnienie wymagań sprawnościowych możliwe jest przy długościach symptomowych sesji diagnostycznych nie mniejszych niż 14 testów (dla każdego modułu), a równocześnie spełnienie wymagań niezawadnościowych możliwe jest przy sesjach symptomowych o długościach nie mniejszych niż 12 testów;
- przy regule SmP spełnienie wymagań sprawnościowych możliwe jest przy długościach symptomowych sesji diagnostycznych nie mniejszych niż 22 testy (dla każdego

modułu), a równocześnie spełnienie wymagań niezawadnościowych możliwe jest przy sesjach symptomowych o długościach nie mniejszych niż 6 testów;

- przy regule SnP spełnienie wymagań sprawnościowych możliwe jest przy długościach symptomowych sesji diagnostycznych nie mniejszych niż 17 testów (dla każdego modułu), a równocześnie spełnienie wymagań niezawadnościowych możliwe jest przy sesjach symptomowych o długościach nie mniejszych niż 18 testów.

Zbliżone wyniki uzyskano w badaniach drugiego i trzeciego modelu.

Warto zauważyć, że zachodzi znaczne podobieństwo wyników przeprowadzonych badań mimo tego, że poszczególne modele różnią się wartością prawdopodobieństwa a priori zdadności modułów oraz zadanymi stanami obiektu.

Na podstawie uzyskanych rezultatów badań można sądzić, że w przypadku gdy ważna jest wysoka niezawadność diagnoz należy stosować regułę SmP. Z kolei w przypadku, gdy ważniejsza jest sprawność diagnozowania korzystnie jest stosować regułę SmW.

Wnioski te potwierdzają się dla różnych stanów obiektu i dla różnych wartości prawdopodobieństw a priori zdadności modułów.

LITERATURA

- [1] Będkowski L., Dąbrowski T.: *Diagnozowanie na podstawie niepewnych syndromów stanu obiektu*. Diagnostyka, PTDT, Vol. 37, 2006, ss. 6.
- [2] Będkowski L.: *Diagnozowanie z dwupoziomą komparacją niepewnych symptomów i syndromu stanu obiektu*. Materiały Konferencji DIAG'2006, ss. 6.



Dr hab. inż. **Tadeusz DĄBROWSKI** jest zatrudniony na stanowisku profesora nadzwyczajnego i pełni funkcję dyrektora Instytutu Systemów Elektronicznych Wydziału Elektroniki WAT. Zainteresowania naukowe koncentruje w obszarze teorii eksploatacji – głównie na diagnostyce technicznej. Do ważniejszych zagadnień, którymi się zajmował i/lub zajmuje należą: diagnostyka systemów antropotechnicznych (w aspekcie użytkowym i bezpieczeństwa); optymalizacja procesów diagnostyczno-obslugowych; komputerowe wspomaganie procesu diagnostyczno-obslugowego.

SIECI STWIERDZEŃ W DIAGNOSTYCE TECHNICZNEJ

Wojciech CHOLEWA

Katedra Podstaw Konstrukcji Maszyn
Politechnika Śląska, ul. Konarskiego 18a, 44-100 Gliwice, e-mail: wch@polsl.pl

Streszczenie

Opisano zwięźle główne klasy systemów bazujących na wiedzy, które mogą być stosowane do wnioskowania w diagnostyce technicznej. Ich szczególne znaczenie jest związane z możliwościami oferowanymi przez systemy stosujące sieci stwierdzeń. Omówiono wybrane zagadnienia związane ze stosowaniem bardzo obiecujących sieci stwierdzeń dynamicznych. Wydaje się, że będą one mogły skutecznie wspomagać między innymi rozwiązywanie zadań diagnostycznych dla obiektów działających w warunkach przejściowych.

Słowa kluczowe: systemy doradcze, sieci stwierdzeń, stwierdzenia dynamiczne.

STATEMENT NETWORKS IN TECHNICAL DIAGNOSTICS

Summary

Paper discusses shortly main classes of knowledge based systems applied for reasoning in technical diagnostics. Their considerable importance is connected with possibilities delivered by systems basing on statement networks. Selected details related to the application of very promising dynamic statement networks are also discussed. It seems that they can effectively aid solving diagnostic tasks for objects running in transient conditions.

Keywords: expert systems, statement networks, dynamic statements.

1. WPROWADZENIE

W celu zapewnienia racjonalnej eksploatacji obiektów technicznych stosowane są między innymi różne rodzaje systemów diagnozujących i monitorujących. Systemy diagnozujące umożliwiają rozpoznawanie stanu lub zmian stanu technicznego rozpatrywanego obiektu na podstawie zgromadzonych wyników obserwacji jego działania. Systemy monitorujące umożliwiają ciągłą obserwację działania obiektu, której wyniki stanowią podstawę do rozpoznawania zachodzących zmian klas stanu technicznego.

Procesy wnioskowania o stanach technicznych obiektów i ich zmianach mogą być wspomagane różnymi sformalizowanymi systemami podejmowania decyzji. Szczególne znaczenie mają systemy nazywane systemami bazującymi na wiedzy. Systemy te dysponując zapisaną wiedzą specjalistów z wybranej dziedziny, mogą ją stosować wielokrotnie, umożliwiając rozwiązywanie zadań bez bezpośredniego udziału specjalisty. W procesie budowania i stosowania takich systemów można wskazać trzy znacznie różniące się role, w jakich występują korzystające z nich osoby lub inne systemy:

- główną rolą jest *użytkownik końcowy*, który stosuje system jako narzędzie wspomagające prowadzony przez tego użytkownika proces

podejmowania decyzji o stanie rozpatrywanego obiektu;

- pozostałe dwie role to *specjalista dziedziny zastosowań* odpowiedzialny za wiedzę zgromadzoną w systemie oraz *inżynier wiedzy* odpowiedzialny za wybór istoty działania systemu i sposobu reprezentacji wiedzy odpowiednio dla uwzględnianej wiedzy.

Rozwojowi tych systemów sprzyja wyraźne rozdzielenie ról specjalisty dziedziny zastosowań i inżyniera wiedzy.

Celem opracowania jest przedstawienie zwięzłego przeglądu wybranych systemów wspomagających wnioskowanie oraz omówienie istoty działania nowych systemów budowanych jako sieci stwierdzeń, a w szczególności sieci stwierdzeń dynamicznych. Sieci te są aktualnie przedmiotem badań. Przewiduje się możliwość ich skutecznego zastosowania w systemach diagnozujących i monitorujących.

2. SYSTEMY DORADCZE

Szczególne miejsce wśród systemów wspomagających podejmowanie decyzji zajmują systemy doradcze, nazywane również systemami ekspertowymi. System doradczy powinien posiadać zdolność do rozpoczęcia i kontynuowania procesu wnioskowania oraz zdolność do wykonywania

działań związanych z tym procesem, wtedy gdy zezwala na to lub wymaga tego użytkownik systemu. Charakterystyczną cechą tych systemów jest możliwość budowania oprogramowania niezależnego od dziedziny zastosowań i występującego w postaci tzw. systemów szkieletowych, zawierających moduły umożliwiające działanie systemu, komunikację ze źródłami danych i dialog z użytkownikiem, w tym objaśnianie wyników działania. W celu uzyskania kompletnego systemu doradczego system szkieletowy uzupełniany jest o bazę wiedzy zawierającą odpowiednio zapisany fragment wiedzy rozpatrywanej dziedziny.

Jednym z najważniejszych elementów systemów doradczych jest układ objaśniający przeznaczony do udostępniania użytkownikowi systemu wszelkich objaśnień, które pozwolą mu na uznanie iż wnioski proponowane przez system są poprawne. Konieczność stosowania objaśnień wynika między innymi stąd, że system doradczy powinien dysponować bazą wiedzy, której zakres znacznie wykracza poza wiedzę posiadaną przez użytkownika co powoduje, że bezpośrednia weryfikacja przez użytkownika wyników działania systemu może być utrudniona lub niemożliwa.

Wiedza zapisana w systemie doradczym może obejmować zarówno wiedzę deklaratywną zawierającą orzeczenia o uznanych faktach, jak i wiedzę proceduralną dotyczącą działania obiektów i ogólnych sposobów postępowania. Wiedza ta może być pozyskiwana [19], [20] z różnych źródeł. Szczególnie ważnymi źródłami wiedzy diagnostycznej są specjaliści tej dziedziny i ich publikacje oraz eksploatacyjne bazy danych i wyniki badań symulacyjnych. Niestety dostępne obecnie metody pozyskiwania wiedzy nie zapewniają jeszcze odpowiedniej efektywności bezpośredniego pozyskiwania wiedzy od specjalistów.

Rozpatrywane mogą być dwie kategorie systemów doradczych [5]:

- systemy statyczne działające w stałym otoczeniu,
- systemy dynamiczne działające w zmieniającym się otoczeniu i przystosowane do realizowania zadań w ograniczonym czasie.

Charakterystyczną cechą systemów doradczych jest to, że sekwencja działań pozwalających na wyznaczenie wniosków nie jest ustalana (programowana) w czasie konstruowania systemu. Kolejne działania są ustalane dynamicznie przez układ wnioskujący systemu, z uwzględnieniem jego aktualnego stanu.

Podstawową klasą systemów doradczych są systemy prowadzące wnioskowanie zgodnie z zasadami logiki klasycznej. W systemach tych baza wiedzy składa się z reguł zapisywanych w postaci:

jeżeli przesłanka to konkluzja (1)

gdzie *przesłanka* jest wyrażeniem logicznym. Reguły takie mogą tworzyć łańcuchy, w których konkluzja wybranej reguły jest jednocześnie przesłanką reguły następującej po niej. Przyjęcie, iż przesłanka pierwszej reguły w łańcuchu może być uznana jako prawdziwa, pozwala na przyjęcie, że konkluzja, czyli równocześnie przesłanka następnej reguły, jest prawdziwa. Pozwala to następnie na tzw. wnioskowanie w przód tzn. na uznawanie, że konkluzje kolejnych reguł są prawdziwe. Podobnie przyjęcie, że konkluzja ostatniej reguły w łańcuchu nie może być uznana jako prawdziwa (tzn. jest fałszywa) pozwala na tzw. wnioskowanie wstecz, czyli na uznawanie, że przesłanki poprzedzających reguł nie są prawdziwe. Zbiory reguł mogą tworzyć struktury inne niż łańcuchy, zwłaszcza wtedy gdy występują w nich przesłanki złożone. Wnioskowanie w takich zbiorach reguł jest prowadzone przez moduły wnioskujące zawarte w systemach szkieletowych.

Literatura omawiająca zagadnienia związane z konstruowaniem i stosowaniem systemów doradczych jest bardzo obszerna. W języku polskim wydano między innymi następujące przeglądowe prace zwarte [1], [5], [6], [7], [8], [14], [18], [21], [23], [25].

3. ZMIENNE OTOCZENIE

Systemy doradcze są przeznaczone do działania w określonym otoczeniu. Większość stosowanych obecnie systemów to systemy statyczne, a dokładniej systemy przeznaczone do działania w statycznym otoczeniu. Użyte określenie "otoczenie statyczne" nie oznacza, że otoczenie systemu jest niezmiennie. Wskazuje ono jedynie, że dla potrzeb działającego w nim systemu uwzględniany jest stan tego otoczenia, aktualny w chwili obserwacji. Równocześnie wiadomo, że zarówno diagnostyka maszyn jak i diagnostyka procesów wymagają uwzględniania informacji o charakterze zmian obserwowanych wielkości fizycznych, związanych z działaniem obserwowanej maszyny lub przebiegiem obserwowanego procesu.

W celu opisanie dynamiki zmian otoczenia stosowanych jest szereg, wprowadzanych ad hoc, rozwiązań. Jednym z najczęściej stosowanych sposobów jest uzupełnianie zbiorów cech uwzględnianych sygnałów o pochodne pierwszego i wyższych rzędów, tych sygnałów. Wartości pochodnych są nośnikami informacji o zachodzących zmianach. Wprowadzenie takich dodatkowych cech sygnałów pozwala na opisywanie otoczenia dynamicznego z zastosowaniem metod przeznaczonych głównie do opisywania otoczenia statycznego. Brak jest jednak uniwersalnych metod postępowania umożliwiających reprezentowanie otoczenia zmieniającego się w czasie. Ciekawy

przegląd związanych z tym zagadnień zamieszczono w [10].

Omawiając systemy, w których wiedza reprezentowana jest za pomocą reguł, należy zwrócić uwagę na to, że systemy te umożliwiają realizowanie wnioskowania monotonicznego. Charakteryzuje się ono tym, że uznanie nowych przesłanek, których wartości logiczne nie były wcześniej znane, nie może prowadzić do zmiany wartości logicznych uznanych dotychczas konkluzji, tzn. nie może wpływać na zmianę uzyskanych już wyników procesu wnioskowania. Własność ta jest szczególnie niedogodna w systemach diagnostycznych, ponieważ wynikiem działania takich systemów mają być wnioski o stanie rozpatrywanego obiektu, który może ulegać zmianie.

W celu uniknięcia ograniczeń wynikających ze stosowania wnioskowania monotonicznego można prowadzić wnioskowanie częściowo niemonotoniczne, zapewniające racjonalne wyznaczanie konkluzji w zmieniających się warunkach zewnętrznych i realizowane jako szereg niezależnych procesów wnioskowania monotonicznego. Procesy te uruchamiane są dla obiektu rozpatrywanego jako obiekt "zamrażany" w kolejnych dyskretnych chwilach czasu. Wynikiem tych procesów są szeregi niezależnie wyznaczonych konkluzji. W każdym z takich szeregów wartości logiczne uznanych konkluzji mogą się zmieniać.

4. PRZYBLIŻONE SYSTEMY DORADCZE

Założenie, iż w bazie wiedzy systemu doradczego występują wyłącznie reguły dokładne zapisywane w postaci (1), jest często nadmiernym ograniczeniem. Źródłem wiedzy mogą być bowiem nie tylko dokładnie znane prawa i fakty. Wiedza rozpatrywanej dziedziny może być udostępniana również w postaci niedokładnych opinii specjalistów, niepewnych wyników obserwacji obiektu lub przybliżonych modeli. Zastosowanie takiej wiedzy wymaga użycia przybliżonych metod wnioskowania [4], które charakteryzują się stosowaniem przybliżonych/niepewnych reguł, przesłanek i konkluzji. Nie opracowano ogólnej metody reprezentowania takich elementów przybliżonych.

Stosowane są dwie klasy przybliżonych metod wnioskowania:

- pierwsza klasa, która nie będzie rozpatrywana w dalszej części opracowania, związana jest ze stosowaniem koncepcji zbiorów i relacji rozmytych [16] [22] [26];
- druga klasa metod polega na stosowaniu różnych kategorii stopni pewności lub stopni prawdziwości elementów występujących w postaci przesłanek, konkluzji oraz reguł.

Bardzo ważnym i trudnym zadaniem jest ustalenie właściwej interpretacji stosowanych miar

(stopni pewności). Miary te mogą być interpretowane jako prawdopodobieństwa zdarzeń polegających na tym, że odpowiedni element jest prawdziwy lub jako stopnie prawdziwości tych elementów. Wprowadzenie prawdopodobieństw pozwala na stosowanie statystycznych metod weryfikacji hipotez. Wymaga to jednak przyjęcia licznych założeń, dla których możliwości weryfikacji ich słuszności są często bardzo ograniczone. Jest to przyczyną rezygnowania ze stosowania prawdopodobieństw definiowanych na podstawie częstości zdarzeń i zastępowania ich prawdopodobieństwami o wartościach ustalanych subiektywnie. Przyjmowane zasady wykonywania działań na stopniach pewności mogą być przyczyną wielu intuicyjnych wątpliwości. Główna wątpliwość związana jest z częstym utożsamianiem wartości logicznej stwierdzenia (prawda, fałsz) z jego stopniem pewności. Liczne wątpliwości związane są również z przyjmowanymi modelami propagacji niepewności i niedokładności w układzie wnioskującym.

5. SIECI PRZEKONAŃ

Przybliżone systemy doradcze umożliwiają stosowanie przybliżonych i/lub niepewnych reguł, przesłanek oraz konkluzji. Stopnie prawdziwości reguł i przesłanek wyznaczane są najczęściej na podstawie subiektywnych opinii specjalistów rozpatrywanej dziedziny i konstruktorów systemu. Wynikiem badań weryfikacyjnych mogą być zalecenia dotyczące niezbędnych modyfikacji tak ustalonych wartości stopni. Proces modyfikowania tych wartości nazywany jest procesem strojenia systemu. Wadą systemów doradczych bazujących na zbiorach reguł są trudności praktycznej realizacji takiego strojenia, zwłaszcza wtedy gdy uwzględniany jest liczny zbiór reguł. Trudności te występują w mniejszym stopniu w systemach, w których bazy wiedzy reprezentowane są w postaci grafów.

Stosowanie grafów jako modelu pozwalającego na reprezentowanie wiedzy ma długą historię, np. [27]. Do ważniejszych koncepcji szczegółowych należy zaliczyć sieci Markova [13] [17], rozpatrywane często w postaci tzw. tablic kontyngencji oraz sieci bayesowskie [24] [2] [12], nazywane również sieciami przekonań. Sieci przekonań spotykają się z coraz większym uznaniem, jako skuteczne narzędzia, przydatne do wnioskowania przybliżonego.

Sieć Markova jest grafem nieskierowanym, w którym każdej gałęzi przypisane jest symetryczne prawdopodobieństwo współwystępowania węzłów łączonych przez tą gałąź. Sieć ta nie pozwala na bezpośrednie reprezentowanie wiedzy dotyczącej zależności przyczynowo-skutkowych nawet wtedy, gdy występują one pomiędzy węzłami.

Sieć przekonań jest acyklicznym (nie zawierającym cykli) grafem skierowanym, składającym się z węzłów i łączących je gałęzi skierowanych. Węzłom przypisywane są kompletne (wyczerpujące) zestawy ich wykluczających się stanów oraz wektory wartości węzłów, określające rozkłady ich stanów i zawierające odpowiednie wartości prawdopodobieństw interpretowanych jako stopnie przekonania o tym, że węzeł znajduje się w określonym stanie. Węzłom przyporządkowane są również tablice zawierające wartości prawdopodobieństw warunkowych dla wszystkich elementów iloczynu kartezjańskiego zestawów stanów węzłów nadrzędnych i węzła posiadającego tablicę. Tablice prawdopodobieństw warunkowych opisują relacje występujące między węzłami. Nie zakłada się że są to relacje przyczynowo skutkowe.

Wnioskowanie z zastosowaniem sieci przekonań, posiadającej kompletne tablice prawdopodobieństw warunkowych, polega na uzgadnianiu prawdopodobieństw przypisanych kolejnym węzłom, w celu wyznaczenia stanu równowagi sieci, w którym spełnione jest twierdzenie Bayesa o prawdopodobieństwach warunkowych. Próby poszukiwania rozwiązań globalnych mogą prowadzić do zadań NP-trudnych. Skutecznym sposobem postępowania jest identyfikowanie węzłów warunkowo niezależnych, a następnie poszukiwanie rozwiązań lokalnych obejmujących kolejne węzły, ich rodziców i dzieci. Ciekawymi pracami omawiającymi sposoby formułowania i rozwiązywania takich zadań są [24], [15].

6. TABLICE OGŁOSZEŃ

W rozwoju sztucznej inteligencji wyjątkową rolę spełniły systemy bazujące na koncepcji tablicy ogłoszeń. Tablica taka jest miejscem udostępniania ogłoszeń (komunikatów, wiadomości) dla ich odbiorców. Źródłami komunikatów i ich odbiorcami mogą być użytkownicy systemu stosującego tablicę ogłoszeń lub inne systemy. Komunikaty dostarczane są przez źródła komunikatów do administratora tablicy. Administrator decyduje o umieszczaniu ogłoszeń na tablicy i o usuwaniu z tablicy ogłoszeń nieaktualnych. Odbiorcy ogłoszeń obserwują tablicę i w odpowiedzi na zmianę jej stanu mogą wykonywać ustalone działania, w tym mogą kierować odpowiednie komunikaty do administratora tablicy. Koncepcja tablicy ogłoszeń, w której ogłoszenia określone były jako źródła wiedzy lub demony, została wprowadzona w systemach przeznaczonych do interpretowania mowy [9] i jest nadal rozwijana [11].

Znanych jest wiele sposobów praktycznej realizacji tablic ogłoszeń. Tablice te mogą być budowane jako hierarchicznie uporządkowane bazy danych, przeznaczone do przechowywania rozwiązań generowanych przez autonomiczne

moduły, uprawnione do użytkowania (stosowania) różnych technik wnioskowania w celu efektywnego osiągnięcia najlepszych rozwiązań [3].

7. SIECI STWIERDZEŃ

Przesłanki i konkluzje występujące w (1) mogą być reprezentowane w różny sposób dla potrzeb systemów doradczych. Należy zwrócić uwagę na rzadko stosowany sposób polegający na zapisywaniu ich w postaci stwierdzeń. Jest to postępowanie szczególnie dogodnie dla systemów przeznaczonych do wspomagania diagnostyki technicznej, wymagającej między innymi rozwiązywania zadań związanych z interpretowaniem wartości danych i ich zmian.

Stwierdzenie jest informacją o uznaniu wypowiedzi orzekającej o obserwowanych faktach lub reprezentującej określoną opinię. Stwierdzenie s może być zapisywane w postaci następującej pary:

$$s = \langle c, b \rangle, \quad (2)$$

gdzie

- c jest *treścią stwierdzenia*, czyli wypowiedzią o tym, że wskazanemu obiektowi przysługuje określony atrybut o ustalonej wartości,
- b jest *wartością stwierdzenia* definiowaną jako stopień prawdziwości lub stopień przekonania o prawdziwości wypowiedzi będącej treścią stwierdzenia.

Zakłada się, że *treść stwierdzenia* jest stała, a *wartość stwierdzenia* może się zmieniać. Stosowanie stwierdzeń pozwala na wyraźne rozróżnianie obiektywnych faktów i wypowiedzi o występujących faktach, które bardzo często, w sposób błędny, są utożsamiane z faktami.

Stwierdzenie może być rozpatrywane jako uogólnienie pojęcia *zdanie* stosowanego w rachunku zdań, gdzie zdanie dotyczy tylko takich wypowiedzi, o których potrafimy orzec czy są prawdziwe czy fałszywe. Omawiane stwierdzenia są stwierdzeniami przybliżonymi ponieważ zbiór dopuszczalnych wartości stwierdzeń nie jest ograniczany do dwóch elementów $\{\text{prawda}, \text{fałsz}\}$.

W procesie wnioskowania, prowadzonym w systemie doradczym na podstawie wartości logicznych przesłanek, przyjmuje się lub odrzuca konkluzje. Rozpatrywany zbiór przesłanek i konkluzji jest zbiorem zamkniętych, ustalonym w czasie budowania bazy wiedzy. Podczas działania systemu nie są generowane lub odkrywane nowe konkluzje. Oznacza to, że procesy wnioskowania prowadzone w systemach doradczych są procesami realizowanymi w tzw. zamkniętych światach. W każdym systemie wszystkie przesłanki i konkluzje występujące w takich procesach mogą być zastąpione skończonym zbiorem stwierdzeń. Wartości tych stwierdzeń mogą być interpretowane

jako wartości logiczne zastępowanych przesłanek i konkluzji.

Rozwijając wyjściową koncepcję tablicy ogłoszeń [9] można przyjąć, że występujące na niej ogłoszenia zostaną zastąpione przez stwierdzenia tworzące sieć stwierdzeń przybliżonych. Zmiany wartości stwierdzeń będą mogły inicjować ciągi działań powodujących zmiany wartości innych stwierdzeń, umożliwiające realizowanie procesu wnioskowania. Stosowane mogą być różne sposoby wnioskowania. Sieci stwierdzeń przybliżonych mogą być na przykład rozpatrywane jako szczególne wersje sieci przekonań, co pozwala na zastosowanie metod wnioskowania opracowanych dla sieci przekonań.

Rozpatrywane stwierdzenia będą należały do jednej z następujących klas:

- *stwierdzenia pierwotne*, których wartości nie zależą (jawnie) od wartości innych stwierdzeń i są zadawane bezpośrednio przez procesy zewnętrzne (np. działające układy pomiarowe lub dialog z użytkownikiem systemu),
- *stwierdzenia wtórne*, których wartości zależą (jawnie) od wartości innych stwierdzeń występujących w sieci i nie są zadawane bezpośrednio przez procesy zewnętrzne.

Część stwierdzeń wtórnych może występować w postaci stwierdzeń ukrytych, pełniących rolę wniosków pośrednich, niedostępnych dla użytkowników systemu

Danymi wejściowymi w procesie wnioskowania w sieciach stwierdzeń są zbiory B_p wartości stwierdzeń pierwotnych należących do zbiorów stwierdzeń pierwotnych S_p

$$B_p = \{b(x) : x \in S_p\}. \quad (3)$$

Celem procesu wnioskowania jest wyznaczenie wartości wszystkich lub wybranych stwierdzeń wtórnych należących do zbioru stwierdzeń wtórnych S_s

$$B_s = \{b(x) : x \in S_s\}. \quad (4)$$

Ze względu na potrzebę uwzględniania zależności występujących pomiędzy stwierdzeniami, wnioskowanie w sieci stwierdzeń rozpatrywane jest najczęściej jako zadanie poszukiwania równowagi w tej sieci, dla danego zbioru wartości stwierdzeń pierwotnych.

8. WARUNKI KONIECZNE I DOSTATECZNE

Sieci stwierdzeń mogą być rozpatrywane jako szczególna wersja baz wiedzy systemów doradczych. Występujące w systemach doradczych reguły (1) zostają zastąpione przez dwie klasy warunków:

- warunki konieczne,
- warunki dostateczne (wystarczające).

Jeżeli uznaniu prawdziwości stwierdzenia x towarzyszy zawsze uznanie prawdziwości stwierdzenia y , lecz niekoniecznie odwrotnie, to x określane jest jako *warunek dostateczny* dla y . Jednocześnie y jest określane jako *warunek konieczny* dla x . Jeżeli x jest równocześnie warunkiem koniecznym i dostatecznym dla y , to także y będzie warunkiem koniecznym i dostatecznym dla x . Podane definicje warunków dostatecznych i koniecznych dotyczą stwierdzeń dokładnych czyli takich, które są jednoznacznie uznane lub nie. Dla stwierdzeń przybliżonych, które mogą być uznane częściowo, czyli dla stwierdzeń x i y o wartościach $b(x)$ i:

$$b(x) \in [0,1], \quad b(y) = [0,1]. \quad (5)$$

Informację o tym, że x jest warunkiem dostatecznym dla y lub, że y jest warunkiem koniecznym dla x możemy zapisywać [5]:

$$b(y) \geq b(x). \quad (6)$$

Warunki konieczne i dostateczne określane dla stwierdzeń dokładnych są szczególnym przypadkiem warunków określanych dla stwierdzeń przybliżonych zgodnie z (6).

Uwzględniając potrzebę stosowania reguł przybliżonych interpretowanych tak, że są one w przybliżeniu warunkami koniecznymi i dostatecznymi, czyli warunkami koniecznymi i dostatecznymi spełnianymi z jakąś (niewielką) niedokładnością, można przekształcić warunek (6) do postaci przybliżonej:

$$b(y) \geq b(x) - \delta; \quad \delta \geq 0, \quad (7)$$

gdzie δ jest wspólną dla wszystkich rozpatrywanych warunków, graniczną nieujemną wartością parametru, określającego dopuszczalny stopień przybliżenia (niedokładności) warunku. Dla warunku dokładnego $\delta=0$.

9. FORMALNA POPRAWNOŚĆ BAZ WIEDZY

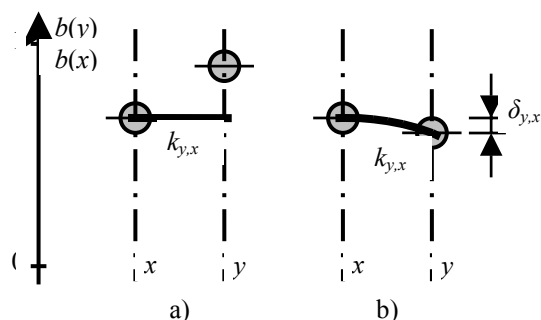
Bazy wiedzy systemów doradczych powinny zawierać wyłącznie elementy niesprzeczne. Formalna poprawność baz wiedzy powinna być przedmiotem weryfikacji w czasie budowania i eksploatacji systemu. W bazach wiedzy stosujących reguły wykrywanie reguł sprzecznych jest zadaniem prostym wtedy gdy rozpatrywane są systemy doradcze stosujące dokładne przesłanki i dokładne reguły. Niestety, jest to zadanie trudne dla przybliżonych systemów doradczych. W celu jego uproszczenia zakłada się, że ewentualne reguły sprzeczne mogą być wykrywane w czasie działania systemu dla ustalonego przypadku, czyli dla danego zbioru uznanych przesłanek. Dla omawianych przybliżonych warunków koniecznych i dostatecznych można zastosować postępowanie

podobne. Polega ono na poszukiwaniu minimalnej wartości parametru δ w (7), zapewniającej spełnienie wszystkich rozpatrywanych warunków koniecznych i dostatecznych.

Proces wnioskowania z uwzględnieniem warunków koniecznych i dostatecznych polega na poszukiwaniu stanu równowagi w sieci stwierdzeń. Wymaga to rozwiązania zadania programowania dynamicznego. Dla rozpatrywanego przypadku wyznaczona minimalna wartość parametru δ może być również interpretowana jako ocena stopnia sprzeczności uwzględnianego zbioru warunków przybliżonych, czyli ocena stopnia sprzeczności występujących pomiędzy elementami bazy wiedzy.

10. MODEL SIECI STWIERDZEŃ

W celu rozwinięcia omawianych metod wnioskowania w sieciach stwierdzeń, rozpatrujemy model będący analogią mechaniczną tych sieci. Przyjmijmy, że stwierdzenia będące elementami sieci stwierdzeń, reprezentowane są przez punkty materialne, które mogą się przemieszczać. Punkty te można przedstawiać graficznie jak na rys. 1 zakładając, że mogą się one przemieszczać wyłącznie w kierunku pionowym, podczas gdy ich odległości w kierunku poziomym są zerowe. Na rys. 1 punkty te pokazano jako "rozsunięte" w kierunku poziomym, wyłącznie w celu umożliwienia ich porównywania i niezależnego obserwowania każdego z nich. Wysokość, na której zlokalizowano każdy z tych punktów, odpowiada wartości stwierdzenia i należy do przedziału $[0, 1]$. Stwierdzenia w pełni uznane za prawdziwe reprezentowane są przez punkty położone na wysokości równej 1. Stwierdzenia uznane za fałszywe reprezentowane są przez punkty położone na wysokości równej 0.



Rys. 1. Model reprezentujący warunek dostateczny dla pary stwierdzeń x i y .
a) spełniony warunek dokładny (6),
b) spełniony warunek przybliżony (8)

Warunki konieczne i dostateczne reprezentowane są przez więzy jednostronne nałożone na rozpatrywane punkty i odpowiadające nierównościom (8).

$$b(y) \geq b(x) - \delta_{y,x}; \quad \delta_{y,x} \geq 0 \quad (8)$$

Nierówność (8) uzyskano w wyniku modyfikacji nierówności (7), polegającej na indywidualizacji parametru δ i założeniu, że może on przyjmować różne wartości dla kolejnych par stwierdzeń związanych z omawianymi warunkami. Na rys. 1 pokazano element podatny o sztywności $k_{y,x}$, który jest zginany wtedy gdy warunek konieczny lub dostateczny (6) musi być uwzględniany jako warunek przybliżony (8) częściowo sprzeczny z innymi warunkami, tzn. wtedy gdy $\delta_{y,x} > 0$.

Wprowadzone sztywności $k_{y,x}$ umożliwiają różnicowanie ważności warunków przybliżonych. Są one jednostkowymi kosztami ewentualnych występujących sprzeczności warunków. Mogą przyjmować równe wartości (np. 1.0), wtedy gdy brak potrzeb różnicowania ważności warunków. Sumaryczny koszt sprzeczności występujących pomiędzy elementami bazy wiedzy:

$$E = \sum_{(y,x)} \delta_{y,x} k_{y,x}, \quad (9)$$

może być interpretowany jako energia potencjalna rozpatrywanego zbioru punktów.

Wnioskowanie w rozpatrywanym modelu sieci stwierdzeń polega na poszukiwaniu stanu równowagi, któremu odpowiada minimalna energia potencjalna E wg (9). Należy zwrócić uwagę, że nieliniowy charakter więzów reprezentujących warunki (więzy jednostronne) powoduje, iż poszukiwanie stanu równowagi sieci jest zadaniem nieliniowym.

11. SIECI STWIERDZEŃ DYNAMICZNYCH

Wspólną cechą opisanych systemów doradczych, sieci przekonań i sieci stwierdzeń jest to, że systemy te stosują statyczne bazy wiedzy i nie posiadają pamięci umożliwiającej wykorzystanie historii ich działań w procesie wnioskowania. Wynik ich działania wyznaczany jest niezwłocznie, bezpośrednio w odpowiedzi na zmianę danych wejściowych. Wyniki działania sieci przekonań i sieci stwierdzeń nie zależą od kolejności zmian danych wejściowych. Nie odpowiada to w pełni procesom wnioskowania prowadzonym przez człowieka.

Podczas wyciągania wniosków na podstawie zbioru dostępnych stwierdzeń człowiek często uwzględnia nie tylko wartości stwierdzeń ale również charakter ich zmian i czasokres utrzymywania się tych wartości bez zmian. Należy zauważyć, że często wyciągamy wnioski (uznajemy stwierdzenia) z pewnym opóźnieniem względem chwili, w której poznaliśmy przesłanki. Opóźnienie takie dotyczy również wniosków pośrednich i może wpływać jakościowo na przebieg złożonego procesu wnioskowania, zwłaszcza wtedy gdy dla

uwzględnianego zbioru warunków istnieje wiele (kilka) rozwiązań statycznych. Występowanie opóźnień wniosków pośrednich może stwarzać między innymi warunki pozwalające na rozróżnianie stanów technicznych charakteryzujących się identycznymi symptomami w stanach ustalonych i różniącymi się wyłącznie przebiegami czasowymi zmian symptomów w stanach przejściowych. Opóźnienia pozwalają również na ograniczanie wpływu zakłóceń na wynik procesu wnioskowania.

Podczas budowania systemów stosujących sieci stwierdzeń można wprowadzać układy wstępnego przetwarzania danych, występujące w postaci filtrów dolnoprzepustowych pomiędzy stwierdzeniami pierwotnymi i odpowiadającymi im źródłami danych. Filtry takie umożliwiają uśrednianie danych wejściowych i pozwalają na uzyskanie omawianego opóźnienia. Zastosowanie podobnych filtrów w celu opóźnienia stwierdzeń wtórnych jest zadaniem bardzo trudnym.

W celu uzyskania możliwości wpływania na dynamikę zmian wartości stwierdzeń wtórnych w procesach wnioskowania w sieciach stwierdzeń, wprowadza się pojęcie *stwierdzenia dynamicznego*. Istotą proponowanej modyfikacji jest uwzględnienie bezwładności stwierdzenia, decydującej o czasie niezbędnym na dokonanie zmiany wartości stwierdzenia wtórnego. Zakłada się, że bezwładność jest cechą stwierdzenia, której wartość jest stała. Stwierdzenia wtórne o dużych bezwładnościach będą zmieniały swoje wartości wolniej niż stwierdzenia wtórne o małych bezwładnościach.

Modelem sieci zawierającej stwierdzenia dynamiczne, czyli modelem sieci stwierdzeń dynamicznych, może być zmodyfikowany model sieci stwierdzeń. Modyfikacja tego modelu polega na przypisaniu mas punktom materialnym, reprezentującym stwierdzenia. W tak przedstawianej sieci możliwe jest realizowanie procesów wnioskowania, polegających na propagacji zmian wartości stwierdzeń, rozpatrywanych jako przemieszczenia oddziaływujących na siebie punktów materialnych o określonych masach.

WNIOSKI

Procesy wnioskowania w systemach diagnostycznych wymagają wspomagania za pomocą systemów bazujących na wiedzy. Szczególne znaczenie, ze względu na możliwości zastosowania, mają systemy doradcze, w których baza wiedzy zorganizowana jest w postaci sieci stwierdzeń. Należy oczekiwać, że zastosowanie znajdą również sieci stwierdzeń dynamicznych pozwalających na wprowadzanie bezwładności stwierdzeń. Zagadnienia dotyczące metod definiowania oraz stosowania tych sieci wymagają dalszych badań.

LITERATURA

- [1] Bubnicki Z.: *Wstęp do systemów ekspertowych*, PWN, Warszawa 1990.
- [2] Charniak E.: *Bayesian networks without tears*. AI Magazine Vol.12 (1991), No.4, s.50-63.
- [3] Cholewa W.: *Tablice ogłoszeń w diagnostycznych systemach doradczych*. Pomiary, Automatyka, Kontrola. 1998, nr 4, s.123-128.
- [4] Cholewa W.: *Wnioskowanie przybliżone w dynamicznych systemach doradczych*. w: "Zbiory rozmyte i ich zastosowanie" red. J. Chojcan, J. Łęski. Wyd. Politechniki Śląskiej, Gliwice 2001.
- [5] Cholewa W.: *Systemy doradcze w diagnostyce technicznej*. w: "Diagnostyka procesów. Modele. Metody sztucznej inteligencji. Zastosowania". WNT, Warszawa 2002, rozdz. 15, s.543-580.
- [6] Cholewa W., Czogała E.: *Podstawy systemów ekspertowych*. Prace Instytutu Biocybernetyki i Inżynierii Biomedycznej PAN, nr 28, Warszawa 1989.
- [7] Cholewa W., Pedrycz W.: *Systemy doradcze*. Wydawnictwo Politechniki Śląskiej, Gliwice 1987.
- [8] Chromiec J., Strzemieczna E.: *Sztuczna inteligencja. Metody konstrukcji i analizy systemów eksperckich*. Akademicka Oficyna Wydawnicza PLJ, Warszawa 1994 (wydanie drugie zmienione 1995).
- [9] Englemore R., Morgan T. (eds): *Blackboard systems*. Addison-Wesley 1988.
- [10] Hajnicz E.: *Reprezentacja logiczna wiedzy zmieniającej się w czasie*. Akademicka Oficyna Wydawnicza PLJ, Warszawa 1996.
- [11] Hayes-Roth B.: *An architecture for adaptive intelligent systems*. Artificial Intelligence, Vol.72 (1995), s.329-365.
- [12] Henrion M., Breese J. S., Horvitz E. J.: *Decision analysis and expert systems*. AI Magazine, Vol. 12 (1991), No.4, s.64-91.
- [13] Isham V.: *An introduction to spatial point processes and Markov random fields*. Intl. Statist. Review, 49 (1981), s.21-43.
- [14] Jagielski J.: *Inżynieria wiedzy w systemach ekspertowych*. Lubuskie Towarzystwo Naukowe, Zielona Góra 2001.
- [15] Jensen F. V.: *Bayesian networks and decision graphs*. Springer 2001.
- [16] Kuncheva L.: *Fuzzy classifier design*. Physica-Verlag, Heidelberg 2000.
- [17] Lauritzen S. L.: *Lectures on contingency tables*. 2nd ed. University of Aalborg Press, Aalborg 1982.
- [18] Moczulski W.: *Metody pozyskiwania wiedzy dla potrzeb diagnostyki maszyn*. Zeszyty Naukowe Politechniki Śląskiej, Gliwice 1997.

- [19] Moczulski W.: *Diagnostyka techniczna. Metody pozyskiwania wiedzy*. Wydawnictwo Politechniki Śląskiej, Gliwice 2002.
- [20] Moczulski W.: *Metody pozyskiwania wiedzy diagnostycznej*. w: "Diagnostyka procesów. Modele. Metody sztucznej inteligencji. Zastosowania". WNT, Warszawa 2002, rozdz. 17, s.623-664.
- [21] Mulawka J. J.: *Systemy ekspertowe*. WNT, Warszawa 1996.
- [22] Negnevitsky M.: *Artificial intelligence. A guide to intelligent systems*. Addison-Wesley, 2002.
- [23] Niederliński A.: *Regulowe systemy ekspertowe*. Wyd. Pracowni Komputerowej Jacka Skalmierskiego. Gliwice 2000.
- [24] Pearl J.: *Probabilistic reasoning in intelligent systems. Networks of plausible inference*. Morgan Kaufmann, San Mateo 1988.
- [25] Rutkowski L.: *Metody i techniki sztucznej inteligencji*. PWN, Warszawa 2005.
- [26] Siler W., Buckley *Fuzzy expert systems and fuzzy reasoning*. John Wiley, New Jersey 2005.
- [27] Wright S.: *The method of path coefficients*. Ann. Math. Statistics. 5 (1934), s. 161-215.



Prof. dr hab. inż.
Wojciech CHOLEWA
prowadzi badania
dotyczące diagnostyki
technicznej, ze szczegó-
lnym uwzględnieniem
metod pomiaru i analizy
sygnałów oraz metod
wnioskowania diagnosty-
cznego. Rozwija metody
i techniki sztucznej

inteligencji związane z reprezentowaniem oraz przetwarzaniem wiedzy niepewnej i przybliżonej. Jest autorem publikacji dotyczących dynamicznych systemów doradczych budowanych dla potrzeb diagnostyki technicznej oraz wspomagania procesów projektowania i konstruowania maszyn.

Jest członkiem Komitetu Budowy Maszyn PAN oraz członkiem zwyczajnym Akademii Inżynierskiej w Polsce.

NADZÓR NAD SPRZĘTEM POMIAROWO-KONTROLNYM - UNORMOWANIA I PRZEPISY PRAWNE

Eliza JARYSZ-KAMIŃSKA

Politechnika Szczecińska, Instytut Technologii Mechanicznej
al. Piastów 19

e-mail: ejarysz@ps.pl, tel. 602-17-65-17

Streszczenie

W artykule przedstawiono przepisy obligatoryjne i unormowania fakultatywne prawa polskiego oraz wymagania międzynarodowe dotyczące nadzoru nad sprzętem pomiarowo-kontrolnym, stosowane w celu zapewnienia prawidłowego przebiegu procesów produkcyjnych i możliwości zapewnienia wiarygodności pomiarów. Omówiono także korzyści dla przedsiębiorstw wynikające z uwzględnienia przepisów obligatoryjnych.

Słowa kluczowe: system zarządzania jakością, legalizacja, wzorcowanie, niepewność pomiaru.

SUPERVISING OF MEASURING-CONTROL EQUIPMENT

Summary

This paper presents obligatory regulations and un compulsory adjustment of polish law and also international claims (demands) concerning supervising of measuring-control equipment, use to providing good course of productive process and ability to provide reliable measurements. It also shows benefits which firms may get by subjecting to un compulsory regulations.

Keywords: quality management systems, legalization, verification, uncertainty of measurement.

1. WPROWADZENIE

W dozowaniu procesów wytwarzania jednym z nieodzownych elementów jest nadzór nad sprzętem pomiarowo-kontrolnym. Określenie zakresu tego nadzoru zależne jest od charakterystyki procesu wytwarzania, od rodzaju wyrobu bądź usługi produkowanej przez przedsiębiorstwo. Nadzór ten, tak jak proces wytwarzania czy cykl życia produktu, związany jest z kolejnymi etapami życia wyrobu. Zaczyna się, więc w momencie planowania realizacji procesu wytwarzania, a kończy na likwidacji wyrobu.

Celem nadzoru nad sprzętem pomiarowo-kontrolnym jest zarządzanie ryzykiem dotyczącym wyposażenia pomiarowego i procesów pomiarowych, które mogłyby doprowadzić do nieprawidłowych wyników wpływających na jakość wyrobów organizacji. Skuteczny system zarządzania pomiarami zapewnia, że wyposażenie pomiarowe jak i procesy pomiarowe, są przystosowane do wyznaczonych im zastosowań. Gospodarka środkami pomiarowo-kontrolnymi musi zapewnić odpowiedni dobór narzędzi pomiarowych do wybranych zadań produkcyjnych, tak, aby cechy metrologiczne danego narzędzia (zakres pomiarowy, klasa dokładności, czułość, dokładność, wartość działki elementarnej, niepewność pomiaru itd.) były zgodne z wymaganiami. Należy jak najdokładniej

określić potrzeby w zakresie parametrów, wykorzystania, sposobów użytkowania, miejsca i czasu przeprowadzania kontroli, jak również ilości niezbędnych przyrządów pomiarowo-kontrolnych, odstępów czasu między kolejnymi kontrolami i sposobów wycofania z eksploatacji.

Gospodarka wyposażeniem kontrolno-pomiarowym jest jednym z elementów systemu jakości, podobnie jak i inne elementy procesów wytwarzania. Gospodarka ta związana jest zarówno z unormowaniami fakultatywnymi, charakterystycznymi dla danej branży i jednocześnie z unormowaniami obligatoryjnymi w formie rozporządzeń wydanych przez odpowiednie organy władzy państwowej dla danej gałęzi przemysłu.

2. PRZEPISY PRAWNE

W Polsce urzędem administracji rządowej zajmującym się sprawami miar i probiernictwa jest Główny Urząd Miar. Prezes Głównego Urzędu Miar zgodnie z ustawą z dnia 11 maja 2001 r. Prawo o miarach [5] nadzorowane jest przez ministra właściwego do spraw gospodarki.

W polskim prawodawstwie nadrzędnym aktem prawnym jest ustawa Prawo o miarach z dnia 11 maja 2001r. [5]. Ten akt prawny ma na celu zapewnienie jednolitości miar i dokładności pomiarów poprzez uregulowanie zagadnień:

1. legalnych jednostek miar i państwowych wzorców jednostek miar,
2. prawnej kontroli metrologicznej przyrządów pomiarowych,
3. kompetencji i zadań organów administracji rządowej właściwych w sprawach miar,
4. sprawowania nadzoru nad wykonywaniem przepisów ustawy.

W ustawie Prawo o miarach [5] zostały wprowadzone pewne zmiany ustawą z dnia 27 maja 2004 r. o zmianie ustawy – Prawo o miarach (Dz. U. Nr 141, poz. 1493), która weszła w życie z dniem 6 lipca 2004r.

Innym aktem prawnym jest ustawa z dnia 30 sierpnia 2002 r. o systemie oceny zgodności (Dz. U. Nr 166, poz. 1360) określająca:

1. zasady funkcjonowania systemu oceny zgodności z zasadniczymi i szczegółowymi wymaganiami dotyczącymi wyrobów, jak również procesów ich wytwarzania, które mogą stwarzać zagrożenie albo służyć ochronie lub ratowaniu życia, zdrowia, mienia oraz środowiska;
2. zasady i tryb udzielania akredytacji oraz autoryzacji;
3. sposób zgłaszania Komisji Europejskiej i państwom członkowskim Unii Europejskiej autoryzowanych jednostek oraz autoryzowanych laboratoriów;
4. zadania Polskiego Centrum Akredytacji;
5. zasady sprawowania nadzoru nad wyrobami podlegającymi ocenie zgodności oraz organy właściwe w tych sprawach.

Dla zapewnienia sprawności i zgodności przyrządów pomiarowych należy zapoznać się zarówno z wymaganiami ustawy Prawo o miarach [5], jak i wszystkimi rozporządzeniami, dotyczącymi wybranego zagadnienia, wydanymi przez uprawnione do tego organy. Można wymienić tu m.in. rozporządzenia, stanowiące akty wykonawcze stosowane w działalności Głównego Urzędu Miar, jak:

- Rozporządzeniem Ministra Gospodarki, Pracy i Polityki Społecznej z dnia 2 kwietnia 2004 r. w sprawie prawnej kontroli metrologicznej przyrządów pomiarowych (Dz. U. Nr 77, poz. 730). Rozporządzenie to określa:

1. warunki i tryb zgłaszania przyrządów pomiarowych do prawnej kontroli metrologicznej;
2. szczegółowy tryb wykonywania prawnej kontroli metrologicznej przyrządów pomiarowych;
3. wzory znaków zatwierdzania typu;
4. wzory oraz okresy ważności dowodów prawnej kontroli metrologicznej.

- Rozporządzenie Ministra Gospodarki i Pracy z dnia 30 marca 2005 r. w sprawie rodzajów przyrządów pomiarowych podlegających prawnej kontroli metrologicznej oraz zakresu tej kontroli. Z chwilą wejścia w życie rozporządzenia moc swoją traci Rozporządzenie Ministra Gospodarki, Pracy i Polityki Społecznej z dnia 20 lutego 2003 r. w sprawie przyrządów pomiarowych podlegających prawnej kontroli metrologicznej oraz rodzajów

przyrządów pomiarowych, które są legalizowane bez zatwierdzenia typu (Dz. U. Nr 41, poz. 351 oraz z 2004 r. Nr 82, poz. 754).

Zarządzenie nr 3 Prezesa Głównego Urzędu Miar z dnia 7 marca 2000 (Dz. U. Miar i Probiernictwa Nr 1 z dnia 14.03.2000) uchyliło wszystkie przepisy i instrukcje sprawdzania przyrządów do pomiaru wielkości geometrycznych. Dało to możliwość użytkownikom przyrządów opracowywania własnych procedur wzorcowania/sprawdzania przyrządów na podstawie obowiązujących norm krajowych lub zagranicznych oraz innych uznawanych dokumentów.

Do tej pory podstawą do opracowania procedur wzorcowania/sprawdzania większości przyrządów do pomiaru długości i kąta były przepisy metrologiczne i instrukcje sprawdzania opracowane w Głównym Urzędzie miar i opublikowane w Dziennikach Urzędowych Miar i Probiernictwa.

Prawna kontrola metrologiczna

Prawną kontrolą metrologiczną jest działanie zmierzające do wykazania, że przyrządy pomiarowe rozumiane w ustawie jako urządzenia, układy pomiarowe lub jego elementy przeznaczone do wykonywania pomiarów samodzielnie lub w połączeniu z jednym lub wieloma urządzeniami pomiarowymi, bądź wzorce miary i materiały odniesienia [5], spełniają wymagania techniczne i metrologiczne określone we właściwych przepisach.

Prawnej kontroli metrologicznej podlegają przyrządy pomiarowe stosowane:

- w ochronie zdrowia, życia i środowiska,
- w ochronie bezpieczeństwa i porządku publicznego
- w ochronie praw konsumenta,
- przy pobieraniu opłat, podatków i nieopodatkowanych należności budżetowych,
- przy dokonywaniu kontroli celnej,
- w obrocie handlowym.

Prawna kontrola metrologiczna przyrządów pomiarowych wykonywana jest przez :

1. zatwierdzenie typu przyrządu pomiarowego na podstawie badania typu – przed wprowadzeniem typu przyrządu do obrotu;
2. legalizację pierwotną albo legalizację jednostkową – przed wprowadzeniem danego egzemplarza przyrządu pomiarowego do obrotu lub użytkowania;
3. legalizację ponowną – w stosunku do przyrządów pomiarowych wprowadzanych do obrotu lub użytkowania [5].

Przyrząd pomiarowy (zgodnie z ustawą Prawo o miarach [5]) jest to urządzenie, układ pomiarowy lub jego elementy, przeznaczone do wykonywania pomiarów samodzielnie lub w połączeniu z jednym lub wieloma urządzeniami dodatkowymi: wzorce miary i materiały odniesienia są traktowane jako przyrządy pomiarowe.

Zatwierdzenie typu jest decyzją Prezesa GUM o potwierdzeniu, że typ przyrządu pomiarowego

spełnia wymagania. Jest badaniem mającym na celu otrzymanie charakterystyki metrologicznej przyrządu, na podstawie której określone zostaną metody legalizacji danego typu przyrządu pomiarowego. Badanie typu przyrządu pomiarowego, na podstawie badań przed wprowadzeniem danego egzemplarza przyrządu pomiarowego do obrotu, obejmuje: analizę dokumentów i badanie charakterystyk technicznych oraz metrologicznych. Wydając decyzję o zatwierdzeniu typu bądź zatwierdzeniu typu z ograniczeniami Prezes GUM nadaje znak zatwierdzenia typu jak również określa miejsce umieszczenia cech legalizacyjnych oraz cech zabezpieczających. Prezes GUM może cofnąć decyzję o zatwierdzeniu typu w przypadku, kiedy przyrząd pomiarowy wykonany zgodnie z zatwierdzonym typem posiada wady utrudniające jego prawidłowe użytkowanie. Okres zatwierdzenia typu wynosi 10 lat, o ile przepisy odrębne nie stanowią inaczej [5].

Legalizacja definiowana jest przez Prawo o miarach [5], jako zespół czynności obejmujących sprawdzenie, stwierdzenie i poświadczenie dowodem legalizacji, że przyrząd pomiarowy spełnia wymagania.

Legalizacja obejmuje:

- legalizację jednostkową obejmuje analizę dokumentów i badań charakterystyk technicznych oraz metrologicznych przyrządu
- legalizację ponowną, czyli każdą kolejną legalizację przyrządu wprowadzonego do użytkowania, wykonywaną po upływie okresu ważności legalizacji pierwotnej, przy zmianie miejsca użytkowania i naprawie przyrządu; wykonywana jest na wniosek użytkownika bądź wykonawcę naprawy lub instalacji przyrządu pomiarowego; obejmuje oględziny przyrządu w celu stwierdzenia czy przyrząd pomiarowy posiada wymagane oznaczenia i znaki oraz czy nie jest uszkodzony i zgodność charakterystyk metrologicznych z wymaganiami.
- legalizacja pierwotna obejmuje sprawdzenie zgodności konstrukcji z dokumentacją techniczno-konstrukcyjną, oznaczeń i znaków, wykonania, materiałów i charakterystyk metrologicznych z zatwierdzonym typem lub wymaganiami.

Legalizacji podlegają przyrządy pomiarowe wymienione w Rozporządzeniu Ministra Gospodarki, Pracy i Polityki społecznej z dnia 20 lutego 2003 roku (Dz. U. Nr 41 poz. 351 zmiana z 2004 roku Dz. U. 141).

Legalizacja jest ważna przez czas określony jednakże, gdy przyrząd pomiarowy podlega tylko zatwierdzeniu typu i legalizacji pierwotnej albo legalizacji jednostkowej, legalizacja jest ważna przez czas nieokreślony [5].

Zasady zgłaszania przyrządów pomiarowych podlegających prawnej kontroli metrologicznej określa w drodze rozporządzenia minister właściwy do spraw gospodarki.

Przyrządy pomiarowe, które nie podlegają prawnej kontroli metrologicznej, mogą być poddane wzorcowaniu, którego poświadczeniem jest świadectwo wzorcowania wystawiane przez organ administracji miar, bądź podmiot upoważniony do przeprowadzania legalizacji. Zgodnie z EN ISO/IEC 17025:2005 [8] (pkt 5.6.2.1 i 5.10.4.2.) świadectwa wzorcowań, wydawane przez laboratoria, powinny zawierać wyniki pomiarów wraz z niepewnością pomiaru i/lub stwierdzenie zgodności z określoną specyfikacją metrologiczną. Ponadto, gdy jest to konieczne do interpretacji wyników wzorcowań, świadectwo wzorcowania powinno zawierać następujące dane: warunki (np. środowiskowe), w których wykonano wzorcowanie, a które mają wpływ na wyniki pomiarów oraz dowód, że zapewniono spójność pomiarową.

Przez wzorcowanie należy rozumieć zbiór operacji ustalających, w określonych warunkach, relację między wartościami wielkości mierzonej wskazanymi przez przyrząd pomiarowy lub układ pomiarowy albo wartościami reprezentowanymi przez wzorzec miary lub przez materiał odniesienia, a odpowiednimi wartościami wielkości realizowanymi przez wzorce jednostki miary. Wynik wzorcowania pozwala na przypisanie wskazaniom odpowiednich wartości wielkości mierzonej lub na wyznaczenie poprawek wskazań [3]. Wzorcowanie wykonuje się po określeniu przez użytkownika metody sprawdzania oraz dla jakich wielkości i w jakich zakresach wielkości mierzonych zostanie wykalibrowany przyrząd pomiarowy. Zgodnie z Prawem o miarach [5] wzorcowanie, na wniosek zainteresowanego podmiotu, może dodatkowo obejmować stwierdzenie zgodności przyrządu pomiarowego ze wskazanymi przez ten podmiot wymaganiami lub specyfikacjami. Z wzorcowaniem przyrządu pomiarowego związana jest nierozdzielnie niepewność pomiaru. Wzorcowanie przyrządów pomiarowych za pomocą wzorców mających powiązanie z wzorcami państwowymi lub międzynarodowymi stwarza gwarancję poprawnego działania wyposażenia pomiarowego stosowanego do pomiarów i kontroli wyrobu.

Sprawdzenie jest czynnością, mającą na celu stwierdzenie zgodności przyrządu pomiarowego z określonymi wymaganiami. Sprawdzenie przyrządu pomiarowego nie jest formą kontroli metrologicznej, choć może rozstrzygać, czy przyrząd spełnia wymagania, czy wymaga on naprawy, regulacji, włączenia bądź wycofania z eksploatacji. Stanowi ono pierwszy etap zarówno legalizacji, jak i uwierzytelnienia

Wypożyczenie pomiarowe stosowane do wzorcowań, badań i kontroli, mające istotny wpływ na niepewność pomiaru, związaną z wynikami tych działań, powinno być wzorcowane przez akredytowane laboratoria pomiarowe (wzorcujące) lub przez krajowy instytut metrologiczny (w Polsce – Główny Urząd Miar). Wzorce odniesienia akredytowanych laboratoriów wzorcujących powinny być kalibrowane w Głównym Urzędzie

Miar lub akredytowanych laboratoriach pomiarowych o odpowiedniej najlepszej możliwości pomiarowej.[9]

W momencie przystąpienia Polski do Unii Europejskiej rząd polski musiał ujednolicić normy polskie z unormowaniami unijnymi. Dlatego też wiele firm w ciągu ostatnich lat podejmowało działania, mające na celu dostosowanie przepisów wewnętrznych, zgodnie z kompatybilnymi unormowaniami unijnymi.

3. NORMY MIĘDZYNARODOWE

Każde przedsiębiorstwo, chcąc nawiązać z kontrahentami długoletnią i trwałą współpracę, buduje systemy zarządzania jakością. Najbardziej rozpowszechnionymi międzynarodowymi normami, omawiającymi zagadnienia systemu zarządzania jakością są normy ISO serii 9000. Normy te określają wymagania dotyczące między innymi nadzorowania wyposażenia do monitorowania i pomiarów. Norma ISO 9001:2000 w punkcie 7.6 przedstawia to w następujący sposób: Tam gdzie niezbędne jest zapewnienie wiarygodnych wyników wyposażenie pomiarowe należy:

- wzorcować lub sprawdzać w wyspecyfikowanych odstępach czasu lub przed użyciem w odniesieniu do wzorców jednostek miary mających powiązanie z międzynarodowymi lub państwowymi wzorcami jednostek miary; jeżeli nie ma takich wzorców, należy prowadzić zapisy dotyczące zastosowanej podstawy wzorcowania lub sprawdzania,
- adiustować lub ponownie adiustować, jeżeli jest to niezbędne,
- zidentyfikować w celu umożliwienia określenia statusu wzorcowania,
- zabezpieczyć przed adiustacjami, które mogłyby unieważnić wyniki pomiaru,
- chronić przed uszkodzeniem i pogorszeniem stanu podczas przemieszczania, utrzymywania przechowywania.[6]

Jednakże skupienie się tylko i wyłącznie na wspomnianych normach ISO serii 9000 nie w każdej gałęzi przemysłu przyniesie pożądany efekt. Zdarza się, że dla zapewnienia klientowi wymaganej przez niego jakości konieczne jest wprowadzenie także innych norm. Tak na przykład producenci samochodów uznali, że systemy jakości, zgodne z międzynarodowymi standardami norm ISO serii 9000, są niewystarczające. Sugerowano, że system ten nie jest w pełni wyspecyfikowany dla przemysłu motoryzacyjnego, tak, aby w pełni zapewnić zaplecze zaopatrzeniowe o oczekiwanej jakości. Skłoniło to Wielką Trójkę, czyli firmy Daimler-Chrysler, Ford i General Motors, do opracowania systemu QS-9000 (Quality System Requirements). Nowy system QS-9000 opublikowano w 1994 r., a następnie zmieniono w marcu 1998 r. System ten oparty został na normie ISO 9001:1994 i zawiera dodatkowe wymagania jakości, wprowadzone i oczekiwane przez Wielką Trójkę. Do

najważniejszych wymagań wykraczających poza normy ISO 9001 możemy zaliczyć:

- wprowadzenie systemu nadzoru w fazie opracowania koncepcji, tworzenia prototypu oraz produkcji;
- obowiązek dokumentowania trendów w zakresie produkcyjności, opłacalności i skuteczności oraz wykorzystywania uzyskanych informacji w tworzeniu ogólnych cech przedsiębiorstwa;
- wprowadzenie pisemnych procedur pozwalających na ocenę poziomu satysfakcji klientów oraz analizowanie danych w kontekście działań konkurencji;
- sporządzenie udokumentowanych instrukcji nadzorowania procesu oraz instrukcji operacyjnych dla wszystkich pracowników odpowiedzialnych za przebieg procesów;
- przyjęcie zasady „zero defektów” jako benchmarku przy określeniu metod i kryteriów badań próby [2].

Wymagania normy QS 9000 dotyczą wszystkich kooperantów, którzy zaopatrują Wielką Trójkę w materiały produkcyjne, części zamienne bądź świadczą dla nich usługi. W celu zapoznania dostawców z wymaganiami systemu jakości QS 9000 w przemyśle motoryzacyjnym twórcy opracowali kilka podręczników do których należą:

- QSA – Quality System Assessments – sterowanie systemem jakości;
- APQP – Advanced Product Quality Planning – planowanie jakości wyrobu;
- PPAP – Production Part Approval Process – proces zatwierdzania części do produkcji;
- FMEA – Potential Failure Mode And Effects Analysis – analiza przyczyn i skutków wad;
- SPC – Statistic Process Control – statystyczne sterowanie procesem;
- MSA – Measurement System Analysis – analiza systemów pomiarowych [1].

Wraz ze zmianami w normach ISO serii 9000 i po wydaniu nowej jej wersji w 2000 roku, dostosowując swoje wymagania z QS 9000, Międzynarodowy Zespół Operacyjny Przemysłu Motoryzacyjnego (International Automotive Task Force, IATF) opracował specyfikację techniczną ISO/TS 16949:2002. System zarządzania jakością – szczególne wymagania dotyczące zastosowania ISO 9001:2000 w przemyśle motoryzacyjnym. Specyfikacja ta została opracowana w celu: ujednolicenia istniejących na świecie norm systemów jakości w branży motoryzacyjnej w obrębie globalnego przemysłu motoryzacyjnego oraz wyeliminowania potrzeby wielokrotnych certyfikacji dla spełnienia wymagań klienta.

W normie tej dozwolone jest wyłączenie dotyczące punktu 7.3 (projektowanie i rozwój), w przypadku gdy przedsiębiorstwo nie ponosi odpowiedzialności z danej dziedziny.

Na rynku motoryzacyjnym uznawane są także inne wymagania jakościowe, takie jak EAQF (Francja), VDA-6.1 (Niemcy) czy AVSQ (Włochy), chociaż ich powszechność jest niewspółmierna wobec QS-9000.

Podobnie można wskazać normę AS9000/ D1-9000 na rynku dostaw dla przemysłu lotniczego, TL 9000 dla dostawców dla telekomunikacji oraz zamierzenia w podobnym zakresie wobec dostawców w wielu innych gałęziach przemysłu.

Mając za cel wysoką jakość wyrobów i gwarancję prawidłowych wyników firmy poprzez skuteczny system zarządzania pomiarami gwarantują swoim klientom, że zarówno wyposażenie jak i procesy pomiarowe przystosowane są do wyznaczonych im zastosowań. Norma ISO 10012:2003 „System zarządzania pomiarami. Wymagania dotyczące procesów pomiarowych i wyposażenia pomiarowego”, zawiera zarówno wymagania i wytyczne dotyczące zarządzania procesami pomiarowymi i potwierdzeniem metrologicznym wyposażenia pomiarowego, używanego do wspomagania oraz wykazywania zgodności z wymaganiami metrologicznymi [7].

Duże koncerny motoryzacyjne, posiadające własne laboratoria, badające ich produkty i wzorcuje aparaturę pomiarową, powinny rozważyć wdrożenie systemu zarządzania jakością w oparciu standard ISO 17025 bądź poddać akredytacji wskazane metody badawcze/wzorcuje. Norma ISO/IEC 17025:2005 Ogólne wymagania dotyczące kompetencji laboratoriów badawczych i wzorcujących [8] jest aktualizacją dotychczasowego standardu zarządzania jakością laboratoriów badawczych i wzorcujących. Norma ISO/IEC 17025:2005, zastąpiła wydanie z roku 1999 (polska wersja PN-EN ISO/IEC 17025: 2001), które stanowiło podstawowe kryterium akredytacji laboratoriów przez Polskie Centrum Akredytacji (PCA). Decyzją ILAC (International Laboratory Accreditation Cooperation) 14 maja 2007 kończy się okres przejściowy, w którym laboratoria posiadające akredytację zobowiązane są do spełnienia wymagań tejże normy.

4. ZWIĄZEK ZDOLNOŚCI JAKOŚCIOWEJ PROCESU POMIAROWEGO Z WYMAGANIAMI NORMATYWNYMI

Celem nadrzędnym dozoru sprzętu pomiarowo-kontrolnego jest zapewnienie wymaganej jakości systemu pomiarowego wykorzystywanego do kontroli procesów produkcyjnych. Zgodnie z normą EN ISO 10012:2003 [7] procesy pomiarowe, które są częścią systemu zarządzania pomiarami powinny być planowane, walidowane, wdrażane, dokumentowane i kontrolowane. Pełna specyfikacja każdego procesu pomiarowego powinna zawierać identyfikację całego istotnego wyposażenia, procedur pomiarowych, oprogramowania pomiarowego, warunków stosowania, możliwości personelu i wszystkich innych czynników wpływających na wiarygodność wyniku pomiaru. Aby móc spełnić wymagania metrologiczne określone przez klienta, organizacje i przepisy prawne, należy tak

zaprojektować wszelkie procesy pomiarowe, aby uniknąć możliwości wystąpienia błędnych wyników pomiarów, umożliwić szybkie wykrycie występujących nieprawidłowości i wprowadzić szybkie działania korygujące. Należy zidentyfikować i określić właściwości wymagane w danych procesach pomiarowych, takie jak: niepewność pomiaru, stabilność, powtarzalność, odtwarzalność czy błąd graniczny dopuszczalny. Normy międzynarodowe wymagają stałej analizy i doskonalenia systemu zarządzania pomiarami, np. normy serii 9000 wymagają auditowania i monitorowania procesów. Audyty zarządzania pomiarami mogą być przeprowadzane zgodnie z normą EN ISO 10012:2003 [7], ale również jako część auditów zarządzania organizacją zgodnie z wytycznymi podanymi w normie EN ISO 9001:2000 [6], jak również w normie PN-EN ISO 19011:2003 Wytyczne dotyczące auditowania systemów zarządzania jakością i/lub zarządzania środowiskowego.

Do czynników wpływających na dobór sprzętu pomiarowego uwzględniającego niepewność pomiaru można zaliczyć:

- niepełną definicję wielkości mierzonej;
- niedoskonałą realizację definicji wielkości mierzonej;
- niepełną znajomość oddziaływań otoczenia na pomiar albo niedoskonały pomiar warunków otoczenia;
- subiektywne błędy w odczytywaniu wskazań przyrządów analogowych;
- skończoną rozdzielczość albo próg pobudliwości przyrządu;
- niedokładne wartości przypisane wzorcom i materiałom odniesienia;
- niedokładne wartości stałych i innych parametrów otrzymanych ze źródeł zewnętrznych do pomiaru, a używanych w procedurach przetwarzania danych;
- przybliżenia i założenia upraszczające tkwiące w metodzie i procedurze pomiarowej;
- zmiany w powtarzalnych obserwacjach wielkości mierzonej w pozornie identycznych warunkach.

Wymienione wyżej przyczyny nie muszą być od siebie niezależne. Nierozpoznane oddziaływanie systematyczne nie może być wzięte pod uwagę przy szacowaniu niepewności wyniku pomiaru, ale wchodzi w skład jego błędu [4].

Ważnym zadaniem jest przeprowadzenie analizy systemu pomiarowego w celu zrozumienia, w jaki sposób błąd pomiaru wpływa na wynik badania. Narzucenie przez określone normy jakości wymogu dozoru systemu pomiarowego a przez to zwrócenie szczególnej uwagi na dokładne określenie wyniku pomiaru pozwala nam na uniknięcie wielu braków, niepotrzebnych kosztów usunięcia braków poprzez badanie właściwości opisanych poprzez charakterystyki techniczne określające przydatność i zakres stosowania systemu i charakterystyki

statystyczne mówiące nam o jego wiarygodności pozwalają nam na szybkie wykrycie nieprawidłowości i szybkie wprowadzenie działań korygujących. Do określenia przydatności systemu pomiarowego wykorzystywane są wskaźniki zdolności jakościowej procesu lub maszyn (C_p , C_{pk}), które są uwzględniane do doboru maszyn i urządzeń w fazie planowania produkcji. Najczęściej wykorzystywany jest wskaźnik %R&R – wskaźnik powtarzalności i odtwarzalności procesu pomiarowego.

Z analizy czynników wpływających na wynikową niepewność pomiarów – zdolność systemu pomiarowego wynika, że można je podzielić na dwie podstawowe grupy. Jedna związana jest z normatywnymi wskaźnikami wynikającymi z klasy przyrządu pomiarowego (dokładność, liniowość itp.). Druga dotyczy pozostałych czynników wpływających m.in. na odtwarzalność i powtarzalność wyników pomiarów (warunki pomiaru, obserwator, edycja i opracowanie wyników). Wynika z tego, że dobór sprzętu pomiarowo kontrolnego uwzględniającego standardy narzucone przez przepisy obligatoryjne jest czynnikiem, który w zasadniczy sposób wpływa na końcową niepewność pomiarów i zdolność systemu pomiarowego tak, aby spełnione były wymagania zawarte w planowanych metodach nadzoru i sterowania procesami produkcyjnymi.

5. KORZYŚCI

Używanie przez kontrahentów, którzy pochodzą z różnych krajów, jednolitych międzynarodowych norm np. serii ISO, zapobiega różnorodnej interpretacji przepisów zawartych w tych normach. Wybierając kontrahenta z kraju zrzeszonego w Europejskim Komitecie Normalizacyjnym, przedsiębiorstwo jest pewne, że reprezentuje ono wymaganą jakość swojej produkcji, a gospodarka środkami kontrolno-pomiarowymi daje pewność, że pomiary wykonywane są prawidłowo. Należy pamiętać, że zarówno sprawdzanie czy wzorcowanie tych przyrządów, jak i same procedury pomiaru i wyznaczania niepewności, są przeprowadzane w jednakowy sposób we wszystkich krajach unijnych. Biorąc pod uwagę wyspecjalizowane branże, jak np. branża motoryzacyjna, znalezienie się w grupie kontrahentów należących do takiej organizacji jak IATF, którą tworzy międzynarodowa grupa producentów pojazdów: BMW Group, DaimlerChrysler, Fiat, Ford Motor Company, General Motors Corporation, PSA Peugeot-Citroen, Renault i Volkswagen, oraz krajowe stowarzyszenia handlowe AIAG (Ameryka), VDA (Niemcy), SMMT (Wielka Brytania), ANFIA (Włochy) i FIEV (Francja), daje możliwość uzyskania wieloletnich kontraktów z dużymi koncernami. Zapewnia także ciągłość produkcji firmom, które posiadają certyfikaty i przestrzegają wymagań, uznanych przez wyższe organizacje.

6. PODSUMOWANIE

Po wejściu Polski do Unii Europejskiej krajowa gospodarka w szybkim tempie musiała dopasować się do wymogów rynku. O sukcesie rynkowym w nowych warunkach decyduje wiele czynników a wśród nich jakość. Normy serii ISO stały się podstawą do budowania systemów zarządzania jakością, które to nie tylko dają możliwość uporządkowania procesów zachodzących w przedsiębiorstwie, ale przede wszystkim pozwalają na efektywne gospodarowanie posiadanymi zasobami. Gospodarka zasobami wymaga pomiarów i kontroli we wszystkich obszarach działalności i procesach wytwarzania, zarówno zasobów materialnych, personelu, dokumentacji, optymalizacji potrzeb w zakresie planowania i zakupów itp., dlatego też każde przedsiębiorstwo, które w praktyce zarządza jakością, musi dokumentować procesy pomiarowe, dokonywać ich szybkiej analizy, musi więc wybrać odpowiednie przepisy i stworzyć na ich podstawie system nadzoru nad sprzętem pomiarowo-kontrolnym, zapewniający pożądaną dokładność pomiarów, utrzymanie wymaganej jakości wyrobów i usług w celu zapewnienia satysfakcji klienta.

LITERATURA

- [1] Gazdecki A. „Analiza MSA (Measurement System Analysis) na przykładzie przemysłu motoryzacyjnego”, IV Sympozjum Klubu Polskie Forum ISO 9000 Kielce 2000.
- [2] Wiśniewski S. „Systemy zarządzania jakością w branży motoryzacyjnej” Zarządzanie jakością 2/2005 str. 23-27.
- [3] Międzynarodowy słownik podstawowych i ogólnych terminów metrologii, Główny Urząd Miar, 1996.
- [4] Przewodnik. Wyrażanie niepewności pomiaru, GUM Warszawa 1999.
- [5] Ustawa Prawo o miarach z dnia 11 maja 2001 r. Dz. U. Nr 63 z 2001r.- zmiana z dnia 27 maja 2004 r. Dz. U. z 2004 r. Nr 243, poz. 2441 – tekst jednolity.
- [6] EN ISO 9001:2000 System Zarządzania Jakością. Wymagania.
- [7] EN ISO 10012:2003 System zarządzania pomiarami. Wymagania dotyczące procesów pomiarowych i wyposażenia pomiarowego.
- [8] EN ISO/IEC 17025:2005 Ogólne wymagania dotyczące kompetencji laboratoriów badawczych i wzorcujących.
- [9] www.pca.gov.pl /da-06.



Mgr inż. **ELIZA JARYSZ-KAMIŃSKA** – słuchacz studiów doktoranckich Politechniki Szczecińskiej na kierunku Budowa i Eksploatacja Maszyn.

WIELOSTANOWE PODEJŚCIE DO ANALIZY BEZPIECZEŃSTWA SYSTEMÓW

Krzysztof KOŁOWROCKI

Akademia Morska w Gdyni

81-225 Gdynia, Morska 81-87, e-mail: katmatkk@am.gdynia.pl

Streszczenie

Zaproponowane jest wielostanowe podejście do zdefiniowania podstawowych pojęć analizy i diagnozowania bezpieczeństwa systemów. Zdefiniowane są funkcja bezpieczeństwa systemu i funkcja ryzyka systemu. Zdefiniowane są podstawowe struktury wielostanowych systemów elementów z pogarszającymi się stanami bezpieczeństwa. Dla tych systemów wyznaczone są funkcje bezpieczeństwa. Proponowane podejście jest zastosowane do oceny funkcji bezpieczeństwa, funkcji ryzyka oraz innych charakterystyk bezpieczeństwa systemu transportu rurociągowego, kabla energetycznego oraz liny stalowej.

Słowa kluczowe: bezpieczeństwo systemów, systemy wielostanowe, diagnostyka.

MULTI-STATE APPROACH TO SYSTEM SAFETY ANALYSIS

Summary

A multi-state approach to defining basic notions of the system safety analysis and diagnosis is proposed. A system safety function and a system risk function are defined. Basic safety structures of multi-state systems of components with degrading safety states are defined. For these systems the multi-state safety functions are determined. The proposed approach is applied to the evaluation of safety functions, risk functions and other safety characteristics of a piping transportation system, an energetic cable and a steel rope.

Keywords: system safety, multi-state systems, diagnostics.

1. WPROWADZENIE

Z uwagi na bezpieczeństwo oraz efektywność eksploatacji systemów technicznych podczas analizy diagnostycznej wskazanym jest odejście od dwustanowego modelu ich bezpieczeństwa. Przyjęcie założenia, że są one wielostanowymi systemami starzejącymi się z powodu pogarszających się w czasie stanów technicznych ich elementów jest podstawą do bardziej dokładnej analizy i diagnozowania procesu eksploatacji tych systemów. Założenie to pozwala na wyróżnienie progowego stanu krytycznego systemu, którego przekroczenie jest niebezpieczne dla otoczenia lub też nie zapewnia odpowiedniego poziomu efektywności eksploatacji tego systemu. Wtedy podstawową charakterystyką bezpieczeństwa systemu staje się rozkład czasu do przekroczenia stanu progowego zwany funkcją ryzyka systemu. Rozkład ten jest ściśle wyznaczony przez wielostanową funkcję bezpieczeństwa systemu.

Wyznaczanie funkcji bezpieczeństwa oraz funkcji ryzyka dużych systemów na podstawie funkcji bezpieczeństwa ich elementów jest podstawowym zadaniem badawczym. Temu zagadnieniu poświęcona jest praca, z odniesieniem się do podstawowych struktur bezpieczeństwa systemów technicznych.

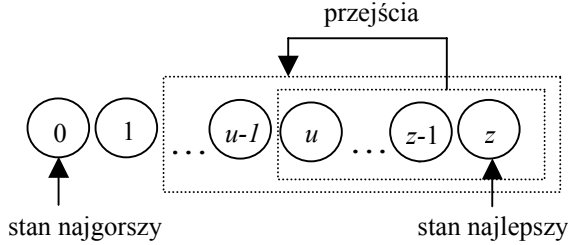
2. POJĘCIA PODSTAWOWE

W celu wprowadzenia wielostanowego podejścia do analizy bezpieczeństwa systemów elementów z pogarszającymi się stanami bezpieczeństwa, podobnie jak w przypadku badania niezawodności tego typu systemów wielostanowych ([1], [2], [3]), przyjmujemy, że:

- E_i , $i = 1, 2, \dots, n$, są elementami systemu,
- wszystkie rozważane elementy oraz system mają zbiór stanów bezpieczeństwa $\{0, 1, \dots, z\}$, $z \geq 1$,
- stany są uporządkowane, 0 jest stanem najgorszym natomiast stan z jest najlepszym,
- $T_i(u)$ są niezależnymi zmiennymi losowymi reprezentującymi czasy przebywania elementów E_i w podzbiórze stanów $\{u, u+1, \dots, z\}$, podczas gdy elementy te w chwili $t = 0$ znajdowały się w stanie z ,
- $T(u)$ jest zmienną losową reprezentującą czas przebywania systemu w podzbiórze stanów $\{u, u+1, \dots, z\}$, podczas gdy w chwili $t = 0$ system ten znajdował w stanie z ,
- stany bezpieczeństwa systemu oraz elementów pogarszają się wraz z upływem czasu t ,
- $E_i(t)$ jest stanem elementu E_i w chwili t , $t \in (-\infty, \infty)$, podczas gdy element ten w chwili $t = 0$ znajdował się w stanie z ,

- $S(t)$ jest stanem systemu w chwili t , $t \in (-\infty, \infty)$, podczas gdy system w chwili $t = 0$ znajdował się w stanie z .

Powyższe założenia oznaczają, że stany bezpieczeństwa elementów i systemu mogą zmieniać się w czasie tylko od lepszych do gorszych. Sposób, w jaki zmieniają się stany bezpieczeństwa elementów i systemu zilustrowany jest na rys. 1.



Rys. 1. Zmiany podzbiorów stanów bezpieczeństwa

Definicja 1. Wektor

$$s_i(t, \cdot) = [s_i(t, 0), s_i(t, 1), \dots, s_i(t, z)]$$

dla $t \in (-\infty, \infty)$, $i = 1, 2, \dots, n$, gdzie

$$s_i(t, u) = P(E_i(t) \geq u \mid E_i(0) = z) = P(T_i(u) > t) \quad (1)$$

dla $t \in (-\infty, \infty)$, $u = 0, 1, \dots, z$, jest prawdopodobieństwem tego, że element E_i w chwili t , $t \in (-\infty, \infty)$, znajduje się w jednym ze stanów podzbioru $\{u, u+1, \dots, z\}$, podczas gdy w chwili $t = 0$ znajdował się w stanie z , nazywamy wielostanową funkcją bezpieczeństwa elementu E_i .

Przy tej definicji oczywista jest następująca własność składowych wielostanowych funkcji bezpieczeństwa elementów

$$s_i(t, 0) \geq s_i(t, 1) \geq \dots \geq s_i(t, z)$$

dla $t \in (-\infty, \infty)$, $i = 1, 2, \dots, n$.

Jeśli

$$p_i(t) = [p_i(t, 0), p_i(t, 1), \dots, p_i(t, z)]$$

dla $t \in (-\infty, \infty)$, $i = 1, 2, \dots, n$,

gdzie

$$p_i(t, u) = P(E_i(t) = u \mid E_i(0) = z)$$

dla $t \in (-\infty, \infty)$, $u = 0, 1, \dots, z$,

jest prawdopodobieństwem tego, że element E_i w chwili t znajduje się w stanie u , podczas gdy w chwili $t = 0$ znajdował się w stanie z , to wobec (1)

$$s_i(t, 0) = 1, s_i(t, z) = p_i(t, z) \quad (2)$$

dla $t \in (-\infty, \infty)$, $i = 1, 2, \dots, n$, oraz

$$p_i(t, u) = s_i(t, u) - s_i(t, u+1) \quad (3)$$

dla $u = 0, 1, \dots, z-1$, $t \in (-\infty, \infty)$, $i = 1, 2, \dots, n$,

a ponadto

$$m_i(u) = \int_0^\infty s_i(t, u) dt \quad (4)$$

dla $u = 1, 2, \dots, z$, $i = 1, 2, \dots, n$,

jest wartością oczekiwaną (średnią) czasu przebywania elementu E_i w podzbiorze stanów $\{u, u+1, \dots, z\}$,

$$\sigma_i(u) = \sqrt{n_i(u) - [m_i(u)]^2} \quad (5)$$

dla $u = 1, 2, \dots, z$, $i = 1, 2, \dots, n$,

gdzie

$$n_i(u) = 2 \int_0^\infty t s_i(t, u) dt \quad (6)$$

dla $u = 1, 2, \dots, z$, $i = 1, 2, \dots, n$,

jest odchyleniem standardowym czasu przebywania elementu E_i w podzbiorze stanów $\{u, u+1, \dots, z\}$ oraz

$$\bar{m}_i(u) = \int_0^\infty p_i(t, u) dt, u = 1, 2, \dots, z, i = 1, 2, \dots, n, \quad (7)$$

jest wartością oczekiwaną czasu przebywania elementu E_i w stanie u , w przypadku gdy całki określone wzorami (4), (6) oraz (7) istnieją. Wtedy, zgodnie z (2), (3), (4) oraz (7), mamy

$$\bar{m}_i(u) = \bar{m}_i(u) - \bar{m}_i(u+1), u = 1, 2, \dots, z-1,$$

$$\bar{m}_i(z) = \bar{m}_i(z), i = 1, 2, \dots, n. \quad (8)$$

Definicja 2. Wektor

$$s_n(t, \cdot) = [s_n(t, 0), s_n(t, 1), \dots, s_n(t, z)], t \in (-\infty, \infty),$$

gdzie

$$s_n(t, u) = P(S(t) \geq u \mid S(0) = z) = P(T(u) > t) \quad (9)$$

dla $t \in (-\infty, \infty)$, $u = 0, 1, \dots, z$,

jest prawdopodobieństwem tego, że system w chwili t , $t \in (-\infty, \infty)$, znajduje się w podzbiorze stanów $\{u, u+1, \dots, z\}$, podczas gdy w chwili $t = 0$ znajdował się w stanie z , nazywamy wielostanową funkcją bezpieczeństwa systemu.

Przy tej definicji prawdziwa jest następująca własność składowych wielostanowej funkcji bezpieczeństwa systemu

$$s_n(t, 0) \geq s_n(t, 1) \geq \dots \geq s_n(t, z), t \in (-\infty, \infty).$$

Jeśli

$$p(t) = [p(t,0), p(t,1), \dots, p(t,z)], t \in (-\infty, \infty),$$

gdzie

$$p(t,u) = P(S(t) = u \mid S(0) = z) \quad (10)$$

dla $t \in (-\infty, \infty)$, $u = 0, 1, \dots, z$,

jest prawdopodobieństwem tego, że system w chwili t znajduje się w stanie u , podczas gdy w chwili $t = 0$ znajdował się w stanie z , to wobec (9) i (10) mamy

$$s_n(t,0) = 1, s_n(t,z) = p(t,z), t \in (-\infty, \infty), \quad (11)$$

oraz

$$p(t,u) = s_n(t,u) - s_n(t,u+1) \quad (12)$$

dla $u = 0, 1, \dots, z-1$, $t \in (-\infty, \infty)$.

Ponadto

$$m(u) = \int_0^\infty s_n(t,u) dt, u = 1, 2, \dots, z, \quad (13)$$

jest średnim czasem przebywania systemu w podzbiorze stanów $\{u, u+1, \dots, z\}$,

$$\sigma(u) = \sqrt{n(u) - [m(u)]^2}, u = 1, 2, \dots, z, \quad (14)$$

gdzie

$$n(u) = 2 \int_0^\infty t s_n(t,u) dt, u = 1, 2, \dots, z, \quad (15)$$

jest odchyleniem standardowym czasu przebywania systemu w podzbiorze stanów $\{u, u+1, \dots, z\}$ oraz

$$\bar{m}(u) = \int_0^\infty p(t,u) dt, u = 1, 2, \dots, z, \quad (16)$$

jest średnim czasem przebywania systemu w stanie u , w przypadku gdy całki (13), (15) i (16) istnieją. Wtedy, zgodnie z (11), (12), (13) i (16), mamy

$$\begin{aligned} \bar{m}(u) &= \bar{m}(u) - \bar{m}(u+1), u = 1, 2, \dots, z-1, \\ \bar{m}(z) &= \bar{m}(z). \end{aligned} \quad (17)$$

Definicja 3. Prawdopodobieństwo

$$r(t) = P(S(t) < r \mid S(0) = z) = P(T(r) \leq t), t \in (-\infty, \infty),$$

tego, że system w chwili t znajduje się w podzbiorze stanów bezpieczeństwa gorszych niż stan krytyczny r , $r \in \{1, \dots, z\}$, podczas gdy w chwili $t = 0$ znajdował

się w stanie z , nazywamy funkcją ryzyka systemu lub krótko ryzykiem.

Przy tej definicji, uwzględniając (9), mamy

$$r(t) = 1 - P(S(t) \geq r \mid S(0) = z) = 1 - s_n(t,r) \quad (18)$$

dla $t \in (-\infty, \infty)$.

Ponadto, jeśli τ jest chwilą, w której ryzyko przekroczy pewien dopuszczalny poziom δ , $\delta \in (0,1)$, to

$$\tau = r^{-1}(\delta), \quad (19)$$

gdzie $r^{-1}(t)$, jeśli istnieje, jest funkcją odwrotną funkcji ryzyka $r(t)$.

2. PODSTAWOWE STRUKTURY BEZPIECZEŃSTWA SYSTEMÓW

Definicja 4. System wielostanowy nazywamy systemem szeregowym, jeśli znajduje się w podzbiorze stanów $\{u, u+1, \dots, z\}$ wtedy i tylko wtedy, gdy wszystkie jego elementy znajdują się w tym podzbiorze stanów.

Wniosek 1. Czas $T(u)$ przebywania wielostanowego systemu szeregowego w podzbiorze stanów $\{u, u+1, \dots, z\}$ określony jest wzorem

$$T(u) = \min_{1 \leq i \leq n} \{T_i(u)\}, u = 1, 2, \dots, z.$$

Wniosek 2. Funkcja bezpieczeństwa wielostanowego systemu szeregowego określona jest wzorem

$$\bar{s}_n(t, \cdot) = [1, \bar{s}_n(t,1), \dots, \bar{s}_n(t,z)], \quad (20)$$

gdzie

$$\bar{s}_n(t,u) = \prod_{i=1}^n s_i(t,u), t \in (-\infty, \infty), u = 1, 2, \dots, z. \quad (21)$$

Definicja 5. Wielostanowy system szeregowy nazywamy jednorodnym, jeśli czasy $T_i(u)$ przebywania jego elementów w podzbiorach stanów $\{u, u+1, \dots, z\}$ mają identyczną dystrybuantę

$$F_i(t,u) = F(t,u), u = 1, 2, \dots, z, t \in (-\infty, \infty), i = 1, 2, \dots, n,$$

tzn. jeśli jego elementy E_i mają tę samą funkcję bezpieczeństwa, czyli gdy

$$s_i(t,u) = s(t,u) = 1 - F(t,u), t \in (-\infty, \infty),$$

dla $u = 1, 2, \dots, z$, $i = 1, 2, \dots, n$.

Wniosek 3. Funkcja bezpieczeństwa jednorodnego wielostanowego systemu szeregowego określona jest wzorem

$$\bar{s}_n(t, \cdot) = [1, \bar{s}_n(t, 1), \dots, \bar{s}_n(t, z)], \quad (22)$$

gdzie

$$\bar{s}_n(t, u) = [s(t, u)]^n \quad (23)$$

dla $t \in (-\infty, \infty)$, $u = 1, 2, \dots, z$.

Wniosek 4. Jeśli elementy jednorodnego wielostanowego systemu szeregowego mają wykładniczą funkcję bezpieczeństwa, tzn., gdy

$$s(t, \cdot) = [1, s(t, 1), \dots, s(t, z)],$$

gdzie

$$s(t, u) = 1 \text{ dla } t < 0,$$

$$s(t, u) = \exp[-\lambda(u)t] \text{ dla } t \geq 0, \lambda(u) > 0$$

$$\text{dla } u = 1, 2, \dots, z,$$

to jego funkcja bezpieczeństwa określona jest wzorem

$$\bar{s}_n(t, \cdot) = [1, \bar{s}_n(t, 1), \dots, \bar{s}_n(t, z)], \quad (24)$$

gdzie

$$\bar{s}_n(t, u) = 1, \quad t < 0,$$

$$\bar{s}_n(t, u) = \exp[-n\lambda(u)t] \text{ dla } t \geq 0, \quad (25)$$

$$\text{dla } u = 1, 2, \dots, z.$$

Definicja 6. System wielostanowy nazywamy systemem równoległym, jeśli znajduje się w podzbiorze stanów $\{u, u+1, \dots, z\}$ wtedy i tylko wtedy gdy co najmniej jeden z jego elementów znajduje się w tym podzbiorze stanów.

Wniosek 5. Czas $T(u)$ przebywania wielostanowego systemu równoległego w podzbiorze stanów $\{u, u+1, \dots, z\}$ określony jest wzorem

$$T(u) = \max_{1 \leq i \leq n} \{T_i(u)\}, \quad u = 1, 2, \dots, z.$$

Wniosek 6. Funkcja bezpieczeństwa wielostanowego systemu równoległego określona jest wzorem

$$s_n(t, \cdot) = [1, s_n(t, 1), \dots, s_n(t, z)], \quad (26)$$

gdzie

$$s_n(t, u) = 1 - \prod_{i=1}^n F_i(t, u), \quad t \in (-\infty, \infty), \quad u = 1, 2, \dots, z. \quad (27)$$

Definicja 7. Wielostanowy system równoległy nazywamy jednorodnym, jeśli czasy $T_i(u)$

przebywania jego elementów w podzbiorach stanów $\{u, u+1, \dots, z\}$ mają identyczną dystrybucję

$$F_i(t, u) = F(t, u), \quad u = 0, 1, \dots, z, \quad t \in (-\infty, \infty), \quad i = 1, 2, \dots, n,$$

tzn. jeśli jego elementy E_i mają tę samą funkcję bezpieczeństwa, czyli gdy

$$s_i(t, u) = s(t, u) = 1 - F(t, u), \quad t \in (-\infty, \infty), \\ \text{dla } u = 0, 1, \dots, z, \quad i = 1, 2, \dots, n.$$

Wniosek 7. Funkcja bezpieczeństwa jednorodnego wielostanowego systemu równoległego określona jest wzorem

$$s_n(t, \cdot) = [1, s_n(t, 1), \dots, s_n(t, z)], \quad (28)$$

gdzie

$$s_n(t, u) = 1 - [1 - s(t, u)]^n, \quad t \in (-\infty, \infty), \quad u = 1, 2, \dots, z. \quad (29)$$

Wniosek 8. Jeśli elementy jednorodnego wielostanowego systemu równoległego mają wykładniczą funkcję bezpieczeństwa, tzn., gdy

$$s(t, \cdot) = [1, s(t, 1), \dots, s(t, z)],$$

gdzie

$$s(t, u) = 1 \text{ dla } t < 0,$$

$$s(t, u) = \exp[-\lambda(u)t] \text{ dla } t \geq 0, \quad \lambda(u) > 0$$

$$\text{dla } u = 1, 2, \dots, z,$$

to jego funkcja bezpieczeństwa określona jest wzorem

$$s_n(t, \cdot) = [1, s_n(t, 1), \dots, s_n(t, z)], \quad (30)$$

gdzie

$$s_n(t, u) = 1 \text{ dla } t < 0,$$

$$s_n(t, u) = 1 - [1 - \exp[-\lambda(u)t]]^n \text{ dla } t \geq 0, \quad (31)$$

$$\text{dla } u = 1, 2, \dots, z.$$

Definicja 8. System wielostanowy nazywamy systemem progowym „ m z n ”, jeśli znajduje się w podzbiorze stanów $\{u, u+1, \dots, z\}$ wtedy i tylko wtedy gdy co najmniej m z jego n elementów znajduje się w tym podzbiorze stanów.

Wniosek 9. Czas $T(u)$ przebywania wielostanowego systemu progowego „ m z n ” w podzbiorze stanów $\{u, u+1, \dots, z\}$ określony jest wzorem

$$T(u) = T_{(n-m+1)}(u), \quad m = 1, 2, \dots, n, \quad u = 1, 2, \dots, z,$$

gdzie $T_{(n-m+1)}(u)$ jest m -tą maksymalną statystyką pozycyjną w ciągu zmiennych losowych $T_1(u), T_2(u), \dots, T_n(u)$.

Wniosek 10. Funkcja bezpieczeństwa wielostanowego systemu progowego określona jest wzorem

$$s_n^{(m)}(t, \cdot) = [1, s_n^{(m)}(t, 1), \dots, s_n^{(m)}(t, z)], \quad (32)$$

gdzie

$$s_n^{(m)}(t, u) = 1 - \sum_{\substack{r_1, r_2, \dots, r_n=0 \\ r_1+r_2+\dots+r_n \leq m-1}}^1 [s_i(t, u)]^{r_i} [1 - s_i(t, u)]^{1-r_i} \quad (33)$$

dla $t \in (-\infty, \infty)$, $u = 1, 2, \dots, z$,

lub wzorem

$$s_n^{(m)}(t, \cdot) = [1, s_n^{(m)}(t, 1), \dots, s_n^{(m)}(t, z)], \quad (34)$$

gdzie

$$s_n^{(m)}(t, u) = \sum_{\substack{r_1, r_2, \dots, r_n=0 \\ r_1+r_2+\dots+r_n \leq n-m}}^1 [1 - s_i(t, u)]^{r_i} [s_i(t, u)]^{1-r_i} \quad (35)$$

dla $t \in (-\infty, \infty)$, $u = 1, 2, \dots, z$.

Definicja 11. Wielostanowy system progowy nazywamy jednorodnym, jeśli czasy $T_i(u)$ przebywania jego elementów w podzbiorach stanów $\{u, u+1, \dots, z\}$ mają identyczną dystrybuantę

$$F_i(t, u) = F(t, u), \quad u = 0, 1, \dots, z, \quad t \in (-\infty, \infty), \quad i = 1, 2, \dots, n,$$

tzn. jeśli jego elementy E_i mają tę samą funkcję bezpieczeństwa, czyli gdy

$$s_i(t, u) = s(t, u) = 1 - F(t, u), \quad t \in (-\infty, \infty),$$

dla $u = 0, 1, \dots, z$, $i = 1, 2, \dots, n$.

Wniosek 8. Funkcja bezpieczeństwa jednorodnego wielostanowego systemu progowego określona jest wzorem

$$s_n^{(m)}(t, \cdot) = [1, s_n^{(m)}(t, 1), \dots, s_n^{(m)}(t, z)], \quad (36)$$

gdzie

$$s_n^{(m)}(t, u) = 1 - \sum_{i=0}^{m-1} \binom{n}{i} [s(t, u)]^i [1 - s(t, u)]^{n-i} \quad (37)$$

dla $t \in (-\infty, \infty)$, $u = 1, 2, \dots, z$,

lub wzorem

$$s_n^{(m)}(t, \cdot) = [1, s_n^{(m)}(t, 1), \dots, s_n^{(m)}(t, z)], \quad (38)$$

gdzie

$$s_n^{(m)}(t, u) = \sum_{i=0}^{n-m} \binom{n}{i} [1 - s(t, u)]^i [s(t, u)]^{n-i} \quad (39)$$

dla $t \in (-\infty, \infty)$, $u = 1, 2, \dots, z$.

Wniosek 12. Jeśli elementy jednorodnego wielostanowego systemu progowego mają wykładniczą funkcję bezpieczeństwa, tzn., gdy

$$s(t, \cdot) = [1, s(t, 1), \dots, s(t, z)],$$

gdzie

$$s(t, u) = 1 \quad \text{dla } t < 0,$$

$$s(t, u) = \exp[-\lambda(u)t] \quad \text{dla } t \geq 0, \quad \lambda(u) > 0$$

$$\text{dla } u = 1, 2, \dots, z,$$

to jego funkcja bezpieczeństwa określona jest wzorem

$$s_n^{(m)}(t, \cdot) = [1, s_n^{(m)}(t, 1), \dots, s_n^{(m)}(t, z)], \quad (40)$$

gdzie

$$s_n^{(m)}(t, u) = 1 \quad \text{dla } t < 0,$$

$$s_n^{(m)}(t, u) = 1 - \sum_{i=0}^{m-1} \binom{n}{i} \exp[-i\lambda(u)t] [1 - \exp[-\lambda(u)t]]^{n-i} \quad (41)$$

dla $t \geq 0$, $u = 1, 2, \dots, z$,

lub wzorem

$$s_n^{(m)}(t, \cdot) = [1, s_n^{(m)}(t, 1), \dots, s_n^{(m)}(t, z)], \quad (42)$$

gdzie

$$s_n^{(m)}(t, u) = 1 \quad \text{dla } t < 0,$$

$$s_n^{(m)}(t, u) = \sum_{i=0}^{n-m} \binom{n}{i} [1 - \exp[-\lambda(u)t]]^i \exp[-(n-i)\lambda(u)t] \quad (43)$$

dla $t \geq 0$, $u = 1, 2, \dots, z$.

Podobnie definiuje i analizuje się inne wielostanowe struktury bezpieczeństwa systemów.

4. ZASTOSOWANIA

Przykład 1. (system transportu rurociągowego)

System rurociągowy jest zbudowany z $n = 80$ pięciostanowych ($z = 4$) segmentów rur.

Segmenty rur występujące w systemie są czterech następujących typów:

20 segmentów posiada wielostanowe wykładnicze funkcje bezpieczeństwa

$$s_i(t, \cdot) = [1, s_i(t, 1), s_i(t, 2), s_i(t, 3), s_i(t, 4)], \\ i = 1, 2, \dots, 20,$$

o składowych

$$s_i(t, 1) = \exp[-0.01t], s_i(t, 2) = \exp[-0.0120t],$$

$$s_i(t, 3) = \exp[-0.015t], s_i(t, 4) = \exp[-0.025t], \\ i = 1, 2, \dots, 20,$$

20 segmentów posiada wielostanowe wykładnicze funkcje bezpieczeństwa

$$s_i(t, \cdot) = [1, s_i(t, 1), s_i(t, 2), s_i(t, 3), s_i(t, 4)], \\ i = 21, 22, \dots, 40,$$

o składowych

$$s_i(t, 1) = \exp[-0.018t], s_i(t, 2) = \exp[-0.019t],$$

$$s_i(t, 3) = \exp[-0.020t], s_i(t, 4) = \exp[-0.023t], \\ i = 21, 22, \dots, 40,$$

10 segmentów posiada wielostanowe weibullowskie funkcje bezpieczeństwa

$$s_i(t, \cdot) = [1, s_i(t, 1), s_i(t, 2), s_i(t, 3), s_i(t, 4)], \\ i = 41, 42, \dots, 50,$$

o składowych

$$s_i(t, 1) = \exp[-0.0005t^2], s_i(t, 2) = \exp[-0.0006t^2],$$

$$s_i(t, 3) = \exp[-0.0010t^2], s_i(t, 4) = \exp[-0.0015t^2], \\ i = 41, 42, \dots, 50,$$

30 segmentów posiada wielostanowe weibullowskie funkcje bezpieczeństwa

$$s_i(t, \cdot) = [1, s_i(t, 1), s_i(t, 2), s_i(t, 3), s_i(t, 4)], \\ i = 51, 52, \dots, 80,$$

o składowych

$$s_i(t, 1) = \exp[-0.00061t^3], s_i(t, 2) = \exp[-0.00062t^3],$$

$$s_i(t, 3) = \exp[-0.00064t^3], s_i(t, 4) = \exp[-0.00070t^3], \\ i = 51, 52, \dots, 80,$$

Średnie czasy przebywania w podzbiorach stanów, liczone w latach, przykładowo dla segmentów rur typu drugiego, po zastosowaniu wzoru (4), odpowiednio wynoszą:

$$m_2(1) = 1/0.018 = 55.56, m_2(2) = 1/0.019 = 52.63,$$

$$m_2(3) = 1/0.020 = 50.00, m_2(4) = 1/0.023 = 43.48,$$

natomiast ich średnie czasy przebywania w poszczególnych stanach, wobec (8), odpowiednio wynoszą:

$$\bar{m}_2(1) = 2.93, \bar{m}_2(2) = 2.63,$$

$$\bar{m}_2(3) = 6.52, \bar{m}_2(4) = 43.48.$$

Zgodnie z **Definicją 4**, rozważany system jest niejednorodnym wielostanowym systemem szeregowym. Zatem, wobec (20)–(21), otrzymujemy następujący wzór na funkcję bezpieczeństwa tego systemu

$$\bar{s}_{80}(t, \cdot) \\ = [1, \bar{s}_{80}(t, 1), \bar{s}_{80}(t, 2), \bar{s}_{80}(t, 3), \bar{s}_{80}(t, 4)],$$

gdzie

$$\bar{s}_{80}(t, 1) = \exp[-0.2t - 0.36t - 0.005t^2 - 0.00183t^3],$$

$$\bar{s}_{80}(t, 2) = \exp[-0.24t - 0.38t - 0.006t^2 - 0.00186t^3],$$

$$\bar{s}_{80}(t, 3) = \exp[-0.3t - 0.4t - 0.01t^2 - 0.00192t^3],$$

$$\bar{s}_{80}(t, 4) = \exp[-0.5t - 0.46t - 0.015t^2 - 0.0021t^3]$$

dla $t \geq 0$.

Oczekiwane średnie czasy przebywania systemu rurociągu w poszczególnych stanach, po zastosowaniu powyższego wyniku, wzoru (13) oraz przybliżonego całkowania, przyjmują następujące wartości:

$$m(1) \cong 1.79, m(2) \cong 1.61,$$

$$m(3) \cong 1.43, m(4) \cong 1.04,$$

a następnie, na podstawie wzoru (17), średnie czasy przebywania systemu w poszczególnych stanach przyjmują wartości:

$$\bar{m}(1) \cong 0.18, \bar{m}(2) \cong 0.18,$$

$$\bar{m}(3) \cong 0.39, \bar{m}(4) \cong 1.04.$$

Jeśli założymy, że krytycznym stanem bezpieczeństwa jest stan $r = 2$, to funkcja ryzyka systemu, na podstawie wzoru (18), przyjmuje postać

$$r(t) = 1 - \exp[-0.24t - 0.38t - 0.006t^2 - 0.00186t^3] \\ \text{dla } t \geq 0.$$

Stąd, po zastosowaniu wzoru (19), otrzymujemy, że

chwilą przekroczenia przez funkcję ryzyka systemu dozwolonego poziomu, np. $\delta = 0.05$, jest

$$\tau = r^{-1}(\delta) \cong 0.083 \text{ roku.}$$

Przykład 2. (system kabla energetycznego)

Przyjmijmy, że kabel energetyczny zbudowany jest z 36 identycznych drutów i może przewodzić prąd, gdy co najmniej jeden z drutów przewodzi prąd. Założenie to, wobec Definicji 7, oznacza, że jest to jednorodny system równoległy zbudowany z $n = 36$ elementów. Przyjmując dalej, że poszczególne druty są elementami czterostanowymi ($z = 3$) posiadającymi weibulłowskie funkcje bezpieczeństwa

$$s_i(t, \cdot) = [1, s_i(t, 1), s_i(t, 2), s_i(t, 3)], \quad i = 1, 2, \dots, 36,$$

których składowe

$$s_i(t, u) = \exp[-\beta(u)t^{\alpha(u)}] \text{ dla } t \geq 0, u = 1, 2, 3,$$

mają parametry

$$\alpha(u) = 2, \beta(u) = (7.07)^{2u-8}, u = 1, 2, 3,$$

po zastosowaniu **Wniosku 5**, otrzymujemy następujący wzór na funkcję bezpieczeństwa kabla

$$[s_{36}(t, 0), s_{36}(t, 1), s_{36}(t, 2), s_{36}(t, 3)]$$

$$= [1, 1 - [1 - \exp[-2t^{0.000008}]]^{36},$$

$$1 - [1 - \exp[-2t^{0.0004}]]^{36}, 1 - [1 - \exp[-2t^{0.02}]]^{36}]$$

dla $t \geq 0$.

Średnie wartości czasów przebywania kabla energetycznego w podzbiorach stanów, po zastosowaniu wzoru (13) oraz przybliżonego całkowania, przyjmują następujące wartości:

$$m(1) \cong 723, m(2) \cong 102, m(3) \cong 14.5,$$

$$\sigma(1) \cong 120, \sigma(2) \cong 17, \sigma(3) \cong 2.4,$$

podczas, gdy średnie czasy przebywania tego systemu w poszczególnych stanach, na podstawie wzoru (17), przyjmują wartości:

$$\bar{m}(1) \cong 621, \bar{m}(2) \cong 87.5, \bar{m}(3) \cong 14.5.$$

Jeśli krytycznym stanem jest stan $r = 2$, to funkcja ryzyka systemu, na podstawie wzoru (18), przyjmuje postać

$$r(t) = [1 - \exp[-2t^{0.0004}]]^{36} \text{ dla } t \geq 0.$$

Stąd, po zastosowaniu wzoru (19), otrzymujemy, że chwilą przekroczenia przez funkcję ryzyka systemu dozwolonego poziomu, np. $\delta = 0.05$, jest

$$\tau = [\log[-\frac{1}{2}\log(1 - \delta^{1/36})]]^{1/0.0004} = 80 \text{ miesięcy.}$$

Wartości funkcji ryzyka systemu kabla energetycznego są dane w Tablicy 1.

Tablica 1. Wartości funkcji ryzyka systemu kabla energetycznego

t	$r(t)$
0	0.000
60	0.000
70	0.017
80	0.048
90	0.276
100	0.514
110	0.731
120	0.856
130	0.934
140	0.968
150	0.985
160	0.993

Przykład 3. (system liny stalowej)

Rozważmy linę stalową składającą się z $n = 36$ czterostanowych ($z = 3$) identycznych splotek mających wielostanowe wykładnicze funkcje bezpieczeństwa

$$s_i(t, \cdot) = [1, s_i(t, 1), s_i(t, 2), s_i(t, 3)], \quad i = 1, 2, \dots, 36,$$

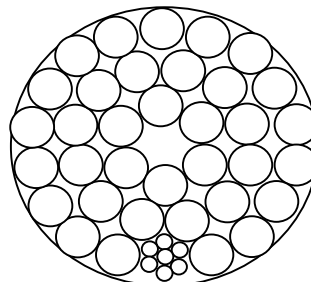
których składowe

$$s_i(t, u) = \exp[-\lambda(u)t] \text{ dla } t \geq 0, u = 1, 2, 3,$$

charakteryzują się następującymi intensywnościami przejść pomiędzy podzbiorami stanów

$$\lambda(u) = 0.2u/\text{rok}, u = 1, 2, 3.$$

Przekrój poprzeczny liny stalowej przedstawiony jest na rys. 2.



Rys. 2. Przekrój poprzeczny liny stalowej

Przyjmując, że lina znajduje się w podzbiorze stanów $\{u, u+1, \dots, z\}$, jeśli przynajmniej $m = 10$ spośród jej splotek znajduje się w tym podzbiorze stanów, zgodnie z Definicją 9, wnioskujemy, że jest ona czterostanowym jednorodnym systemem progowym "10 z 36". Wtedy, zgodnie ze wzorami (40)–(41), funkcja bezpieczeństwa liny dana jest wzorem

$$s_{36}^{(10)}(t, \cdot) = [1, s_{36}^{(10)}(t, 1), s_{36}^{(10)}(t, 2), s_{36}^{(10)}(t, 3)],$$

gdzie

$$s_{36}^{(10)}(t, u) = 1 \text{ dla } t < 0, u = 1, 2, 3,$$

oraz

$$s_{36}^{(10)}(t, 1) = 1 - \sum_{i=0}^9 \binom{36}{i} \exp[-i0.2t][1 - \exp[-0.2t]]^{36-i},$$

$$s_{36}^{(10)}(t, 2) = 1 - \sum_{i=0}^9 \binom{36}{i} \exp[-i0.4t][1 - \exp[-0.4t]]^{36-i},$$

$$s_{36}^{(10)}(t, 3) = 1 - \sum_{i=0}^9 \binom{36}{i} \exp[-i0.6t][1 - \exp[-0.6t]]^{36-i}$$

dla $t \geq 0$.

Średnie wartości czasów przebywania liny stalowej w podzbiorach stanów, po zastosowaniu wzoru (13) oraz przybliżonego całkowania, przyjmują następujące wartości:

$$m(1) \cong 6.66, m(2) \cong 3.33, m(3) \cong 2.22,$$

$$\sigma(1) \cong 1.62, \sigma(2) \cong 0.81, \sigma(3) \cong 0.54,$$

podczas, gdy średnie czasy przebywania tego systemu w poszczególnych stanach, na podstawie wzoru (17), przyjmują wartości:

$$\bar{m}(1) \cong 3.33, \bar{m}(2) \cong 1.11, \bar{m}(3) \cong 2.22.$$

Jeśli krytycznym stanem jest stan $r = 2$, to funkcja ryzyka systemu, na podstawie wzoru (18), przyjmuje postać

$$r(t) = \sum_{i=0}^9 \binom{36}{i} \exp[-i0.4t][1 - \exp[-0.4t]]^{36-i} \text{ dla } t \geq 0.$$

Stąd, po zastosowaniu wzoru (19), otrzymujemy, że chwilą przekroczenia przez funkcję ryzyka systemu dozwolonego poziomu, np. $\delta = 0.05$, jest

$$\tau \cong 2.07 \text{ roku.}$$

Wartości składowych wielostanowej funkcji bezpieczeństwa liny stalowej oraz jej funkcji ryzyka są dane w Tablicy 2.

Tablica 2. Wartości składowych wielostanowej funkcji bezpieczeństwa oraz funkcji ryzyka liny stalowej

t	$s_{36}^{(10)}(t, 1)$	$s_{36}^{(10)}(t, 2)$	$s_{36}^{(10)}(t, 3)$	$r(t)$
0.2	1.000	0.999	0.999	0.001
1.4	0.999	0.996	0.945	0.004
1.8	0.999	0.980	0.776	0.020
2.2	0.999	0.927	0.493	0.073
2.6	0.997	0.815	0.231	0.185
3.0	0.994	0.642	0.080	0.358
3.4	0.986	0.444	0.021	0.556
3.8	0.971	0.267	0.004	0.733
4.2	0.945	0.141	0.000	0.859
4.6	0.906	0.065	0.000	0.935
5.0	0.849	0.027	0.000	0.973
5.4	0.776	0.010	0.000	0.990
5.8	0.689	0.003	0.000	0.997
6.2	0.593	0.001	0.000	0.999
7.0	0.396	0.000	0.000	1.000

LITERATURA

- [1] Kołowrocki, K., *Reliability of Large Systems*, Elsevier: Amsterdam-Boston - Heidelberg - London - New York - Oxford - Paris - San Diego - San Francisco - Singapore - Sydney - Tokyo, 2004.
- [2] Soszynska, J., *Reliability of large series-parallel system in variable operation conditions*. Proc. European Safety and Reliability Conference, ESREL 2005, 27-30 June, 2005, Tri City, Poland.
- [3] *Advances in Safety and Reliability*, Edited by K. Kołowrocki, Volume 2, 1869-1876.
- [4] A. A. Balkema Publishers: Leiden - London - New York - Philadelphia - Singapore, 2005.
- [5] Xue J., Yang K., Dynamic reliability analysis of coherent multi-state systems. *IEEE Transactions on Reliability* 4, 44, 1995, 683-688.

Krzysztof KOŁOWROCKI jest profesorem



zwyczajnym oraz kierownikiem Katedry Matematyki na Wydziale Nawigacyjnym Akademii Morskiej w Gdyni. Jest także wiceprezesem Polskiego Towarzystwa Bezpieczeństwa i Niezawodności. Jego obszarem zainteresowań naukowych jest modelowanie bezpieczeństwa i niezawodności

złożonych systemów i procesów. W tym zakresie opublikował ponad 200 prac naukowych. Więcej danych można znaleźć na jego stronie internetowej <http://www.am.gdynia.pl/~katmatkk/>

METODA OPTYMALIZACJI PRZEGLĄDÓW OKRESOWYCH TRANSPORTOWYCH SYSTEMÓW NADZORU

Adam ROŚIŃSKI

Politechnika Warszawska, Wydział Transportu, Zakład Telekomunikacji w Transporcie
00-662 Warszawa, ul. Koszykowa 75
Tel. 022-6607036, email: adro@it.pw.edu.pl

Streszczenie

Transportowe systemy nadzoru są to systemy, których celem jest wykrywanie zagrożeń występujących w procesie transportowym. Zapewniają bezpieczeństwo ludziom, a także ochronę przewożonym towarom zarówno w obiektach stałych, jak i ruchomych. Zachowanie zdolności jest istotnym warunkiem ich poprawnej pracy. W referacie zaprezentowano metodę optymalizacji przeglądów okresowych, która uwzględnia wybrane własności niezawodnościowe (intensywność uszkodzeń) i eksploatacyjne (intensywność napraw, intensywności realizacji przeglądów). Kryterium optymalizacyjne maksymalizuje wartość wskaźnika gotowości.

Słowa kluczowe: eksploatacja, gotowość.

EXPLOITATION STRATEGIES OF MONITORING TRANSPORT SYSTEMS

Summary

Monitoring transport systems are the systems which task is detection of hazards occurring in the transport process. They assure the safety of people as well as protection of the carried goods in the stationary and mobile objects. Preservation of their ability is the essential condition of their correct operation. Paper presents method of optimization of the periodical surveys, which takes into consideration the chosen factors of reliability (failure rate) and exploitation (repair rate, intensity of survey realization). Optimization criterion is the maximization of the rate of availability.

Keywords: exploitation, availability.

1. WSTĘP

Głównym celem transportu jest przemieszczanie osób i ładunków. Proces ten powinien cechować się wysokim poziomem niezawodności i bezpieczeństwa. Dlatego też niezbędne jest stosowanie transportowych systemów nadzoru, których zadaniem jest zwiększenie bezpieczeństwa przewożonych osób i ładunków.

Transportowe systemy nadzoru są to systemy, których celem jest wykrywanie zagrożeń występujących w procesie transportowym (dla obiektów stacjonarnych i ruchomych). Są one coraz częściej stosowane w procesie transportowym, gdzie oprócz zapewnienia bezpieczeństwa ludziom (np. systemy nadzoru zainstalowane w obiektach stałych lotniska, dworcach kolejowych, itp.), zapewniają także ochronę przewożonym towarom zarówno w obiektach stałych (np. bazy logistyczne, terminale przeładunkowe, itp.), jak i ruchomych (np. w połączeniu z systemem GPS – monitorowanie stanu ładunku i trasy przejazdu danego pojazdu).

Transportowe systemy nadzoru pracują w zróżnicowanych warunkach eksploatacyjnych. Ich poprawne funkcjonowanie jest uzależnione nie tylko

od niezawodności poszczególnych części składowych tworzących system, ale także od przyjętych do realizacji przeglądów okresowych [1, 2, 4]. Maksymalizacja wskaźnika gotowości możliwa jest m.in. poprzez opracowanie metody optymalizacji przeglądów okresowych.

2. OPTYMALIZACJA PRZEGLĄDÓW OKRESOWYCH

Projektując systemy nadzoru można wpływać na wskaźniki niezawodności m.in. poprzez zastosowanie odpowiednich struktur niezawodnościowych [3]. W ten sposób korygujemy wartość intensywności uszkodzeń zaprojektowanego systemu. Pomimo takiego podejścia przy projektowaniu tego typu systemów, niezbędne jest także uwzględnienie wskaźnika gotowości, zgodnie ze wzorem :

$$K_g = \frac{T_m}{T_m + T_n} \quad (1)$$

gdzie: T_m - średni czas poprawnej pracy między uszkodzeniami,

T_n - średni czas naprawy.

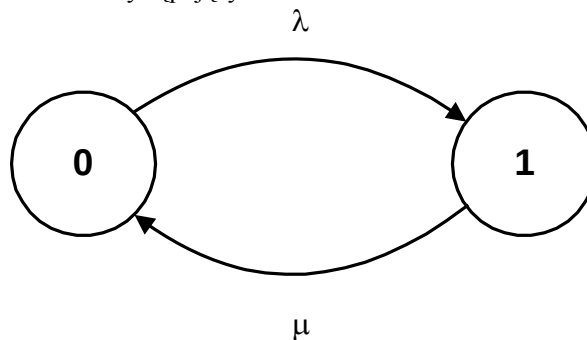
Z przedstawionych zależności wynika, że system może znajdować się w jednym z dwóch stanów (rys. 1):

- stan użytkowania (sprawności),
- stan naprawy.

Przedstawiony na rys. 1 graf przejść nie zawiera wszystkich możliwych i występujących

w rzeczywistości stanów. Dlatego należy go uzupełnić o następujące dwa stany (rys. 2):

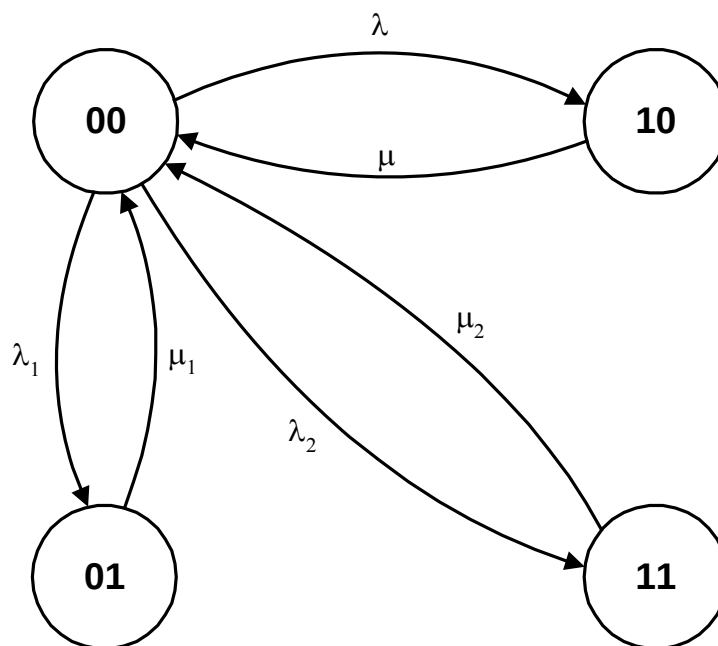
- stan 01 (podczas którego wykonuje się podstawowe czynności wymagane zakresem obsługi),
- stan 11 (podczas którego wykonuje się rozszerzony zakres czynności).



Rys. 1. Graf przejść między stanem użytkowania i naprawy

Oznaczenia na rys.: λ - intensywność uszkodzeń,

μ - intensywność napraw



Rys. 2. Graf przejść między stanem użytkowania, naprawy, przeglądu I i II rodzaju

Oznaczenia na rys.: λ - intensywność uszkodzeń,

μ - intensywność napraw,

λ_1 - intensywność przeglądów I rodzaju,

μ_1 - intensywność realizacji przeglądów I rodzaju,

λ_2 - intensywność przeglądów II rodzaju,

μ_2 - intensywność realizacji przeglądów II rodzaju

Dla grafu przejść przedstawionego na rys. 2 można zapisać następujące równania:

$$\begin{aligned}
 -\lambda \cdot P_{00} + \mu \cdot P_{10} - \lambda_1 \cdot P_{00} + \mu_1 \cdot P_{01} - \lambda_2 \cdot P_{00} + \mu_2 \cdot P_{11} &= 0 \\
 \lambda \cdot P_{00} - \mu \cdot P_{10} &= 0 \\
 \lambda_1 \cdot P_{00} - \mu_1 \cdot P_{01} &= 0 \\
 \lambda_2 \cdot P_{00} - \mu_2 \cdot P_{11} &= 0
 \end{aligned} \tag{2}$$

W zapisie macierzowym można przedstawić to następująco:

$$\begin{bmatrix}
 -(\lambda + \lambda_1 + \lambda_2) & \mu & \mu_1 & \mu_2 \\
 \lambda & -\mu & 0 & 0 \\
 \lambda_1 & 0 & -\mu_1 & 0 \\
 \lambda_2 & 0 & 0 & -\mu_2
 \end{bmatrix} \cdot \begin{bmatrix} P_{00} \\ P_{10} \\ P_{01} \\ P_{11} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{3}$$

Przekształcając otrzymujemy:

$$\begin{aligned}
 P_{10} &= \frac{\lambda}{\mu} \cdot P_{00} \\
 P_{01} &= \frac{\lambda_1}{\mu_1} \cdot P_{00} \\
 P_{11} &= \frac{\lambda_2}{\mu_2} \cdot P_{00}
 \end{aligned}$$

Oczywiście:

$$P_{00} + P_{10} + P_{01} + P_{11} = 1$$

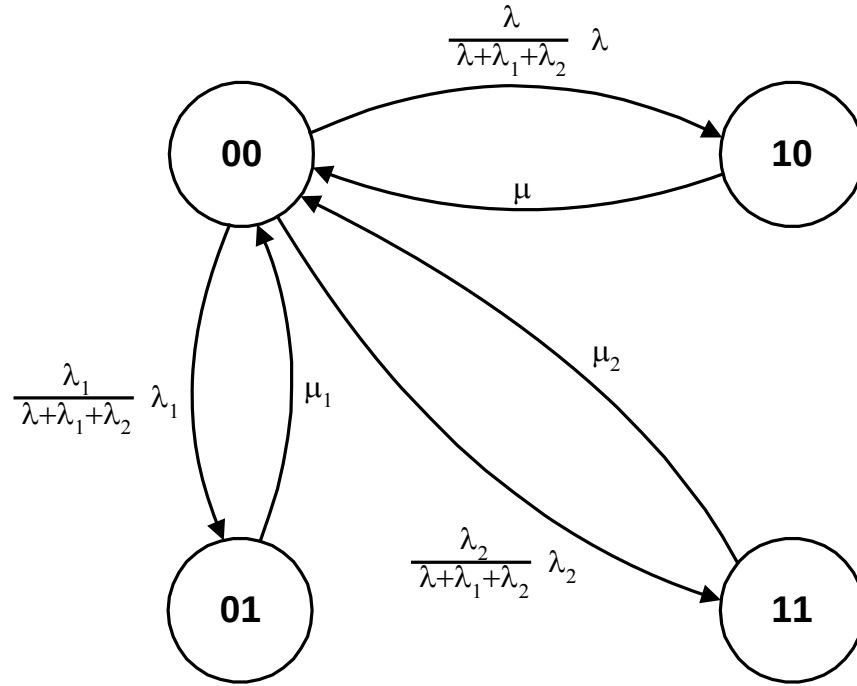
Zatem:

$$P_{00} \cdot \left(1 + \frac{\lambda}{\mu} + \frac{\lambda_1}{\mu_1} + \frac{\lambda_2}{\mu_2}\right) = 1 \tag{4}$$

$$P_{00} = \frac{1}{\left(1 + \frac{\lambda}{\mu} + \frac{\lambda_1}{\mu_1} + \frac{\lambda_2}{\mu_2}\right)} = \frac{\mu \cdot \mu_1 \cdot \mu_2}{\mu \cdot \mu_1 \cdot \mu_2 + \lambda \cdot \mu_1 \cdot \mu_2 + \lambda_1 \cdot \mu \cdot \mu_2 + \lambda_2 \cdot \mu \cdot \mu_1} \tag{5}$$

Wprowadźmy współczynnik, który uzależni wartości intensywności λ , λ_1 i λ_2 od siebie (jeśli jedna z tych wartości będzie wzrastać, to pozostałe powinny maleć). Współczynnikiem tym będzie

stosunek danej intensywności przejścia do sumy wszystkich intensywności naprawy, przeglądu I i II rodzaju. Graf przejść przedstawiony na rys. 2 będzie miał teraz następującą postać (rys. 3):



Rys. 3. Graf przejść między stanem użytkowania, naprawy, przeglądu I i II rodzaju (uwzględniono współczynnik korygujący)

Przeprowadzając rozważania tak jak poprzednio możemy zapisać:

$$\begin{aligned}
 & -\lambda \cdot \frac{\lambda}{\lambda + \lambda_1 + \lambda_2} \cdot P_{00} + \mu \cdot P_{10} - \lambda_1 \cdot \frac{\lambda_1}{\lambda + \lambda_1 + \lambda_2} \cdot P_{00} + \\
 & + \mu_1 \cdot P_{01} - \lambda_2 \cdot \frac{\lambda_2}{\lambda + \lambda_1 + \lambda_2} \cdot P_{00} + \mu_2 \cdot P_{11} = 0 \\
 & \lambda \cdot \frac{\lambda}{\lambda + \lambda_1 + \lambda_2} \cdot P_{00} - \mu \cdot P_{10} = 0 \\
 & \lambda_1 \cdot \frac{\lambda_1}{\lambda + \lambda_1 + \lambda_2} \cdot P_{00} - \mu_1 \cdot P_{01} = 0 \\
 & \lambda_2 \cdot \frac{\lambda_2}{\lambda + \lambda_1 + \lambda_2} \cdot P_{00} - \mu_2 \cdot P_{11} = 0
 \end{aligned} \tag{6}$$

W zapisie macierzowym można przedstawić to następująco:

$$\begin{bmatrix}
 -\left(\frac{\lambda}{\lambda + \lambda_1 + \lambda_2} \cdot \lambda + \frac{\lambda_1}{\lambda + \lambda_1 + \lambda_2} \cdot \lambda_1 + \frac{\lambda_2}{\lambda + \lambda_1 + \lambda_2} \cdot \lambda_2\right) & \mu & \mu_1 & \mu_2 \\
 \frac{\lambda}{\lambda + \lambda_1 + \lambda_2} \cdot \lambda & -\mu & 0 & 0 \\
 \frac{\lambda_1}{\lambda + \lambda_1 + \lambda_2} \cdot \lambda_1 & 0 & -\mu_1 & 0 \\
 \frac{\lambda_2}{\lambda + \lambda_1 + \lambda_2} \cdot \lambda_2 & 0 & 0 & -\mu_2
 \end{bmatrix} \cdot \begin{bmatrix} P_{00} \\ P_{10} \\ P_{01} \\ P_{11} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{7}$$

Przekształcając otrzymujemy:

$$P_{10} = \frac{\lambda}{\lambda + \lambda_1 + \lambda_2} \cdot \frac{\lambda}{\mu} \cdot P_{00}$$

$$P_{01} = \frac{\lambda_1}{\lambda + \lambda_1 + \lambda_2} \cdot \frac{\lambda_1}{\mu_1} \cdot P_{00}$$

$$P_{11} = \frac{\lambda_2}{\lambda + \lambda_1 + \lambda_2} \cdot \frac{\lambda_2}{\mu_2} \cdot P_{00}$$

Oczywiście:

$$P_{00} + P_{10} + P_{01} + P_{11} = 1,$$

zatem:

$$P_{00} \cdot \left(1 + \frac{\lambda}{\lambda + \lambda_1 + \lambda_2} \cdot \frac{\lambda}{\mu} + \frac{\lambda_1}{\lambda + \lambda_1 + \lambda_2} \cdot \frac{\lambda_1}{\mu_1} + \frac{\lambda_2}{\lambda + \lambda_1 + \lambda_2} \cdot \frac{\lambda_2}{\mu_2}\right) = 1 \quad (8)$$

$$P_{00} = \frac{1}{\left(1 + \frac{\lambda}{\lambda + \lambda_1 + \lambda_2} \cdot \frac{\lambda}{\mu} + \frac{\lambda_1}{\lambda + \lambda_1 + \lambda_2} \cdot \frac{\lambda_1}{\mu_1} + \frac{\lambda_2}{\lambda + \lambda_1 + \lambda_2} \cdot \frac{\lambda_2}{\mu_2}\right)} = \quad (9)$$

$$= \frac{(\lambda + \lambda_1 + \lambda_2) \cdot \mu \cdot \mu_1 \cdot \mu_2}{(\lambda + \lambda_1 + \lambda_2) \cdot \mu \cdot \mu_1 \cdot \mu_2 + \lambda^2 \cdot \mu_1 \cdot \mu_2 + \lambda_1^2 \cdot \mu \cdot \mu_2 + \lambda_2^2 \cdot \mu \cdot \mu_1}$$

Otrzymana zależność pozwala określić wpływ przyjętych do realizacji intensywności przeglądów I i II rodzaju na wskaźnik gotowości danego systemu (przy znanej intensywności uszkodzeń, intensywności realizacji przeglądów I i II rodzaju). Jeśli funkcja posiada maksimum, to wyznaczenie odpowiadającej mu wartości współrzędnych jakimi są intensywność przeglądu I i intensywność

przeglądu II rodzaju jest wskazane, ponieważ umożliwi zwiększenie wskaźnika gotowości. Wartości te byłyby wtedy wartościami optymalnymi, zapewniającymi maksymalizację wskaźnika gotowości.

Sprawdźmy zatem czy funkcja posiada maksimum. Obliczmy pochodne funkcji:

$$\frac{dP_{00}}{d\lambda_1} \text{ i } \frac{dP_{00}}{d\lambda_2}$$

$$\frac{dP_{00}}{d\lambda_1} = \frac{\mu \cdot \mu_1 \cdot \mu_2 \cdot [(\lambda + \lambda_1 + \lambda_2) \cdot \mu \cdot \mu_1 \cdot \mu_2 + \lambda^2 \cdot \mu_1 \cdot \mu_2 + \lambda_1^2 \cdot \mu \cdot \mu_2 + \lambda_2^2 \cdot \mu \cdot \mu_1] - (\lambda + \lambda_1 + \lambda_2) \cdot \mu \cdot \mu_1 \cdot \mu_2 \cdot (\mu \cdot \mu_1 \cdot \mu_2 + 2 \cdot \lambda_1 \cdot \mu \cdot \mu_2)}{[(\lambda + \lambda_1 + \lambda_2) \cdot \mu \cdot \mu_1 \cdot \mu_2 + \lambda^2 \cdot \mu_1 \cdot \mu_2 + \lambda_1^2 \cdot \mu \cdot \mu_2 + \lambda_2^2 \cdot \mu \cdot \mu_1]^2} \quad (10)$$

$$\frac{dP_{00}}{d\lambda_2} = \frac{\mu \cdot \mu_1 \cdot \mu_2 \cdot [(\lambda + \lambda_1 + \lambda_2) \cdot \mu \cdot \mu_1 \cdot \mu_2 + \lambda^2 \cdot \mu_1 \cdot \mu_2 + \lambda_1^2 \cdot \mu \cdot \mu_2 + \lambda_2^2 \cdot \mu \cdot \mu_1] - (\lambda + \lambda_1 + \lambda_2) \cdot \mu \cdot \mu_1 \cdot \mu_2 \cdot (\mu \cdot \mu_1 \cdot \mu_2 + 2 \cdot \lambda_2 \cdot \mu \cdot \mu_1)}{[(\lambda + \lambda_1 + \lambda_2) \cdot \mu \cdot \mu_1 \cdot \mu_2 + \lambda^2 \cdot \mu_1 \cdot \mu_2 + \lambda_1^2 \cdot \mu \cdot \mu_2 + \lambda_2^2 \cdot \mu \cdot \mu_1]^2} \quad (11)$$

Warunkiem koniecznym, aby funkcja $P_{00}(\lambda_1, \lambda_2)$ miała ekstremum w punkcie $P_{00}(\lambda_{1optym}, \lambda_{2optym})$ jest, aby pierwsze pochodne cząstkowe funkcji w tym punkcie były równe zero, tj.:

$$\begin{cases} \frac{dP_{00}}{d\lambda_1}(\lambda_{1optym}, \lambda_{2optym}) = 0 \\ \frac{dP_{00}}{d\lambda_2}(\lambda_{1optym}, \lambda_{2optym}) = 0 \end{cases} \quad (12)$$

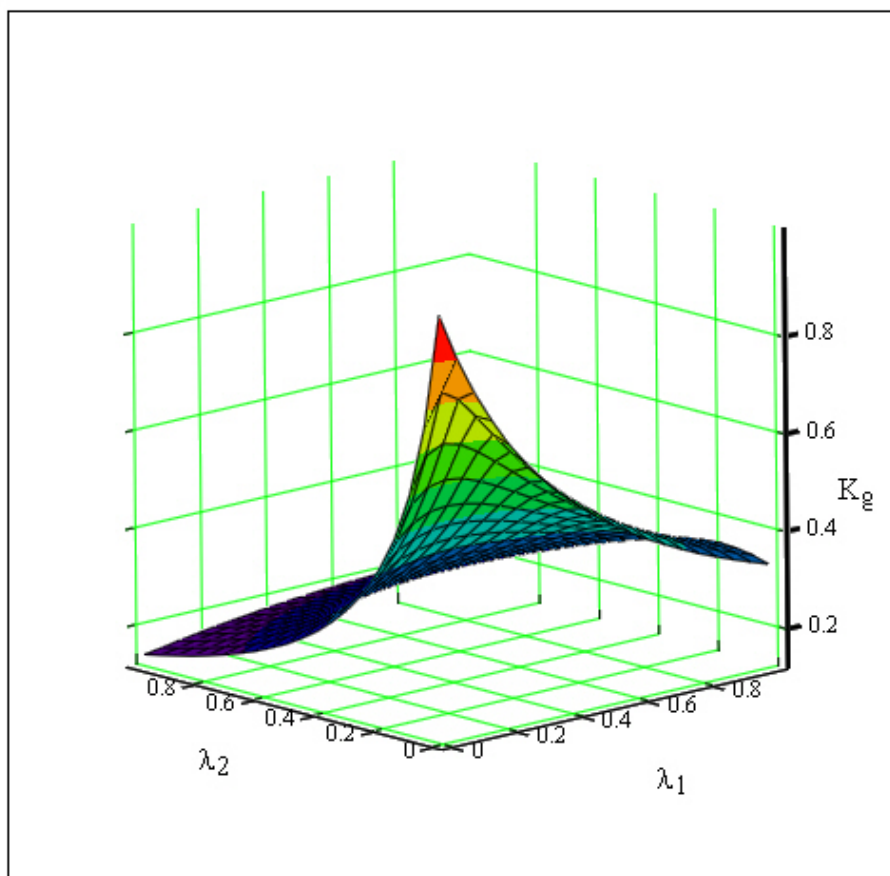
Wyznaczenie analityczne rozwiązania układu równań jest bardzo złożone. Dlatego też zastosowano wspomaganie komputerowe i wykorzystano program Mathcad 2001 Professional. Umożliwia on zobrazowanie graficzne funkcji $P_{00}(\lambda_1, \lambda_2)$. Pozwala to na przybliżone wyznaczenia wartości współrzędnych $(\lambda_{1optym}, \lambda_{2optym})$ odpowiadających maksimum. Zobrazowano to przykładem 1.

Przykład 1

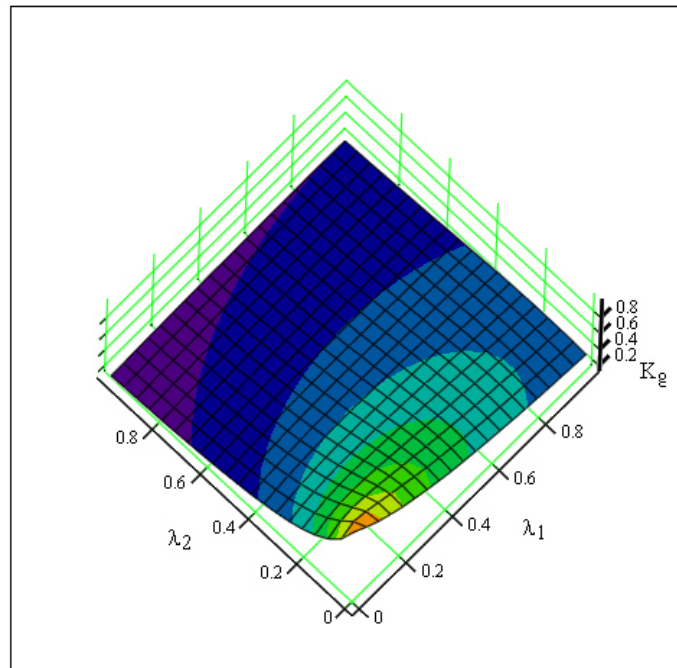
Przyjęto, że:

- intensywność uszkodzeń
 $\lambda = 1,2027 \cdot 10^{-5} \left[\frac{1}{h} \right]$ (odpowiada to systemowi, którego niezawodność wynosi 0,9),
- intensywność napraw $\mu = 0,0666 \left[\frac{1}{h} \right]$ (odpowiada to czasowi naprawy 15 [h]),
- intensywność realizacji przeglądów I rodzaju $\mu_1 = 0,5 \left[\frac{1}{h} \right]$ (odpowiada to czasowi przeglądu 2 [h]),
- intensywność realizacji przeglądów II rodzaju $\mu_2 = 0,1666 \left[\frac{1}{h} \right]$ (odpowiada to czasowi przeglądu 6 [h]).

Dla tak przyjętych założeń otrzymano wykres, który został przedstawiony na rys. 4 (widok od strony osi poziomych układu współrzędnych) i 5 (widok od „góry”).



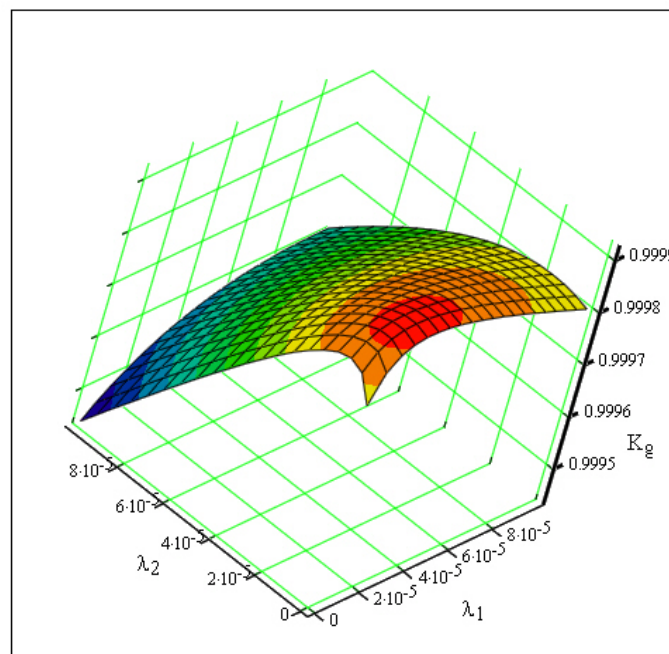
Rys. 4. Zależność wskaźnika gotowości K_g w funkcji intensywności przeglądów I rodzaju λ_1 i intensywności przeglądów II rodzaju λ_2



Rys. 5. Zależność wskaźnika gotowości K_g w funkcji intensywności przeglądów I rodzaju λ_1 i intensywności przeglądów II rodzaju λ_2 (widok z „góry”)

Analizując je można stwierdzić, że wartość wskaźnika gotowości K_g osiąga maksimum dla wartości intensywności przeglądu I i II rodzaju zawartych w przedziale $(0, 0.1)$. Aby precyzyjnie (graficznie) określić wartości poszukiwanych

intensywności niezbędne jest zmniejszenie maksymalnych wartości na osiach poziomych układu współrzędnych. Po analizie przyjęto je jako wartość: 0,0001. Dla tych warunków otrzymano wykres przedstawiony na rys. 6.



Rys. 6. Zależność wskaźnika gotowości K_g w funkcji intensywności przeglądów I rodzaju λ_1 i intensywności przeglądów II rodzaju λ_2 (dla max. wartości poziomych układu współrzędnych = 0,0001)

Wartości współrzędnych $(\lambda_{1optym}, \lambda_{2optym})$ odpowiadające maksimum są następujące:

$$\lambda_{1optym} = 2 \cdot 10^{-5} \left[\frac{1}{h} \right]$$

$$\lambda_{2optym} = 6 \cdot 10^{-6} \left[\frac{1}{h} \right]$$

Zostały one odczytane z rys. 6. Dla tych współrzędnych wartość wskaźnika gotowości jest maksymalna.

3. WNIOSKI

W artykule zaprezentowano metodę optymalizacji przeglądów okresowych, która uwzględnia wybrane własności niezawodnościowe (intensywność uszkodzeń) i eksploatacyjne (intensywność napraw, intensywności realizacji przeglądów). Umożliwia ona wyznaczenie optymalnych intensywności z jakimi powinny się one odbywać. Kryterium optymalizacyjne maksymalizuje wartość wskaźnika gotowości.

W rzeczywistych warunkach pracy transportowych systemów nadzoru przeglądy dokonywane są z różnymi intensywnościami, które nie są wartościami optymalnymi. Wynika to z konieczności poniesienia nakładów finansowych przez właściciela za dokonane przeglądy. Dlatego też rzeczywiste przeglądy tych systemów przeprowadzane są z intensywnościami mniejszymi niż optymalne. W przyszłych badaniach tego zagadnienia zachodzi więc potrzeba określenia zależności pomiędzy intensywnościami przeglądów a wskaźnikiem gotowości przy uwzględnieniu współczynnika nakładów finansowych.

LITERATURA

- [1]. Niziński S.: *Eksploatacja obiektów technicznych*. Wydaw. i Zakł. Poligrafii Inst. Technologii Eksploatacji, Radom 2002.
- [2]. Rosiński A.: *Analiza eksploatacji systemów nadzoru*. Międzynarodowa Konferencja Naukowa - Transport XXI wieku. Warszawa, wrzesień 2004.
- [3]. Rosiński A.: *Analiza niezawodnościowa struktur transportowych systemów nadzoru*. Prace naukowe Politechniki Radomskiej 3(23) 2005, Radom 2005.
- [4]. Szulc W., Rosiński A.: *Problemy eksploatacyjno-niezawodnościowe rozproszonego systemu bezpieczeństwa*. Zabezpieczenia Nr 1 (47)/2006, wyd. AAT, Warszawa 2006 r., str. 80 ÷ 82.



Mgr inż. Adam ROSIŃSKI pracuje na Politechnice Warszawskiej na Wydziale Transportu. Jego zainteresowania naukowe obejmują analizę niezawodnościowo-eksploatacyjną systemów bezpieczeństwa. W dorobku naukowym posiada kilkanaście publikacji naukowo-technicznych z dziedziny bezpieczeństwa. W bieżącym roku przystąpi do obrony rozprawy doktorskiej z zakresu optymalizacji strategii eksploatacji systemów nadzoru.

Praca naukowa finansowana ze środków na naukę w latach 2006/2007 jako projekt badawczy.

ZASTOSOWANIE DIAGNOSTYKI DO DECYZYJNEGO STEROWANIA PROCESEM EKSPLOATACJI URZĄDZEŃ

Jerzy GIRTLEK

Wydział Oceanotechniki i Okrętownictwa
Katedra Siłowni Okrętowych
Politechniki Gdańskiej, ul. G. Narutowicza 11/12, 80-952 Gdańsk
e-mail: jgirtl@pg.gda.pl

Streszczenie

Decyzyjne sterowanie procesem eksploatacji urządzeń technicznych należy do najistotniejszych problemów w eksploatacji tych urządzeń. Rozwiązanie tego problemu i zapewnienie racjonalnego sterowania wspomnianym procesem wymaga zastosowania diagnostyki technicznej. W artykule scharakteryzowano stochastyczną sytuację decyzyjną wynikającą z funkcjonowania tak istotnych urządzeń, jakimi są silniki spalinowe, sprężarki, pompy, kotły parowe, itp. w dowolnym systemie eksploatacyjnym i na tym tle przedstawiono propozycje zastosowania diagnostyki technicznej do sterowania procesem eksploatacji tych urządzeń.

Słowa kluczowe: diagnostyka techniczna, proces eksploatacji, sterowanie decyzyjne.

APPLICATION OF DIAGNOSTICS FOR DECISION CONTROL OVER THE PROCESS OF SYSTEMS OPERATION

Summary

Decision control of technical systems operation process is one of the most important problems in operating systems. Solving this problem and ensuring the rational control of the mentioned process demand employment of technical diagnostics. The paper provides description of a stochastic decision situation following from operating such significant systems like combustion engines, compressors, pumps, steam boilers, etc. of any operation system and on this background there are presented proposals of applying technical diagnostics for controlling the operation process of the systems.

Keywords: technical diagnostics, process of operation, decision control.

1. WSTĘP

Zastosowanie diagnostyki w fazie eksploatacji do decyzyjnego sterowania procesem eksploatacji wielu urządzeń technicznych (np. silników o zapłonie samoczynnym, turbinowych silników spalinowych, pomp, sprężarek, kotłów parowych i innych) ma szczególne znaczenie. Wynika to z dążenia ich użytkowników do utrzymania wysokiej trwałości tych urządzeń, także ich niezawodności i bezpieczeństwa działania. Zastosowanie diagnostyki technicznej umożliwia racjonalne sterowanie, czyli uporządkowane oddziaływanie użytkowników oraz urządzeń systemu sterującego (ST) na parametry struktury konstrukcyjnej urządzeń technicznych (jako systemów sterowanych – STN i zarazem diagnozowanych – SDN), zmierzające do uzyskania racjonalnego (a jeśli to możliwe optymalnego) przebiegu zachodzących w nich procesów (np. tarcia, zużycia) i w konsekwencji – procesów ich

eksploatacji. Sterowanie takie jest możliwe w przypadku zastosowania odpowiednich systemów diagnozujących (SDG), generujących sygnały niezbędne do zapewnienia ciągłości automatycznego sterowania wspomnianymi urządzeniami za pomocą odpowiednich systemów sterujących (STR), z uwzględnieniem sytuacji powodujących zagrożenia. Wobec tego zastosowanie diagnostyki wymaga opracowania i wdrożenia odpowiednich systemów sterowania (ST) i diagnostycznych (SD).

Każdy system sterowania (ST) jest wobec tego systemem składającym się z dwóch podsystemów: sterowanego (STN), którym jest diagnozowane urządzenie i sterującego (STR). Również każdy system diagnostyczny (SD) składa się z dwóch podsystemów: diagnozowanego (SDN), czyli wspomnianego urządzenia i diagnozującego (SDG). Wobec tego istnieje także potrzeba uwzględnienia, w procesie decyzyjnym, niezawodności systemu sterującego (STR) i systemu diagnozującego (SDG). Umożliwi to określenie wiarygodności diagnoz

generowanych przez *SDG* [3, 6]. System taki powinien być dostosowany do potrzeb identyfikacji stanów technicznych i energetycznych urządzeń jako *STN* i zarazem *SDN*. System diagnozujący (*SDG*) powinien przy tym wymuszać pożądane funkcjonowanie systemu sterowania (sterującego i sterowanego), stosownie do przebiegu implikujących się wzajemnie procesów zmian stanów technicznych i energetycznych urządzenia [3, 5].

Z rozważań tych wynika, że zastosowanie diagnostyki technicznej w decyzyjnym sterowaniu procesem eksploatacji urządzeń wymaga starannej identyfikacji tego problemu.

Ze względu na to, że silniki spalinowe o zapłonie samoczynnym należą do urządzeń najbardziej rozpowszechnionych, zatem wskazane jest dokonanie takiej identyfikacji urządzeń technicznych na przykładzie tego rodzaju silników.

2. FUNKCJONOWANIE URZĄDZEŃ A DIAGNOSTYKA

Sterowanie procesem eksploatacji dowolnego urządzenia technicznego, np. silnika spalinowego o zapłonie samoczynnym, rozumianym podobnie jak w publikacjach [5, 6, 8], wymaga między innymi sterowania zarówno jego stanami technicznymi jak również energetycznymi w czasie eksploatacji. Sterowanie takie polega na uporządkowanym oddziaływaniu na parametry jego struktury konstrukcyjnej, urządzeń sterujących i bezpośredniego użytkownika (systemu sterującego), według zasad, programów i algorytmów przyjętych przez decydenta, a dostosowanych do wykonywanych zadań, których podejmowanie wymaga uwzględnienia ograniczeń (*H*) funkcji kryterialnej (*F*) oraz wyników diagnozowania w formie diagnoz (*D*), bądź diagnoz i zarazem prognoz (*P*), bądź też diagnoz i ewentualnych genez (*G*), albo diagnoz pełnych (czyli łącznie diagnoz, prognoz i genez), generowanych przez system diagnozujący (*SDG*). Sytuacja ta została zobrazowana w formie schematu przedstawionego w pracy [3]. Każde zadanie (*Z*) należące do zbioru wspomnianych zadań może być interpretowane jako poprawne działanie (*D_Z*) urządzenia, w określonych warunkach eksploatacji (*W_R*) i w ustalonym czasie (*t*). Sterownie to powinno zapewnić racjonalną pracę urządzenia w sytuacjach zwykłych (normalnych) i szczególnych (skomplikowanych, niebezpiecznych i awaryjnych a nawet katastroficznych) [4].

Racjonalne sterowanie działaniem (pracą) dowolnego urządzenia technicznego oraz związane z nim podejmowanie decyzji eksploatacyjnych (*DE*) stosownie do istniejącego stanu technicznego i energetycznego tego urządzenia wymagają łącznej (jednocześnie) znajomości aktualnej diagnozy (*D*), genezy (*G*) i prognozy (*P*), uzależnionej od wewnętrznych i zewnętrznych ograniczeń (*H*),

przyjętej funkcji kryterialnej (*F*) i różnych zakłóceń: systemu sterowanego i diagnozowanego (*K_D*), systemu diagnozującego (*K_G*), systemu decyzyjnego (*K_C*), czyli określonego decydenta, a także zakłóceń systemu sterującego (*K_T*) [3]. Opracowanie koncepcji takiego sterowania jest możliwe w przypadku zastosowania statystycznej teorii decyzji bądź teorii decyzyjnych (sterowanych) procesów semimarkowskich [3, 4, 6].

Koncepcja decyzyjnego sterowania oraz uwarunkowania sprawnego sterowania (z uwzględnieniem diagnozowania) parametrami struktury konstrukcyjnej dowolnego silnika o zapłonie samoczynnym oraz procesem eksploatacji tego silnika (jako systemu sterowanego i zarazem diagnozowanego), a tym samym jego działaniem, została przedstawiona w pracy [3].

We wspomnianej koncepcji, która może być dostosowana do potrzeb sterowania procesem eksploatacji dowolnego urządzenia technicznego z uwzględnieniem diagnostyki, istotne jest uwzględnienie faktu, że sterowanie procesem eksploatacji każdego takiego urządzenia może być nieracjonalne, jeśli opracowana dla potrzeb tego sterowania diagnoza (bądź prognoza) będzie niewłaściwa. W praktyce eksploatacyjnej wielu urządzeń, podobnie jak w przypadku wspomnianych silników, istnieje zawsze określone prawdopodobieństwo zdarzenia *B* oznaczającego opracowanie niewłaściwej (błędnej) diagnozy, które można oszacować korzystając z następującego wzoru:

$$P(B) = \sum_{i=1}^4 P(A_i)P(B/A_i) \quad (1)$$

gdzie:

B – zdarzenie oznaczające opracowanie niewłaściwej (błędnej) diagnozy;

A₁ – zdarzenie oznaczające uszkodzenie systemu diagnozującego (*SDG*);

A₂ – zdarzenie oznaczające zakłócenia procesu diagnozowania, w wyniku istnienia losowych zakłóceń sterowania i zasilania urządzeń;

A₃ – zdarzenie oznaczające zająście stanu urządzenia (czyli *SDN*), który nie został uwzględniony w zadaniu diagnostycznym (*ZD*);

A₄ – zdarzenie oznaczające popełnienie błędu przez użytkownika diagnozy, np. skutek zmęczenia, nieodpowiednich jego kwalifikacji, itp.

Z rozważań tych wynika, że przy podejmowaniu decyzji eksploatacyjnych powinna być uwzględniana wiarygodność diagnozy. Zagadnienie to zostało przedstawione w publikacjach [4, 6]. Konieczność dążenia do określenia wiarygodności diagnozy wynika stąd, że prawidłowo sformułowana diagnoza o stanie technicznym (np. *S*) ma formę hipotezy. Hipoteza taka jest zdaniem empirycznym zawierającym implikację syntaktyczną. W zdaniu tym, ta jego część, która poprzedza wspomnianą implikację jest prawdopodobną przyczyną

obserwowanego faktu, zawartego w części zdania następującej po tej implikacji. Faktem tym jest stwierdzenie, że obserwowany jest taki a nie inny wektor wartości parametrów diagnostycznych (np. K). Wobec tego prawidłowo sformułowana diagnoza ma (w ujęciu ogólnym) następującą treść: *urządzenie znajduje się w takim (np. S) a nie innym stanie technicznym dlatego, ponieważ obserwowany jest taki (np. K) a nie inny wektor wartości parametrów diagnostycznych*. W przypadku, gdy można stwierdzić, że podczas badania diagnostycznego urządzenia i opracowania diagnozy o jego stanie system diagnozujący (SDG) był w stanie pełnej zdadności (a więc działał niezawodnie) powinna być uwzględniana trafność diagnozy [6]. Wynika także to, że racjonalne uwzględnienie diagnostyki w procesie sterowania urządzeniami technicznymi wymaga dokonania identyfikacji diagnostycznej poszczególnych urządzeń, w wyniku której powinny zostać opracowane informacje umożliwiające skonstruowanie odpowiednich ich modeli diagnostycznych [2, 7, 8, 9, 11, 12, 13].

Wymaga także identyfikacji procesu diagnozowania, w celu uzyskania informacji określających specyfikę tego procesu. Informacje te mają istotne znaczenie, ponieważ przebieg tego procesu wpływa istotnie na wiarygodność (ewentualnie trafność) diagnozy o stanie technicznym poszczególnych urządzeń technicznych. Proces ten może być różnie postrzegany. Na podstawie dotychczasowych rozważań [1, 2, 6, 7, 10, 11, 12, 13] można uznać, że proces diagnozowania jest dwuwymiarowym procesem stochastycznym, który składa się z procesu $\{B(t): t \geq 0\}$ funkcjonowania SDG i procesu pozyskiwania informacji o stanie SDN

$\{C(\mathcal{G}): \mathcal{G} \geq 0\}$. Zatem proces ten można określić następująco:

$$D(t, \mathcal{G}) = [B(t), C(\mathcal{G})]; t, \mathcal{G} \in R_+ \quad (2)$$

gdzie:

$B(t)$ – składowa procesu, która jest rozpatrywana w czasie funkcjonowania (użytkowania) SDG, czyli składowa rozpatrywana w czasie długim (w czasie pracy SDG, w którym nie muszą być generowane diagnozy chwilowe);

$C(\mathcal{G})$ – składowa procesu, która jest rozpatrywana w czasie realizowania pomiaru i wnioskowania diagnostycznego, czyli składowa rozpatrywana w czasie krótkim (w czasie pracy SDG, w którym uzyskiwana jest diagnoza);

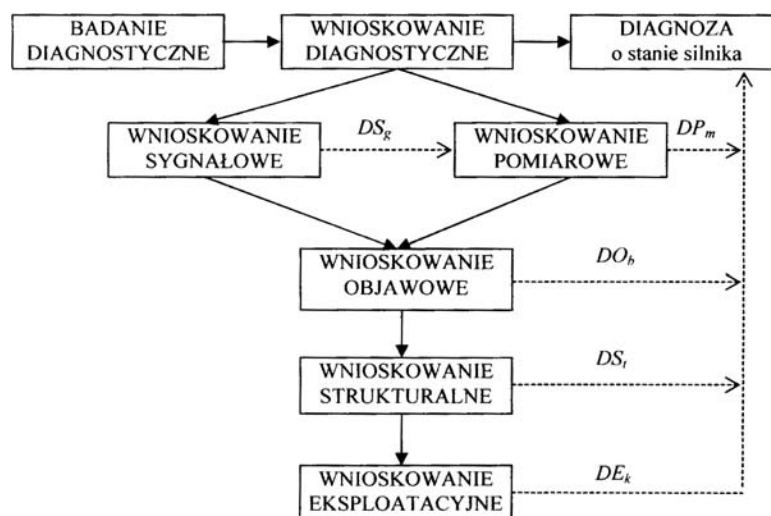
R_+ – zbiór liczb rzeczywistych nieujemnych.

Proces $\{C(\mathcal{G}): \mathcal{G} \geq 0\}$ zawsze tworzą następujące realizacje:

- badanie diagnostyczne urządzenia (np. silnika),
- wnioskowanie diagnostyczne o jego stanie.

Można wyróżnić takie rodzaje wnioskowania jak: sygnałowe, pomiarowe, objawowe, strukturalne i eksploatacyjne [1]. Schemat takiego wnioskowania, np. w przypadku silnika, przedstawiono na rys. 1.

Wobec tego we wspomnianym procesie $\{C(\mathcal{G}): \mathcal{G} \geq 0\}$ można wyróżnić wartości (stany), jak badanie diagnostyczne oraz wymienione wnioskowania: sygnałowe, pomiarowe, objawowe, strukturalne i eksploatacyjne. Z kolei czas trwania tych rodzajów wnioskowania jest czasem realizacji wymienionych stanów. Zatem jest to proces klasy: dyskretny w stanach i ciągły w czasie.



Rys. 1. Schemat badania i wnioskowania diagnostycznego umożliwiającego opracowanie diagnozy o stanie technicznym dowolnego urządzenia (SDN) i tym samym podjęcia decyzji eksploatacyjnej: urządzenie (np. silnik spalinowy) „zdadne”, „częściowo zdadne”, bądź „niezdadne” do zastosowania zgodnie z przeznaczeniem; DS_g – diagnoza sygnałowa, DP_m – diagnoza pomiarowa, DO_b – diagnoza objawowa, DS_i – diagnoza strukturalna, DE_k – diagnoza eksploatacyjna

Z dotychczasowych badań wynika, że proces ten może być uznany za semimarkowski [4, 6].

Błędy popełnione w każdym z wspomnianych rodzajów wnioskowania mogą się kumulować. Ponadto są tym większe im więcej stanów *SDN* zawartych jest w zadaniu diagnostycznym. Ich znaczenie jest tym większe im bardziej potrzebna jest diagnoza o dużej wiarygodności. Fakty te powodują, że najczęściej rozpatrywane są trzy klasy stanów technicznych *SDN*: pełnej zdatności (s_1), częściowej (niepełnej) zdatności (s_2) oraz niezdatności (s_3).

Diagnozy o stanie *SDN* generowane przez *SDG* mogą być, ze względu na przydatność eksploatacyjną, podzielone ogólnie na użytkową i obsługową.

Diagnoza użytkowa zawiera informację o zdolności *SDN* do wykonywania zadań, natomiast diagnoza obsługowa – informację o zakresie potrzeb związanych z odnową utraconego stanu zdatności (s_1 bądź s_2).

Opracowanie wymienionych rodzajów diagnoz (wiarygodnych) wymaga zbudowania adekwatnego modelu diagnostycznego urządzenia oraz odpowiedniego do tego modelu – systemu diagnostycznego (*SD*).

Szczególne znaczenie ma wybór (ze względu na przyjęty cel badań) rodzaju modelu diagnostycznego dla poszczególnych urządzeń technicznych. Należy starannie uzasadnić w drodze wnioskowania dedukcyjnego bądź indukcyjnego czy model ten ma być stochastyczny, czy pseudo-deterministyczny, symptomowy, odwrotny, rozmyty, neuronowy, symulacyjny, czy też inny [2, 9, 11, 12, 13].

3. MODEL DIAGNOSTYCZNY URZĄDZENIA DLA POTRZEB STEROWANIA

Najistotniejszą fazą procesu tworzenia modelu diagnostycznego każdego urządzenia, z uwzględnieniem potrzeb jego sterowania jest sformułowanie zadania diagnostycznego. Najważniejszą częścią tego zadania jest opracowanie zbioru stanów urządzenia, do których rozpoznania należy dążyć w czasie jego eksploatacji. Zbiór takich stanów można opracować na podstawie informacji o uszkodzeniach urządzeń danego typu już eksploatowanych bądź urządzeń do nich podobnych. Równie ważną częścią tego zadania jest wydzielenie ze zbioru procesów wyjściowych (*PW*) urządzenia tych, które mogą być przyjęte za sygnały diagnostyczne (K^*). Następnym etapem tworzenia modelu diagnostycznego urządzenia jest wyodrębnienie spośród wielu cech sygnałów (K^*) takich, które zawierają najwięcej

informacji diagnostycznych. Te właśnie cechy, jako symptomy diagnostyczne należy następnie zweryfikować według kryteriów odwzorowania diagnostycznego (*KO*). Po takiej weryfikacji symptomy te stanowią parametry diagnostyczne (*K*). W ten sposób opracowany model diagnostyczny urządzenia może być przysposobiony do zastosowania klasycznej bądź uproszczonej (asocjacyjnej) metody diagnostycznej [3, 7, 8]. Przykładowy schemat tworzenia modelu diagnostycznego urządzenia na przykładzie silnika o zapłonie samoczynnym został przedstawiony na rys. 2.

Stosowanie uproszczonej metody diagnostycznej (asocjacyjnej) jest możliwe dlatego, ponieważ można uznać za prawdziwe następujące twierdzenie [3]:

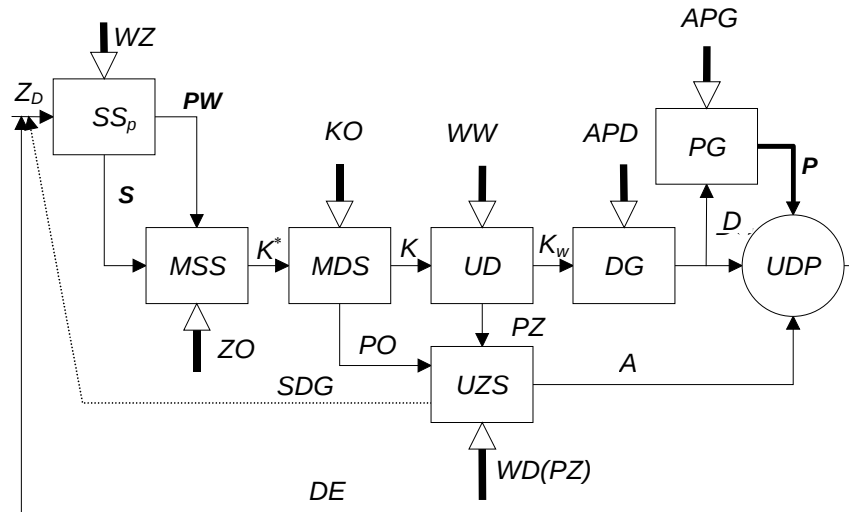
jeżeli zgodnie z przyjętym kryterium podziału stanów dowolnego urządzenia technicznego określone jego stany w dwóch różnych chwilach są sobie bliskie, to bliskie są również sygnały diagnostyczne wysyłane przez to urządzenie w tych samych dwóch chwilach.

Za sygnał diagnostyczny można uznać wektor wartości parametrów diagnostycznych, które mogą być zarejestrowane przez *UD SDG* podczas diagnozowania urządzenia jako *SDN* [3, 6, 12, 13].

Twierdzenie to, po zmodyfikowaniu umożliwiającym porównywanie stanu wielu urządzeń, można przedstawić w następującej wersji:

jeżeli określone zgodnie z tym samym kryterium stany różnych urządzeń tego samego typu w tej samej chwili są sobie bliskie, to bliskie są także generowane przez nie w tej samej chwili ich sygnały diagnostyczne.

Zgodnie z tymi założeniami obowiązuje zasada, według której *SDG* kwalifikuje badany stan techniczny urządzenia (*SDN*) do tej klasy stanów diagnostycznych wzorcowych, której charakterystyczny sygnał diagnostyczny w danej chwili jest najbardziej skorelowany z sygnałem wygenerowanym przez diagnozowane urządzenie i odebrany w dowolnej chwili przez wspomniany *SDG*.



Rys. 2. Schemat tworzenia MDS i jego przykładowego zastosowania: ZD – zasilanie, WZ – warunki zewnętrzne, SS_p – silnik spalinowy, S – stany silnika, PW – procesy wyjściowe, MSS – model stanów silnika, ZD – zadanie diagnostyczne, K^* – sygnały diagnostyczne, MDS – model diagnostyczny silnika, KO – kryteria odwzorowania diagnostycznego, K – parametry diagnostyczne, PO – program zmian obciążenia, PZ – parametry zabezpieczenia, WW – wartości wzorcowe K, K_w – wyniki sprawdzeń, UD – urządzenie diagnozujące, DG – diagnoza, APD – algorytmy i programy diagnozowania, APG – algorytmy i programy prognozowania, D – diagnoza, PG – prognoza, KP – kryteria prognozy, P – prognoza, UDP – użytkownik diagnozy i prognozy, UZS – układ zabezpieczeń silnika, A – sygnał alarmowy, WD(PZ) – wartości dopuszczalne PZ, SDG – sterowanie obciążeniem w stanach granicznych, DE – decyzje eksploatacyjne (użytkowe i obsługowe) [3].

4. ZASTOSOWANIE DIAGNOSTYKI DO STEROWANIA PROCESEM EKSPLOATACJI URZĄDZEŃ

Z rozważań przedstawionych w różnych publikacjach wynika, że niewłaściwe sterowanie procesami roboczymi urządzeń technicznych może doprowadzić do szybkich zmian parametrów ich struktury konstrukcyjnej a w rezultacie do przedwczesnych uszkodzeń elementów tych urządzeń [3, 4, 7, 12].

Zapobieganie szybkiemu zużyciu elementów dowolnego urządzenia i tym samym przedwczesnym ich uszkodzeniom jest możliwe, jeżeli będą stosowane odpowiednie obsługi profilaktyczne. Przewidywanie potrzeby wykonania takich usług i ich racjonalne planowanie wymaga zastosowania odpowiednich systemów diagnozujących [3, 4, 8, 12, 13].

Planowanie racjonalne usług profilaktycznych jest trudne z uwagi na to, że czynniki oddziałujące na urządzenia tak uwarunkowane, jak też nieuwarunkowane ich funkcjonowaniem mają naturę losową. Wskutek tego również własności struktury konstrukcyjnej urządzeń są wielkościami losowym (zmiennymi losowymi). W praktyce eksploatacyjnej należy się spodziewać, że każde urządzenie z danego zbioru będzie miało w tej samej chwili inny stan techniczny. Ponadto ich obciążenia cieplno-mechaniczne istniejące w czasie ich pracy będą procesami losowymi. Wynika to z tego, że obciążenia te zależą przede wszystkim od zadań,

jakie wykonuje urządzenie techniczne oraz własności otoczenia, w którym jest ono eksploatowane. Niekiedy zależą istotnie także od kwalifikacji i psychofizycznych predyspozycji użytkowników poszczególnych urządzeń. W zależności od obciążeń i początkowego stanu technicznego zmiana struktury konstrukcyjnej każdego urządzenia, w tym samym przedziale czasu, pracy będzie inna. Wynika to z tego, że zmiana stanu technicznego urządzeń jest słabo skorelowana ze czasem ich pracy do najbliższej obsługi profilaktycznej. Stąd oczywisty wniosek, że czas działania (pracy) urządzeń nie może być dobrą miarą zużycia ich struktury konstrukcyjnej. Zatem planowanie usług profilaktycznych urządzeń technicznych po upływie stałych przedziałów czasu poprawnej pracy umożliwia wprowadzenie sterowania procesem eksploatacji tych urządzeń, ale obniża ich współczynnik gotowości technicznej oraz wskaźniki sprawności funkcjonowania [4, 5]. Zatem korzystniejsze jest sterowanie procesem usług profilaktycznych urządzeń technicznych polegające na zastosowaniu metod realizowania tych usług według określonego planu, ale po uprzednim zidentyfikowaniu stanów technicznych poszczególnych urządzeń za pomocą odpowiednich SDG [3, 4].

Wobec tego sterowanie działaniem urządzenia technicznego, jako systemu sterowanego (STN) i diagnozowanego (SDN), powinno być nadzorowane przez system diagnozujący (SDG), sprzężony z systemem sterującym (STR). Sprzężenie

takie umożliwia systemowi sterującemu oddziaływanie na urządzenie techniczne według diagnoz systemu diagnozującego i kształtowanie stanów energetycznych tego urządzenia w zależności od zmian jego stanu technicznego. Opracowanie systemu diagnostycznego (SD) jako SDN i SDG o odpowiednich relacjach (R) między nimi, wymaga zbudowania diagnostycznego modelu urządzenia, w którym byłyby zawarte zarówno potrzeby identyfikacji stanów technicznych i energetycznych danego urządzenia, jak również kontroli (nadzoru) funkcjonowania systemu sterującego jego pracą. Ze względu na oddziaływanie, na wszystkie wspomniane systemy, czynników destrukcyjnych, opracowane według przyjętych zasad programy i algorytmy sterowania powinny być korygowane w zależności od aktualnych stanów, w których się te systemy znajdują.

Rozpoznawanie i przewidywanie stanów technicznych i energetycznych dowolnego urządzenia, a tym samym jego procesu eksploatacji wymaga zastosowania odpowiedniego systemu diagnostycznego (SD), adekwatnego do potrzeb sterowania [1, 3, 4, 6, 12, 13]. System ten jak już wspomniano, może być określony następująco:

$$SD = \langle SDN, SDG, R \rangle$$

przy czym:

$$R = \{R_k, R_f, R_n, R_d\}$$

gdzie:

SD – system diagnostyczny, SDN – podsystem diagnozowany (urządzenie), SDG – podsystem diagnozujący, R – struktura SD , R_k – struktura konstrukcyjna, R_f – struktura funkcjonalna, R_n – struktura niezawodnościowa, R_d – struktura diagnostyczna.

Podsystem (system) diagnozowany SDN (urządzenie) jest przedmiotem diagnozowania (PmD), zaś podsystem (system) diagnozujący SDG składa się z podmiotu diagnozowania (PdD) oraz środków diagnozowania ($\acute{S}D$).

Zatem system diagnozujący (SDG) może być interpretowany następująco [3]:

$$SDG = \langle PdD, \acute{S}D, R_{P\acute{S}} \rangle$$

gdzie:

$R_{P\acute{S}}$ – relacje między PdD a $\acute{S}D$, tworzące strukturę SDG (konstrukcyjną, funkcjonalną, niezawodnościową oraz diagnostyczną).

Środkami diagnozowania ($\acute{S}D$) są narzędzia diagnozy. Środki te można określić w formie zbioru:

$$\acute{S}D = \{MB, TP, UD, TD, AD, DG, PG, GN\}$$

gdzie:

MB – metody diagnozowania, TP – technologia pomiarów, UD – urządzenia diagnozujące (pomiarowe), TD – technologia przetwarzania danych i opracowania diagnozy, AD – algorytmy diagnostyczne, DG – diagnozer, PG – prognozer, GN – genezer.

Elementy $\acute{S}D$ są z reguły złożone. W ujęciu formalnym można je przedstawić jako zbiory. Na przykład urządzenia diagnozujące (UD) stanowią zbiór:

$$UD = \{C, P, A, L, S, J\}$$

natomiast algorytmy diagnostyczne (AD) – zbiór

$$AD = \{ADg, ADz, APr, AGn\}$$

gdzie:

C – czujniki, P – przewody, A – analizatory sygnału, L – urządzenia logistyczne, S – sygnalizatory, J – rejestratory, ADg – algorytmy diagnozowania, ADz – algorytmy dozoru, APr – algorytmy prognozowania, AGn – algorytmy genezowania.

W tak rozumianym SDG , takie jego elementy jak: MB , TP , UD i TD powinny zapewnić wiarygodne badanie diagnostyczne urządzenia, zaś elementy: AD , DG , PG , GN – prawidłowe wnioskowanie diagnostyczne (sygnałowe, pomiarowe, objawowe, strukturalne oraz eksploatacyjne) prowadzące do sformułowania diagnozy o określonej wiarygodności.

System diagnostyczny powinien zapewniać: diagnozę, prognozę oraz genezę w takim zakresie, który umożliwiłby racjonalny przebieg zarówno użytkowania, jak i obsługiwanego urządzeń. Użytkowanie, jako uporządkowany zbiór zainicjowanych oraz podtrzymywanych procesów roboczych i organizacyjnych umożliwiających realizację celu, jest działaniem pożądanym w przeciwieństwie do obsługiwanego, które jest niepożądane, ale konieczne.

Użytkowanie może być czynne (w urządzeniu zachodzi przemiana energii w formie pracy bądź ciepła) lub bierne (urządzenie oczekuje na uruchomienie). W pierwszym przypadku urządzenie znajduje się w stanie eksploatacyjnym e_1 , a w drugim – w stanie eksploatacyjnym e_2 . W czasie użytkowania aktywnego powinno istnieć dozoruwanie. Zapewnienie tej formy diagnostycznego działania może być uzyskane przez wykorzystanie w procesie wnioskowania diagnostycznego parametrów diagnostycznych ogólnych (PDG_o), które umożliwiają ustalenie, czy jest ono w stanie zdatości pełnej s_1 , czy też w stanie zdatości częściowej s_2 [3, 4]. Wtedy można skorzystać z następującej reguły oceny stanu technicznego urządzenia:

jeżeli wartości parametrów diagnostycznych ogólnych (PDG_o) są w przedziałach wartości dopuszczalnych (czyli nie przekraczają ustalonych dla nich wartości granicznych), to znaczy, że urządzenie jest w stanie s_1 ; jeżeli tak nie jest, to znaczy, że jest ono w stanie s_2 .

Oznacza to, że proces eksploatacji urządzenia $\{Y(t): t \in T\}$ może znajdować się w stanie $z_2 = (s_1, e_1)$ lub w stanie $z_3 = (s_2, e_1)$ [3, 4].

Ze względu na to, że procesowi diagnostycznemu towarzyszą różne zakłócenia [1, 3, 4, 6, 7, 12, 13] może się zdarzyć tak, że urządzenie znajdujące się w stanie s_1 zostanie uznane za będące w stanie s_2 . W tej sytuacji, aby przeciwdziałać zakłóceniom i wynikającym z ich istnienia skutkom i zwiększyć tym samym zdolność rozdzielczą systemu diagnozującego i w związku z tym wiarygodność diagnozy, należy doprowadzić między innymi do nadmiaru informacji. Można ją uzyskać przez wykorzystanie parametrów diagnostycznych dodatkowych, które można nazwać szczegółowymi (PDG_s). Wtedy można byłoby przyjąć zasadę, że w przypadku, gdy wartości parametrów diagnostycznych szczegółowych (PDG_s) mieszczą się w przedziale wartości dopuszczalnych, czyli nie przekraczają wartości granicznych, to można uznać, że:

- ta część $UD\ SDG$, która umożliwia rejestrację wartości PDG_o jest uszkodzona;
- urządzenie można nadal użytkować, jeżeli nie jest potrzebna obsługa profilaktyczna ze względu na wymagania wynikające z konieczności podjęcia określonego zadania.

W przypadku, gdy istnieje potrzeba wykonania obsługi profilaktycznej, w celu dokonania odnowy stanu technicznego, urządzenie może znajdować się, w stanach: zdatości częściowej (s_2) lub niezdatności (s_3). Oba te stany implikują stan obsługi profilaktycznej (e_3). Oznacza to, że proces eksploatacji urządzenia $\{Y(t): t \geq 0\}$ może znajdować się w stanie $z_4 = (s_2, e_3)$ lub $z_5 = (s_3, e_3)$. Po zakończeniu prac profilaktycznych, należy przeprowadzić prognozowanie stanów urządzenia przyjmując odpowiednie kryterium, na przykład dotyczące czasu poprawnej pracy. W przypadku, gdy istotne jest to, aby czas poprawnej pracy (T_p) był większy lub co najmniej równy czasowi (t_z) potrzebnemu do wykonania zadania, to jeżeli:

- $T_p > t_z$, można uznać, że urządzenie jest przygotowane do użytkowania;
- $T_p < t_z$, należy przyjąć, że urządzenie jest w stanie niezdatności (a więc s_3), który implikuje stan eksploatacyjny e_4 (obsługi wymuszonej przez uszkodzenie) tego urządzenia i tym samym uważać, że proces jego eksploatacji $\{Y(t): t \geq 0\}$ znajduje się w stanie $z_6 = (s_3, e_4)$.

W przypadku stwierdzenia stanu z_6 , należy ustalić przyczyny zajścia tego stanu. Jeśli zakres prac obsługowych zostanie zakończony należy ocenić, czy usuwanie stwierdzonych przyczyn jest opłacalne. Zatem trzeba odpowiedzieć na pytanie, czy koszt odnowy (K_o) urządzenia w systemie eksploatacji jest co najwyżej równy kosztowi dopuszczalnemu (K_D), a więc czy $K_o < K_D$. Jeśli warunek ten jest spełniony, to wykonanie obsługi tego urządzenia jest uzasadnione. W przeciwnym razie, gdy $K_o > K_D$, odnowa jest nieopłacalna, więc

powinno ono ulec kasacji (złomowaniu). W razie stwierdzenia, że obsługiwane jest uzasadnione należy doprowadzić do pełnej odnowy urządzenia. Po jej wykonaniu trzeba ponownie dokonać prognozowania jego stanu technicznego.

Z kolei, gdy urządzenie techniczne znajduje się w stanie s_2 (co oznacza, że nie wszystkie PDG_s przyjmują wartości odpowiadające stanowi s_1) należy zidentyfikować przyczyny zajścia stanu. Następnie należy przeprowadzić prognozowanie możliwości zajścia stanu s_3 . W przypadku, gdy małe jest prawdopodobieństwo zajścia tego stanu, a stan s_2 wystarczy do wykonania zadania, można urządzenie dopuścić do dalszego użytkowania. W przeciwnym razie należy ponowić działanie, które zostało opisane wcześniej.

W przedstawionej koncepcji działania należy przyjąć zasadę, że urządzenie może znajdować się w stanie eksploatacyjnym e_2 tylko wtedy, gdy jego stan techniczny może być zaliczony do stanu s_1 . Oznacza to, że wtedy proces eksploatacji urządzenia $\{Y(t): t \in T\}$ znajduje się wtedy w stanie $z_1 = (s_1, e_2)$. Realizacja tej koncepcji wymaga zastosowania systemów diagnozujących o odpowiedniej przydatności eksploatacyjnej [3, 8, 9, 11, 12, 13]. Koncepcja wykorzystania wszystkich wspomnianych form diagnostycznego działania została przedstawiona w pracy [3].

5. UWAGI KOŃCOWE I WNIOSKI

Zapewnienie informacji o stanie urządzeń technicznych, niezbędnych dla potrzeb sterowania ich procesem eksploatacji wymaga wykorzystania w praktyce eksploatacyjnej wszystkich czterech form diagnostycznego działania, a mianowicie: diagnozowania, dozoru, prognozowania i genezowania. Zastosowanie tych form działania jest możliwe tylko w przypadku dysponowania w systemie eksploatacji urządzeń odpowiednim systemem diagnostycznym.

Zbudowanie systemu diagnostycznego przydatnego do racjonalnego sterowania procesem eksploatacji urządzeń technicznych wymaga z reguły opracowania ich probabilistycznego modelu diagnostycznego. Opracowanie takiego modelu jest możliwe dopiero po dokonaniu identyfikacji diagnostycznej danego typu urządzeń, w wyniku której powinno dojść do ustalenia zbioru uszkodzeń i wynikającego z niego zbiorów: stanów urządzenia, które należy rozpoznać, a także parametrów diagnostycznych oraz relacji odwzorowania zbioru wyróżnionych stanów silników w zbiór parametrów diagnostycznych i odwrotnie, z uwzględnieniem własności losowych tych relacji.

Przy opracowaniu modelu diagnostycznego urządzeń o małej liczności, na ogół brak jest informacji o uszkodzeniach i dlatego należy wykorzystać dane dotyczące urządzeń podobnych.

W przedstawionej koncepcji zastosowania diagnostyki do sterowania procesem eksploatacji urządzeń technicznych odzwierciedlone są dwa procesy [1, 3, 6]:

- proces diagnozowania $SDN \{C(\mathcal{G}): \mathcal{G} \geq 0\}$, który stanowi przebieg następujących po sobie i powiązanych przyczynowo w czasie działań tworzących następujący łańcuch diagnostyczny: badanie diagnostyczne $\rightarrow$ wnioskowanie diagnostyczne, które składać się może z kolejno następujących po sobie wnioskowań: sygnałowego, pomiarowego, objawowego, strukturalnego oraz eksploatacyjnego;
- proces użytkowania $SDG \{B(t): t \geq 0\}$, którego stany mogą być interpretowane następująco: stan d_1 , który oznacza użytkowanie SDG , gdy znajduje się on w stanie pełnej zdatności, stan d_2 , który oznacza użytkowanie SDG , gdy znajduje się on w stanie innym niż pełnej zdatności (czyli w stanie częściowej zdatności bądź niezdatności), stan d_3 , który oznacza użytkowanie SDG , gdy znajduje się on w stanie pełnej zdatności, lecz pojawił się stan urządzenia nie uwzględniony w zadaniu diagnostycznym [6, 11, 12].

Potrzeba wyróżnienia dodatkowo wnioskowania sygnałowego wynika stąd, że nie można pominąć wrażeń zmysłowych użytkowników urządzeń technicznych, którzy są istotnymi elementami SDG , nadzorującymi przebieg procesu diagnozowania za pomocą swoich zmysłów.

6. LITERATURA

- [1] Będkowski L.: *Elementy diagnostyki technicznej*. Wyd. 2. WAT, Warszawa 1992.
- [2] W. Cholewa, J. Kiciński: *Diagnostyka techniczna. Odwrotne modele diagnostyczne*. Monografia. Wyd. Politechniki Śląskiej, Gliwice 1977.
- [3] Girtler J.: *Diagnostyka jako warunek sterowania eksploatacją okrętowych silników spalinowych*. Studia Nr 28. WSM, Szczecin 1997.
- [4] Girtler J., Kuszmidler S., Plewiński L.: *Wybrane zagadnienia eksploatacji statków morskich w aspekcie bezpieczeństwa żeglugi*. Monografia. WSM, Szczecin 2003.
- [5] Girtler J., Kitowski Z., Kuriata A.: *Bezpieczeństwo okrętu na morzu. Ujęcie systemowe*. WKiŁ, Warszawa 1995.
- [6] Girtler J.: *Zastosowanie wiarygodności diagnozy do podejmowania decyzji w procesie*

eksploatacji urządzeń. Materiały V Krajowej Konferencji „Diagnostyka Techniczna Urządzeń i Systemów”. Wydział Elektroniki WAT. Zakład poligraficzny „Aldan”, Warszawa 2003, s.100-109.

- [7] Michalski R.: *Diagnostyka Maszyn Roboczych*. Wyd. ITE, Radom 2004.
- [8] Niziński S.: *Eksploatacja obiektów technicznych*. Wyd. ITE. Radom 2002.
- [9] Żółtowski B.: *Podstawy diagnostyki maszyn*. Wyd. ATR, Bydgoszcz 1996.
- [10] *Analiza ryzyka i diagnostyka procesów degradacyjnych i zmęzeniowych*. Praca zbiorowa pod redakcją S. Radcowskiego. Wyd. Politechnika Warszawska, Wydział Samochodów i Maszyn Roboczych, Instytut Podstaw Budowy Maszyn, 2004.
- [11] *Diagnostyka maszyn. Zasady ogólne. Przykłady zastosowań*. Praca zbiorowa pod redakcją C. Cempla i F. Tomaszewskiego. Wyd. Międzyresortowe Centrum Naukowe Majątku Trwałego. Radom 1992.
- [12] *Inżynieria diagnostyki Maszyn*. Praca zbiorowa po redakcją B. Żółtowskiego i C. Cempla. PTDT. Wyd. ITE, Warszawa, Bydgoszcz, Radom 2004.
- [13] *Modelowanie i diagnostyka oddziaływań mechanicznych, aerodynamicznych i magnetycznych w turbozespołach energetycznych*. Praca zbiorowa pod redakcją J. Kicińskiego. Wydział IV PAN. Wyd. IMP PAN, Gdańsk 2005.



Prof. dr hab. inż. **Jerzy GIRTLEK** jest kierownikiem Katedry Siłowni Okrętowych na Wydziale Oceanotechniki i Okrętownictwa Politechniki Gdańskiej. Zajmuje się zagadnieniami dotyczącymi, niezawodności, bezpieczeństwa, diagnostyki urządzeń siłowni okrętowych (ze szczególnym uwzględnieniem silników o zapłonie samoczynnym) oraz decyzyjnego sterowania procesami eksploatacji tych urządzeń. Jest przewodniczącym Rady Naukowej kwartalnika o zasięgu międzynarodowym *Polish Maritime Research*, także przewodniczącym Zespołu Techniki Morskiej STŚT KT PAN i Zespołu Środowiskowego SPE KBM PAN.

IDENTYFIKACJA STANU ZDATNOŚCI SYSTEMU

Jerzy JAŻWIŃSKI, Sławomir KLIMASZEWSKI

Instytut Techniczny Wojsk Lotniczych
01-494 Warszawa, ul. Księcia Bolesława 6
tel./fax: (0 22) 685 163

Streszczenie

Pod pojęciem identyfikacji stanu zdolności systemu rozumiemy relacje $C \subset W$ (zbiór cech C zawiera się w zbiorze wymagań W). Cechy mogą być zdeterminowane – c , zmienne losowe – C , procesy losowe – $C(t)$. Wymagania mogą być zdeterminowane – w , zmienne losowe W , procesy losowe $W(t)$. Zbiór cech i zbiór wymagań wyznaczają dziewięć różnych relacji. W referacie omówiono wszystkie relacje, podając ich modele matematyczne oraz interpretację.

Słowa kluczowe: identyfikacja stanu, cecha, wymaganie, zmienna losowa, proces losowy.

IDENTIFICATION OF THE ABILITY STATE OF THE SYSTEM

Summary

Under the notion of identification of the ability state of the system we understand the relations $C \subset W$ (the set of characteristics C is included in the set of requirements W). The characteristics can be determined – c , random variables – C , random processes – $C(t)$. The requirements can be determined – w , random variables W , random processes $W(t)$. The set of characteristics and the set of requirements determine nine different relations. All the relations have been considered in the paper as well as their mathematical models and interpretation have been given.

Keywords: identification of the state, characteristic, requirements, random variable, random process.

1. OGÓLNE PROBLEMY IDENTYFIKACJI

Ludzie w czasie swojej egzystencji muszą podejmować wiele różnorodnych decyzji. Dotyczą one bezpośrednio osoby, która je podejmuje, a odnoszą się do relacji z innymi ludźmi, ze środowiskiem naturalnym, techniką itp., tworząc układy mogące znajdować się w wielu różnych stanach. Podejmowane decyzje mają na celu zachowanie określonych stanów układu lub wprowadzenie zmian tych stanów. Aby dokonać zmiany stanu układu należy go w pierwszej kolejności zidentyfikować. Każdy stan układu wyznaczony jest przez zbiór wymagań $W_i[w_{i1}, w_{i2}, \dots, w_{ij}, \dots, w_{im_i}]$ odniesiony do zbioru cech układu $C_i[C_{i1}, \dots, C_{ik}, \dots, C_{im_i}]$. Układ znajduje się w i -tym stanie jeżeli zbiór cech C_i spełnia wymagania W_i , to znaczy $C_i \subset W_i$; $i = 1, 2, \dots, N$. W praktyce, człowiek podejmujący decyzje nie zna pełnego zbioru cech układu oraz wymagań wyznaczających stany układu.

Po dokładnym przyjrzeniu się opisanym przypadkom, można zaobserwować pewną prawidłowość:

- Do podjęcia decyzji o zachowaniu lub zmianie układu, niezbędna jest identyfikacja tego stanu (Umownie można powiedzieć – pomiar stanu.).

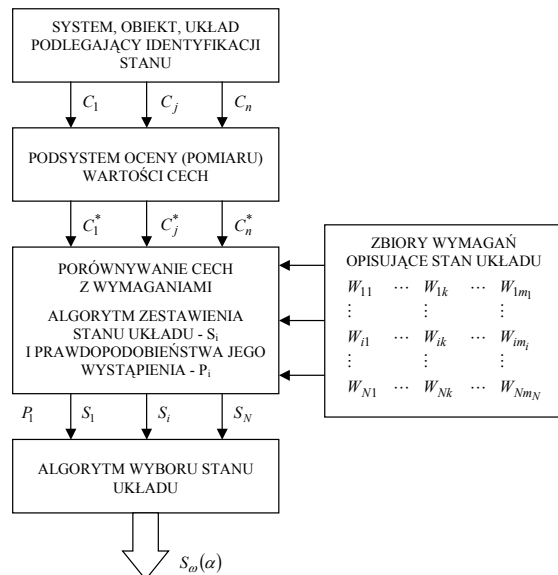
- Obserwuje się różnorodne ilościowe i jakościowe cechy układu. Wartości niektórych cech zobrazowane są przy pomocy przyrządów pomiarowych. Niektóre cechy są obserwowane bezpośrednio lub pośrednio przez nasze zmysły.
- Odebraną informację porównuje się z różnorodnymi zestawami wymagań opisujących różne stany układu.
- W wyniku dokonanych porównań wyznacza się stan układu.
- Zebrana informacja o cechach oraz wymaganiach jest niepełna do określenia stanu układu. Stan układu zostaje wyznaczony z określonym błędem. Zazwyczaj wyznacza się podzbiór stanów.

Z przytoczonego algorytmu postępowania wynika, że do pomiaru wyznaczania stanu układu niezbędny jest specyficzny system pomiarowy. Specyfika tego systemu upoważnia do użycia umownego terminu metasytem pomiarowy.

Metasytem pomiarowy składa się z podsystemu oceny ilościowych i jakościowych cech układu, banku zestawu wymagań opisujących stan układu, zestawu algorytmów porównania cech z wymaganiami, zestawu algorytmów podejmowania decyzji przy niepełnej informacji o stanie układu z określonym błędem. Na rys. 1 przedstawiono jeden z możliwych wariantów

metasystemu pomiarowego dokonującego identyfikacji stanu systemu.

Przykładem metasystemu pomiarowego w szczególnym przypadku jest operator technicznego środka transportu (samolotu, samochodu) wyposażony w zmysły oraz określone mierniki pokładowe.



Rys. 1. Przykładowe zobrazowanie metasystemu pomiarowego wyznaczającego stan układu z błędem α

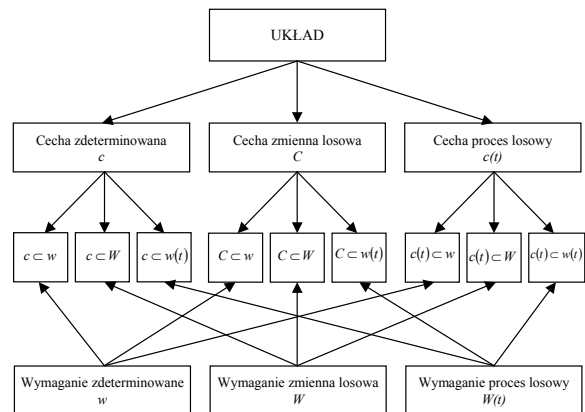
Istotnym problemem MSP jest zbiór cech ilościowych i jakościowych charakteryzujących układ. Cechom ilościowym mogą być przypisane określone wartości, natomiast cechy jakościowe nie mogą w sposób bezpośredni być opisane ilościowo. Przykładem cech jakościowych jest: kształt, kolor, zapach itp. Podział ten nie jest jednak precyzyjny. W technicznych środkach transportu, dla istotnych cech zamontowane są przyrządy pomiarowe, np. prędkościomierz. Operator z obserwacji technicznego środka transportu odbiera różne sygnały (wibracje, sygnały dźwiękowe, zapachowe itp.), świadczące o niepożądanych procesach i wymagające określonych interwencji.

Są to wartości zdeterminowane, zmienne losowe i procesy losowe. Ogólnie można powiedzieć, że w stanie statycznym układu - cechy charakteryzujące układ mogą być w przybliżeniu traktowane jako wartości zdeterminowane, w stanie dynamicznym układu - cechy te nabierają zazwyczaj charakteru zmiennych losowych lub procesów losowych. Sytuacja się komplikuje jeżeli rozpatrywany jest układ jako reprezentacja pewnego zbioru układów i mierzone cechy są reprezentacją zbiorowości cech układu. W takim przypadku nabierają one charakteru losowego.

Charakter cechy w wielu przypadkach wyznacza wymagania. Wymagania mogą być zdeterminowane lub zmiennymi losowymi i procesami losowymi. Na rys. 2 przedstawiono różne relacje zachodzące pomiędzy cechami i wymaganiami. Każdej

rozpatrywanej relacji przynależności $C \subset W$ towarzyszy złożony algorytm jej realizacji. W procesie wykorzystania relacji mogą być popełnione dwa rodzaje błędów:

- Błąd pierwszego typu polega na przyjęciu, że pomierzona cecha C^* nie spełnia wymagań $C^* \not\subset W$, gdy w rzeczywistości cecha ta spełnia wymagania $C \subset W$.
- Miarą błędu pierwszego typu jest prawdopodobieństwo $\alpha = P(C^* \not\subset W | C \subset W)$.
- Błąd drugiego typu polega na przyjęciu, że pomierzona cecha C^* spełnia wymagania $C^* \subset W$, gdy w rzeczywistości nie spełnia ona wymagań $C \not\subset W$.
- Miarą błędu drugiego typu jest prawdopodobieństwo $\beta = P(C^* \subset W | C \not\subset W)$.



Rys. 2. Relacja przynależności pomiędzy różnymi rodzajami cech i wymaganiami

Błędem pierwszego i drugiego typu towarzyszą bezbłędności pierwszego i drugiego typu:

- Bezbłędność pierwszego typu polegająca na właściwym zakwalifikowaniu pomierzonej cechy C^* spełniającej wymagania: $C^* \subset W$, dane jest wzorem:

$$\bar{\alpha} = 1 - \alpha = P(C^* \subset W | C \subset W).$$

- Bezbłędność drugiego typu polegająca na właściwym zakwalifikowaniu mierzonej cechy C^* nie spełniającej wymagań: $C^* \not\subset W$, dane jest wzorem:

$$\bar{\beta} = 1 - \beta = P(C^* \not\subset W | C \not\subset W).$$

Reasumując można powiedzieć:

1. Identyfikacja stanu: systemu, układu, urządzenia, jest najbardziej rozpowszechnionym działaniem człowieka.
2. Utożsamianie identyfikacji stanu z pomiarem pozwala wykorzystać z jednej strony bogaty dorobek teorii systemów, a z drugiej strony bogaty dorobek teorii pomiarów.

3. Koniecznością odróżnienia identyfikacji stanu od pomiaru jest wprowadzenie terminu metasystem pomiarowy.
4. W odróżnieniu od dobrze zdefiniowanego pojęcia systemu pomiarowego, metasystem pomiarowy nie może być dobrze zdefiniowany. Wynika to z tego, że nie dysponujemy pełnym zbiorem cech i wymagań, opisujących stan systemu oraz posiłkujemy niekompletnymi algorytmami sprawdzenia relacji zgodności cech z wymaganiami. Zazwyczaj nie posiadamy definicji wszystkich możliwych stanów układu, zdefiniowane są tylko istotne stany względnie podzbiory stanów. Z powyżej wymienionych powodów metasystem pomiarowy funkcjonuje na dużym poziomie niepewności.
5. Wydaje się, że koncepcja MSP zasługuje na uwagę. Istotne znaczenie mają prace częściowo rozwiązujące określone problemy, np. identyfikacja podzbioru stanów w sensie „zdalny – niezdalny”, „funkcjonuje – nie funkcjonuje” itp.

Można rozważyć następujące przykłady relacji pomiędzy cechami i wymaganiami:

1. Wymijanie - oznacza manewr, podczas którego mijają się pojazdy jadące po tej samej drodze, lecz w przeciwnych kierunkach. Wymijanie oznacza również przechodzenie obok użytkowników podążających w przeciwnych kierunkach. Podstawową dyrektywą wymijania jest zachowanie bezpiecznej odległości od wymijanego użytkownika drogi. Ścisłej definicji podać nie można. Problem należy do kategorii techniczno – taktycznej. W pracy [1] podaje się boczny odstęp w czasie wymijania. Nie powinien on być mniejszy od tylu centymetrów ile wynosi szybkość względna pomiędzy wymijającymi się pojazdami, a więc tyle centymetrów ile wynosi suma szybkości obydwóch pojazdów w km/h. Przykładowo jeżeli pojazdy mijają się z prędkością 50 km/k to odległość wymijania nie powinna być mniejsza od 100 cm. W przypadku wymijania jako cechę przyjmuje się prędkość obu pojazdów, którą ze względu na brak informacji o pojeździe wymijanym można potraktować jako zmienną losową. Wymaganie w takiej sytuacji można również potraktować jako zmienną losową.

2. Omijanie - przejeżdżanie jak również przechodzenie obok znajdujących się na drodze pojazdów lub innych nie poruszających się użytkowników drogi lub przedmiotów. O ile wyprzedzanie lub wymijanie dotyczy innych uczestników ruchu, o tyle omijanie może dotyczyć innych znajdujących się na drodze przedmiotów. Omijany przedmiot, pojazd czy pieszy nie może znajdować się w ruchu, gdyż wówczas nastąpiłoby wyprzedzanie lub wymijanie, a nie omijanie. Za moment omijania przyjmuje się chwilę, gdy omijający znajduje się na tej samej wysokości co omijany użytkownik drogi, względnie omijany przedmiot (słup, latarnia, kamień).

W pracy [1] podano przykłady odległości omijania (tabela 1).

Tabela 1. Przykłady odległości omijania

Rzeczywista szybkość w km/h	80	60	40	30
Wielkość bezpiecznego odstępu w metrach	1	0,75	0,6	05

Odległości te powinny być zwiększone przy występowaniu ograniczonego zaufania (omijanie autobusów, pojazdów konnych, osób nietrzeźwych itp.). Problem omijania należy rozpatrywać w kategoriach doświadczenia kierowcy, według algorytmu „hamowanie – zwalnianie – omijanie”. W danym przypadku cechą jest prędkość, która może być traktowana jako wartość zdeterminowana, a wymaganiem jest odległość od przeszkody, oceniana wzrokowo.

2. RELACJE POMIĘDZY CECAMI I WYMAGANIAMI

2.1. Wymaganie zdeterminowane

Przy wymaganiach zdeterminowanych (*W*) cecha może być również zdeterminowana (*c*), zmienną losową (*C*), procesem losowym [*C(t)*]. Jest to najczęściej spotykany przypadek w praktyce. Urządzenia pomiarowe zamontowane na technicznych środkach transportu w wielu przypadkach posiadają na skali zaznaczony przedział. Jeżeli wartość cechy zawiera się w wartości przedziału, to uznaje się, że obiekt jest zdalny względnie obserwowane zjawisko przebiega prawidłowo. W przeciwnym przypadku przyjmuje się, że obiekt jest niezdalny.

Cecha zdeterminowana

Przypadek cechy zdeterminowanej dotyczy w zasadzie wymiarów obiektu, np. długości i szerokości stołu. Bardziej złożona sytuacja występuje przy opisywaniu cech geometrycznych technicznych środków transportowych, np. skrzydła samolotu.

Cecha zmienna losowa

Rozpatrywany jest element (przeład, urządzenie), dla którego związek sygnału wyjściowego *y* i sygnału wejściowego *x* jest określony zależnością ciągłą, najczęściej zależnością liniową:

$$y = cx \quad (1)$$

Jeżeli na wejściu takiego obiektu podawany jest sygnał o stałej wartości $x = x_1 = const$, to na wyjściu powinien występować również sygnał o wartości stałej, a mianowicie $y_1 = cx_1 = const$. Jednakże w wyniku zewnętrznych zmian obiektu wielkość *c* nie jest stała w czasie i może być większa lub mniejsza od ustalonej wartości średniej $\bar{C}$. Sygnał wyjściowy również będzie ulegał zmianom

w czasie, wokół pewnej średniej wartości $\bar{y}_1 = \bar{C} x_1$. Jak wykazuje doświadczenie, dla większości obiektów odchylenie wielkości y_1 od wartości średniej $\bar{y}_1$ określa się przez rozkład normalny.

Zwykle wymaga się aby obiekt pracował z określoną dokładnością, to jest aby sygnał wyjściowy obiektu nie miał większych odchyżeń od zadanej wartości niedopuszczalnej $\pm \Delta$. Prawdopodobieństwo tego, że wartość sygnału wyjściowego y_1 będzie znajdowała się w granicach dopuszczalnych jeżeli rozkład wartości sygnału wyjściowego przyjmiemy za normalny, wynosi:

$$R = \frac{1}{\sigma\sqrt{2\pi}} \int_{\bar{y}_1 - \Delta}^{\bar{y}_1 + \Delta} \exp\left[-\frac{(y_1 - \bar{y}_1)^2}{2\sigma^2}\right] dy \quad (2)$$

gdzie: σ - odchylenie standardowe wartości y .

Cecha – proces losowy

Oznaczmy proces losowy obciążenia konstrukcji $C(t)$. Zakładamy, że wytrzymałość konstrukcji w jest stała. Warunkiem bezpieczeństwa konstrukcji jest nieprzekroczenie przez proces $U(t) = C(t) - w$ poziomu zerowego. W wielu przypadkach ważny jest czas trwania τ przekroczenia poziomu zerowego.

W ogólnym przypadku bezpieczeństwo konstrukcji można wyznaczyć ze wzoru [6]:

$$R_B(t_0) = R_B(0) \{1 - [1 - P_p(t_0)] P(\tau > \tau_{kr})\} \quad (3)$$

gdzie:

- $R_B(t)$ - prawdopodobieństwo nie zniszczenia konstrukcji do chwili t ;
- $R_B(0)$ - prawdopodobieństwo nie zniszczenia konstrukcji w chwili $t = 0$;
- P_p - prawdopodobieństwo nie wystąpienia poziomu zerowego w czasie t ;
- $P(\tau > \tau_{kr})$ - prawdopodobieństwo przekroczenia poziomu czasu krytycznego;
- τ_{kr} - czas trwania przekroczenia poziomu czasu krytycznego, po którym występuje zniszczenie konstrukcji.

Do wyznaczenia bezpieczeństwa konstrukcji niezbędna jest znajomość wartości wchodzących do wzoru (3).

Do wyznaczenia prawdopodobieństwa $P(\tau > \tau_{kr})$ niezbędna jest znajomość gęstości prawdopodobieństwa $f(\tau)$. Wykazuje się, że gęstość prawdopodobieństwa czasu przekroczenia poziomu zerowego jest w przybliżeniu opisana rozkładem Rayleigha o postaci [4]:

$$f(\tau) = \frac{\tau}{\sigma^2} \exp\left[-\frac{\tau^2}{2\sigma^2}\right] \quad (4)$$

$$P_\tau = P(\tau > \tau_{kr}) = \int_{\tau_{kr}}^{\infty} f(\tau) d\tau = \exp\left[-\frac{\pi \tau_{kr}^2}{4\bar{\tau}^2}\right] \quad (5)$$

$$\text{gdzie: } \bar{\tau} = \sigma \sqrt{\frac{\pi}{12}}.$$

Jeżeli znana jest średnia liczba przekroczeń poziomu zerowego $\bar{n}_0(t_0)$, to prawdopodobieństwo $P_p(t)$ dane jest wzorem:

$$P_p(t_0) = \exp\left[-\int_0^{t_0} \bar{n}_0(\tau) d\tau\right] \quad (6)$$

W ogólnym przypadku wyznaczenie prawdopodobieństw P_τ, P_p następuje wielu trudności.

W praktyce spotykamy się z procesami deterministycznymi. Są to funkcje o losowych argumentach zależnych od czasu. Procesy te są niestacjonarne. Rozważmy deterministyczny proces liniowy o postaci [5]:

$$\omega(t) = At + B \quad (7)$$

A i B to niezależne zmienne losowe. W przypadku gdy parametr $A = a$ jest zdeterminowany, a parametr B jest zmienną losową, wówczas proces ma przebieg równomierny. Gdy parametr A jest losowy, a parametr $B = b$ jest zdeterminowany, to proces ma przebieg wachlarzykowy.

Jeżeli przyjmiemy, że parametry A i B mają rozkłady normalne o parametrach $N(m_a, \sigma_a)$, $N(m_b, \sigma_b)$, to proces zużycia również ma rozkład normalny o parametrach $N(m_\omega, \sigma_\omega)$, gdzie:

$$m_\omega = m_a t + m_b \quad (8)$$

$$\sigma_\omega^2 = t^2 \sigma_a^2 + \sigma_b^2 \quad (9)$$

Jeżeli h jest dopuszczalnym poziomem zużycia, to:

$$R(t) = P(T > t) = P(AT + B \leq h) \quad (10)$$

Z normalności procesu zużycia $\omega(t)$ wynika [2]:

$$P(T \leq t) = 1 - R(t) = \Phi\left[\frac{t - \frac{h - m_b}{m_a}}{\sqrt{\frac{\sigma_a^2 t^2 + \sigma_b^2}{m_a^2}}}\right] \quad (11)$$

Podstawiając:

$$d = \frac{\sigma_a^2}{m_a^2}; \quad e = \frac{\sigma_b^2}{m_a^2}; \quad c = \frac{h - m_b}{m_a}, \quad (12)$$

otrzymuje się:

$$F(t) = \Phi\left[\frac{t - c}{\sqrt{d t^2 + e}}\right] \quad (13)$$

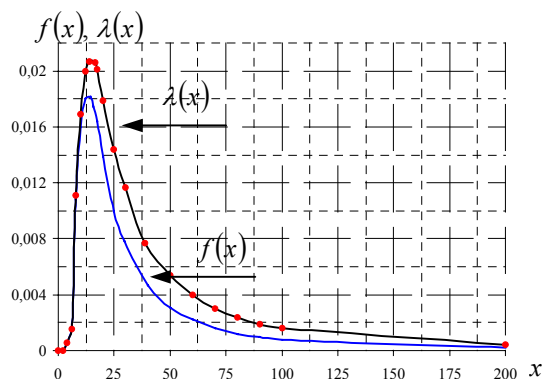
Funkcja gęstości prawdopodobieństwa dana jest wzorem:

$$f(t) = \frac{(e + c d t)}{[e + d t^2]^{3/2} \sqrt{2\pi}} \exp\left[-\frac{(t - c)^2}{2(d t^2 + e)}\right] \quad (14)$$

a intensywność uszkodzeń:

$$\lambda(t) = \frac{f(t)}{1 - F(t)} \quad (15)$$

Zobrazowanie funkcji gęstości prawdopodobieństwa $f(t)$ i funkcji intensywności uszkodzeń $\lambda(t)$ przedstawiono na rys. 3.



Rys. 3. Wykres funkcji gęstości prawdopodobieństwa $f(x)$ i funkcji intensywności zdarzeń $\lambda(x)$ dla liniowego procesu deterministycznego o normalnym rozkładzie parametrów $N(m_a = 0,01; \sigma_a = 0,05)$ oraz $N(m_b = 1; \sigma_b = 0,2)$ i wartości granicznej $h = 2$

2.2. Wymaganie zmienna losowa

W wielu przypadkach wymagania są zmiennymi losowymi. Taka sytuacja występuje wówczas, gdy cechą jest obciążenie, a wymaganiem jest wytrzymałość konstrukcji. Uwzględniając niejednorodność materiału, wytrzymałość jest zmienną losową. W innym przypadku gdy cechą jest czas wykonywania działania, wymaganiem wówczas jest czas dyspozycyjny, który jest zmienną losową na skutek różnorodności sytuacji, która działaniu towarzyszy. W rozpatrywanych przypadkach cecha może być zdeterminowana, może być zmienną losową względnie procesem losowym.

Cecha zdeterminowana

Przykładem takiej sytuacji jest stałe obciążenie działające na konstrukcję o losowej wytrzymałości. Rozważmy parę przykładów porównywania cechy c z wymaganiami W . Miara realizacji pomiędzy cechą i wymaganiem jest prawdopodobieństwo $P(c < W)$, że wytrzymałość jest większa od obciążenia.

Rozważmy parę przykładów:

1. Wytrzymałość ma rozkład Weibulle'a o postaci:

$$F(w) = 1 - \exp \left[- \left(\frac{w - \delta}{\Theta - \delta} \right)^\alpha \right] \quad (16)$$

Prawdopodobieństwo $F(w) = P(W \leq c)$ dla $\alpha = 4$, $\Theta = 2000$, $\delta = 1000$ i $c = 1500$ wynosi: $F(1500) = 0,061$ oraz dla $F(1100) = 0,000099$ lub, że $R(w) = 1 - F(w) = P(W > c) = 0,9999$.

2. Wytrzymałość ma rozkład gamma o postaci:

$$F(w) = 1 - \sum_{k=0}^{m-1} \frac{(\lambda w)^k \exp(-\lambda w)}{k!} \quad (17)$$

Dla parametrów $m = 3$, $\lambda = 0,005$, $c = 24$,

$R(w) = P(W > c) = 0,9927$ wytrzymałość konstrukcji jest zmienną losową na skutek niejednorodności technologii wytwarzania.

W innym obszarze, na przykład sterowania obiektem, cechą jest czas wykonania operacji sterującej, w określonych warunkach może być przyjęty jako wartość zdeterminowana. Jako wymaganie może być przyjęty czas dyspozycyjny, w którym działanie sterujące winno być wykonane. Aby zadanie było wykonane, działanie sterujące winno być krótsze od czasu dyspozycyjnego. Przykładowo czas odpalenia rakiety ziemia – powietrze, dla zniszczenia celu powietrznego, można przyjąć za zdeterminowany (cecha układu). Przebywanie celu w strefie rażenia wyznacza czas dyspozycyjny (wymaganie). Relacje pomiędzy tymi wielkościami decydują o wykonaniu zadania.

Cecha zmienna losowa

Relacją pomiędzy cechą C i wymaganiem W jest prawdopodobieństwo $P(W \leq C)$ dane wzorem [4]:

$$F = 1 - R = \int_{-\infty}^{\infty} F_w(s) f_c(s) ds = P(W \leq C) \quad (18)$$

gdzie: F – prawdopodobieństwo, że cecha C jest większa od wymaganej. Jeżeli cecha C jest obciążeniem, a wymaganiem W – wytrzymałością f_c , to F jest miarą zawodności konstrukcji. Jeżeli cecha C ma sens pracy obiektu zabezpieczanego do zapotrzebowania na działanie obiektu zabezpieczającego, a wymaganiem jest czasem pracy do uszkodzenia obiektu zabezpieczającego, to prawdopodobieństwo F jest miarą niedyspozycyjności obiektu zabezpieczającego.

Jeżeli cecha C oznacza czas wykonania operacji sterującej, a wymaganiem W oznacza czas dyspozycyjny, to prawdopodobieństwo F oznacza nieskuteczność działania. Odpowiednio R oznacza bezpieczeństwo konstrukcji, dyspozycyjność obiektu zabezpieczającego, efektywność działania. Rozważmy parę przykładów wyznaczania prawdopodobieństwa R dla różnych rozkładów $f_c(c)$ i $f_w(w)$:

1. Cecha i wymaganie mają rozkłady normalne:

$$f_c(c) = \frac{1}{\sigma_c \sqrt{2\pi}} \exp \left[- \frac{1}{2} \left(\frac{c - \mu_c}{\sigma_c} \right)^2 \right] \quad (19)$$

$$f_w(w) = \frac{1}{\sigma_w \sqrt{2\pi}} \exp \left[- \frac{1}{2} \left(\frac{w - \mu_w}{\sigma_w} \right)^2 \right] \quad (20)$$

gdzie:

μ_c i σ_c – wartość oczekiwana i odchylenie standardowe cechy C ;

μ_w i σ_w – wartość oczekiwana i odchylenie standardowe wymagań.

Wprowadźmy zmienną losową $Y = W - C$. Wiadomym jest, że zmienna losowa Y posiada rozkład normalny z wartością oczekiwaną:

$$\mu_y = \mu_w - \mu_c \quad (21)$$

i odchyleniem standardowym:

$$\sigma_y = \sqrt{\sigma_w^2 + \sigma_c^2} \quad (22)$$

Wówczas prawdopodobieństwo R dane jest wzorem:

$$R = P(Y > 0) = \int_0^{\infty} \frac{1}{\sigma_y \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{y - \mu_y}{\sigma_y}\right)^2\right] dy \quad (23)$$

Powyższą zależność można napisać w postaci:

$$R = \frac{1}{\sqrt{2\pi}} \int_{\frac{\mu_w - \mu_c}{\sqrt{\sigma_w^2 + \sigma_c^2}}}^{\infty} \exp\left[-\frac{z^2}{2}\right] dz \quad (24)$$

Dla wyznaczenia wartości R można wykorzystać dane z istniejących tablic rozkładu normalnego.

2. Jeżeli cecha i wymaganie mają rozkłady wykładnicze o parametrach λ_c i λ_w , to:

$$R = \frac{\lambda_c}{\lambda_w + \lambda_c} \quad (25)$$

3. Jeżeli wymaganie ma rozkład normalny $N(\mu_w, \sigma_w)$, a cecha ma rozkład wykładniczy o parametrze λ_c , to:

$$R = 1 - \exp\left(\mu_w \lambda_c + \frac{1}{2} \lambda_c^2 \sigma_w^2\right) \quad (26)$$

W ogólnym przypadku prawdopodobieństwo R można wyznaczyć graficznie. Wprowadźmy oznaczenia:

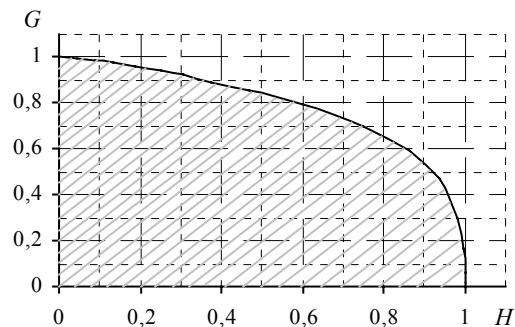
$$G = \int_w^{\infty} f_w(w) dw = 1 - F_w(w) \quad (27)$$

$$H = \int_0^c f_c(c) dc = F_c(c) \quad (28)$$

Wówczas:

$$R = \int_0^1 G dH \quad (29)$$

Powyższy wzór wskazuje, że prawdopodobieństwo R równe jest polu pod krzywą G i H . Relację ilustruje rys. 4.



Rys. 4. Wykres funkcji $G = f(H)$

Cecha proces losowy

W praktyce często spotykamy się z sytuacją kiedy cecha $C(t)$ jest procesem losowym, a wymaganie jest zmienną losową W . Rozważmy niektóre przykłady:

1. W czasie startu samolotu na konstrukcję oddziałują impulsy obciążenia. Wytrzymałość konstrukcji jest zmienną losową. Konstrukcja ulega uszkodzeniu jeżeli impuls obciążenia przekroczy wytrzymałość konstrukcji.
2. Napięcie zasilania energią elektryczną na samolocie jest procesem losowym. Wynika to z losowej liczby odbiorników włączanych w danym momencie oraz z oddziaływania różnych czynników wymuszających na odbiorniki energii elektrycznej. Wymaganie jest również zmienną losową, uwarunkowaną odpornością izolacji na zwarcia różnych odbiorników energii elektrycznej.
3. Czas wykonywania działania sterującego w wielu przypadkach może być traktowany jako proces losowy. Wynika to między innymi ze stanu psychicznego operatora, jego podatności na stres informacyjny, czasowy oraz zmęczenia.

Miarą relacji „cecha $C(t)$ – wymaganie W ” jest prawdopodobieństwo:

$$R = P(W - C(t) > 0 \mid t \leq T) \quad (30)$$

Przyjmując, że wymaganie i cecha są nie skorelowane, znajdujemy [6]:

$$R \approx \exp\left\{-\frac{T}{2\pi} \sqrt{\left[\frac{d^2 \rho_c(\tau)}{d\tau^2}\right]_{\tau=0}} \exp\left[-\frac{(m_w - m_c)^2}{2(\sigma_w^2 + \sigma_c^2)}\right]\right\} \quad (31)$$

gdzie:

$\rho_c(\tau)$ - unormowana funkcja korelacji cechy $C(t)$;

m_w, m_c - wartości oczekiwane wymagań W i cechy $C(t)$;

σ_w^2, σ_c^2 - wariancja wymagań W i cechy $C(t)$.

Wyrażenie powyższe daje dolne oszacowanie prawdopodobieństwa R . Bardziej dokładne oszacowanie uzyskuje się ze wzoru:

$$R = \int_0^{\infty} f_w(x) \exp[-\bar{n}_w T] dx \quad (32)$$

gdzie:

$f_w(x)$ - gęstość prawdopodobieństwa zmiennej losowej W ;
 $\bar{n}_w$ - średnia liczba przekroczeń przez obciążenie poziomu W .

2.3. Wymaganie proces losowy

Wymaganie jako proces losowy występuje w wielu różnych sytuacjach. W przypadku gdy jako wymaganie przyjmuje się wytrzymałość konstrukcji, to na skutek zachodzących w konstrukcji procesów starzenia i zużycia, wytrzymałość konstrukcji jest zależna od czasu i winna być traktowana jako proces losowy. W czasie realizacji zadania transportowego wymaganie odnośnie takiej cechy jak prędkość jest procesem losowym, gdyż o prędkości decydują znaki drogowe, zachowanie się innych użytkowników dróg, stan dróg, warunki klimatyczne itp. W rozpatrywanych przypadkach cecha może być zdeterminowana np. stałe obciążenie konstrukcji, zmienną losową, procesem losowym, np. obciążenie jest zależne od czasu.

Cecha zdeterminowana

1. Cecha to stałe obciążenie konstrukcji. Wytrzymałość zależy od temperatury lub starzenia zachodzącego w materiale konstrukcji. Wytrzymałość jest procesem losowym.
2. Profesjonalne działanie wykonywane przez operatora w czasie sterowania obiektem można przyjąć jako zdeterminowane. Czas dyspozycyjny jako wymaganie zależy od czasu (pora dnia, warunki klimatyczne, itp.) i jest procesem losowym.

Rozważmy problem identyfikacji takiego układu. Obiekt spełnia wymagania jeżeli cecha nie przekroczy procesu opisującego wymagania. Obiekt nie spełnia wymagań jeśli cecha c przekroczy poziom procesu opisującego wymagania i przebywa w tym stanie przez czas τ .

Średnia liczba przekroczeń wymaganego poziomu przez cechę $\bar{n}$ i czas przebywania w stanie przekroczenia τ , dane są wzorami [6]:

$$\bar{n} = A \exp\left[-(c - m_w)^2 \mid K_w(0)\right] \quad (33)$$

$$\tau = \frac{1}{A} \left[1 - \phi\left(\frac{c - m_w}{\sqrt{K_w(0)}}\right) \right] \exp\left[\frac{(c - m_w)^2}{2K_w(0)}\right] \quad (34)$$

$$A = \frac{1}{2\pi} \sqrt{-\left[\frac{1}{K_w(\tau)} \cdot \frac{dK_w(\tau)}{d\tau^2}\right]_{\tau=0}} \quad (35)$$

Parametr A można wyrazić przez gęstość spektralną $S_w(\omega)$:

$$K_w(\tau) = \int_0^\infty S_w(\omega) \cos \omega \tau d\omega \quad (36)$$

gdzie: ω - częstość kątowna.

Przy założeniu, że przekroczenie przez cechę poziomu wymagań jest zdarzeniem

małoprawdopodobnym, to prawdopodobieństwo przekroczenia przez cechę poziomu wymagań dane jest wzorem:

$$R = \exp\left\{-AT \exp\left[-(c - m_w)^2 \mid 2K_w(0)\right]\right\} \quad (37)$$

lub:

$$R \approx e^{-\bar{n}T} \quad (38)$$

gdzie:

$\bar{n}$ - średnia liczba przekroczeń poziomu cechy C ;
 $K_w(\tau)$ - funkcja korelacyjna wymagań;
 $S_w(\omega)$ - gęstość spektralna;
 m_w - wartość oczekiwana wymagań;
 C - wartość cechy.

Cecha zmienna losowa

Sytuacja taka występuje gdy np. obciążenie jest zmienną losową, a wymaganie jest procesem losowym. Przyjmując funkcję gęstości cechy $f_c(x)$:

$$n_{sr} = \int_0^\infty f_c(c) \bar{n}(c) dc \quad (39)$$

$$\tau_{sr} = \int_0^\infty f_c(c) \tau(c) dc \quad (40)$$

$$R_{sr} = \int_0^\infty f_c(c) R(c) dc \quad (41)$$

gdzie:

n_{sr} , τ_{sr} , R_{sr} - odpowiednio: $\bar{n}(c)$ - dane wzorem (33), $\tau(c)$ - dane wzorem (34) i $R(c)$ - dane wzorem (37) dla średniej wartości cechy C .

Cecha proces losowy

Rozważmy przypadek ogólny, kiedy cecha i wymaganie są procesami losowymi. Rozważmy proces losowy:

$$U(t) = C(t) - W(t) \quad (42)$$

Rozważa się krótkotrwale przejścia przez zero funkcji $U(t)$. Krótkotrwale przekroczenie nie zawsze powoduje uszkodzenie systemu. Dlatego niezbędnym jest uwzględnienie czasu trwania przekroczenia. Zakłada się, że zniszczenie systemu następuje gdy $\tau > \tau_{kr}$ (τ_{kr} - krytyczne). Dla wyznaczenia prawdopodobieństwa $P(\tau > \tau_{kr})$ niezbędna jest znajomość funkcji gęstości prawdopodobieństwa $f(\tau)$. Wykazuje się, że gęstość prawdopodobieństwa czasu trwania przekroczenia w przybliżeniu opisuje się rozkładem Rayleigha w postaci:

$$f(\tau) = \frac{\tau}{\sigma^2} \exp\left(-\frac{\tau^2}{2\sigma^2}\right) \quad (43)$$

Parametr σ wyznacza się z zależności:

$$E[\tau] = \bar{\tau} = \sqrt{\frac{\pi}{12}} \sigma \quad (44)$$

Wartość $\bar{\tau}$ wyznacza się z doświadczenia. Przy ustalonym τ_{kr} prawdopodobieństwo P_τ dane jest wzorem:

$$P_{\tau} = P(\tau > \tau_{kr}) = \int_{\tau_{kr}}^{\infty} f(\tau) d\tau = \exp\left[-\frac{\pi \tau_{kr}^2}{4T^2}\right] \quad (45)$$

Ogólnie prawdopodobieństwo niezniszczenia obiektu w czasie T dane jest wzorem [6]:

$$P(T) = P(0)\{1 - [1 - P_1(\tau)]P_{\tau}\} \quad (46)$$

gdzie:

$$P_1(T) = P\{U(t) < 0 \mid t \leq T\} \quad (47)$$

Przy stacjonarnym procesie $U(t)$ prawdopodobieństwo $P_1(t)$ dane jest wzorem:

$$P_1 \approx \exp\left\{-\frac{T}{2\pi} \sqrt{-\left(\frac{d^2 \rho_U^{(\tau)}}{d\tau^2}\right)} \Big|_{\tau=0} \exp\left[-\frac{\sigma_U^2}{2\sigma_U^2}\right]\right\} \quad (48)$$

gdzie: σ_U^2 i $\rho_U^{(\tau)}$ - odpowiednio wariancja i unormowana funkcja korelacji procesu losowego $U(t)$, przy czym:

$$\sigma_U^2 = \sigma_C^2 + \sigma_W^2 - 2\sigma_C \sigma_W \sigma_{UC} \quad (49)$$

$$\rho_W(\tau) = \frac{1}{\sigma_W^2} \left[\sigma_C^2 \rho_C^{(\tau)} + \sigma_W^2 \rho_W - 2\sigma_C \sigma_W \rho_{CW}^{(\tau)} \right] \quad (50)$$

gdzie:

σ_C^2 i σ_W^2 - wariancja cechy C i wymagań W ;

$\rho_W^{(\tau)}$ i $\rho_{CW}^{(\tau)}$ - unormowana autokorelacja cechy C i wymagań W i unormowana funkcja korelacyjna procesu $C(t)$ i wymagań $W(t)$.

3. WNIOSKI

W procesie identyfikacji stanu systemu istotną rolę odgrywa metasytem pomiarowy.

Metasytem pomiarowy występuje wówczas gdy nie dysponujemy pełną informacją o zestawie cech, a wiele cech jest oszacowywanych przy pomocy zmysłów operatora. Dokonana identyfikacja jest przybliżona. Częste zagrożenia bezpieczeństwa w transporcie są skutkiem niedokładnej identyfikacji. Problem metasytemu pomiarowego wymaga dalszych badań..

LITERATURA

- [1] Dobrzyński A.: „Bez zielonego listka”. Krajowa Agencja Wydawnicza, 1982.
- [2] Gerbach I. B., Kordoński C. B.: „Modele niezawodności obiektów technicznych”. WNT, Warszawa 1968.
- [3] Jaźwiński J. i inni: Sprawozdanie nr 1389/50 z pracy pt.: „Identyfikacja stanu transportowego w procesie realizacji zadania transportowego”. ITWL, Warszawa 2003.
- [4] Kapur L., Lanbercov: „Nadexnost' i proektirovanie system”. Mir, Moskwa 1980.
- [5] Prohorenko V. A., Smirnov A. H.: „Prognozirowanie ka'estwa sistem”. „Nauka i tehnika”, Minsk 1976.
- [6] Volkov L. I., Wiwkevi' A. M.: „Nadexnost' letatel'nyh aparator”. Wyswa/ Wkola, 1975.



Prof. dr hab. inż. Jerzy JAŻWIŃSKI jest pracownikiem naukowym Zakładu Niezawodności i Bezpieczeństwa Instytutu Technicznego Wojsk Lotniczych w Warszawie. Zajmuje się problematyką niezawodności, bezpieczeństwa i logistyki systemów transportowych, ze szczególnym uwzględnieniem statków powietrznych. Jest autorem kilkuset publikacji, w tym kilkunastu książkowych. Ważniejsze monografie: Jaźwiński Jerzy, Borgoń Jan: „Niezawodność eksploatacyjna i bezpieczeństwo lotów”, WKiŁ, Warszawa 1989; Ważyńska-Fiok Krystyna, Jaźwiński Jerzy: „Niezawodność systemów technicznych”, PWN, Warszawa 1990; Jaźwiński Jerzy, Ważyńska-Fiok Krystyna: „Bezpieczeństwo systemów”, PWN, Warszawa 1993; Grabski Franciszek, Jaźwiński Jerzy: „Metody bayesowskie w niezawodności i diagnostyce”, WKiŁ, Warszawa 2001; Jaźwiński Jerzy, Grabski Franciszek: „Niektóre problemy modelowania systemów transportowych”, Instytut Technologii Eksploatacji, Radom 2003. Wypromował 15 doktorów.



Dr inż. Sławomir KLIMASZEWSKI jest pracownikiem naukowym Zakładu Niezawodności i Bezpieczeństwa Instytutu Technicznego Wojsk Lotniczych w Warszawie. Zajmuje się problematyką niezawodności, bezpieczeństwa, modelami zużycia techniki lotniczej. Jest autorem ponad stu prac publikowanych i niepublikowanych, które znalazły szerokie zastosowanie w praktyce, między innymi do przedłużania zasobów pracy techniki lotniczej. Jest współautorem monografii: „Symulacyjne metody badania bezpieczeństwa lotów”, Wydawnictwo Naukowe ASKON, Warszawa 1998. Był kierownikiem i wykonawcą wielu projektów badawczych przyznawanych przez Komitet Badań Naukowych.

METROLOGICZNE ASPEKTY DIAGNOZOWANIA

Wiktor OLCHOWIK

Instytut Systemów Elektronicznych, Wydział Elektroniki
Wojskowa Akademia Techniczna, ul. S. Kaliskiego 2, 00-908 Warszawa, fax (22) 68 39 125
e-mail: wolchowik@wel.wat.edu.pl

Streszczenie

W niniejszym artykule dokonano analizy wpływu niepewności pomiaru na wiarygodność diagnozy przy założeniach, że przedział dopuszczalnych wartości badanego sygnału jest ściśle określony a wszystkie relacje pomiędzy wynikiem badania a diagnozą eksploatacyjną są jednoznaczne.

Słowa kluczowe: pomiar, diagnostyka obiektów technicznych.

METROLOGICAL ASPECTS OF DIAGNOSING

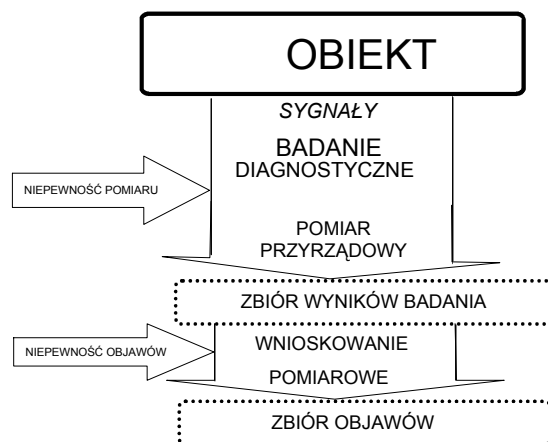
Summary

This article contains analyses of the influence the measurement's uncertainty on the reliability diagnoses assuming that: the range of allowed values measured signal is determined, all relations between the result of measurements and diagnosis are unequivocal.

Keywords: measure, technical system diagnostics.

1. WPROWADZENIE

Diagnozowanie jest złożonym, kilkustopniowym procesem przetwarzania pierwotnej informacji o wielkościach opisujących obiekt na informację o stanie obiektu, czyli na diagnozę. Na każdym etapie mogą występować czynniki powodujące niejednoznaczność tego przetwarzania. W efekcie końcowym powoduje to, że diagnoza jest obarczona niepewnością lub inaczej mówiąc charakteryzuje się obniżoną wiarygodnością. Szczególnie istotny wpływ na wiarygodność diagnozy mają początkowe etapy procesu diagnostycznego: badanie diagnostyczne i wnioskowanie pomiarowe (rys. 1).

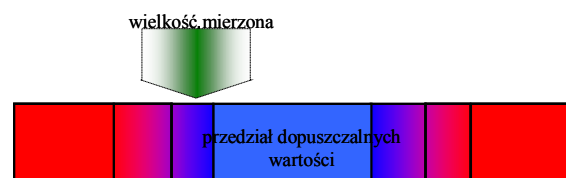


Rys. 1. Początkowe dwa etapy łańcucha działań w procesie diagnozowania

Badanie diagnostyczne realizowane jest najczęściej jako pomiar przyrządowy i obarczone jest niepewnością pomiaru. Rzeczywista wartość

sygnału diagnostycznego może więc być inna niż wartość zmierzona.

Wnioskowanie pomiarowe w najprostszym przypadku polega na porównaniu wyników badania diagnostycznego (czyli zmierzonych wartości sygnału diagnostycznego) z odpowiednimi przedziałami wartości dopuszczalnych dla wyróżnionych stanów obiektu. Przedziały te często nie są jednoznacznie określone – są przedziałami o wartościach brzegowych rozmytych (rys. 2). Stwierdzenie, że zmierzona wartość sygnału (czyli wynik badania) zawiera się w określonym przedziale wartości stanowi stwierdzenie istnienia określonego objawu stanu. Tym samym niepewność zaliczenia wyniku pomiaru do określonego przedziału przekłada się na niepewność stwierdzenia określonego objawu stanu.



Rys. 2. Ilustracja niepewności pomiaru w aspekcie niepewności objawu

Z wymienionego powodu cały proces wnioskowania diagnostycznego jest realizowany w warunkach informacji niepewnej, co wpływa na zmniejszenie wiarygodności diagnozy.

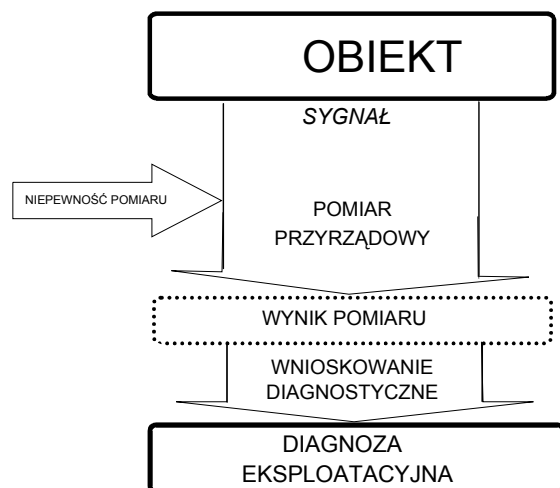
2. ZAŁOŻENIA

Teoretyczna analiza wiarygodności diagnozy nawet w aspekcie nałożenia się dwóch czynników niepewności jest stosunkowo skomplikowana. Można natomiast przeprowadzić analizę wpływu każdego z czynników osobno.

W celu przeprowadzenia obserwacji wpływu niepewności pomiaru na wiarygodność diagnozy przyjęto następujące założenia upraszczające:

1. Badanie diagnostyczne polega na pomiarze tylko jednej wielkości fizycznej.
2. Relacja **sygnał – wynik pomiaru** jest obciążona niepewnością pomiaru.
3. Niepewność pomiaru jest znana.
4. Przedział dopuszczalnych wartości badanego sygnału jest ściśle określony.
5. Wszystkie relacje pomiędzy wynikiem badania a diagnozą eksploatacyjną są jednoznaczne.

W takim przypadku schemat łańcucha działań diagnostycznych można przedstawić jak na rys. 3.



Rys. 3. Schemat uproszczonego łańcucha działań w procesie diagnozowania

Przy tak sformułowanych założeniach można przyjąć, że:

1. Przedział dopuszczalnych wartości jest przedziałem zdatności obiektu.
2. Jeżeli wartość rzeczywista wielkości mierzonej mieści się w przedziale dopuszczalnych wartości, to obiekt jest zdatny.
3. Zdatność obiektu jest określona z prawdopodobieństwem zależnym od warunków pomiaru.

3. ANALIZA PRAWDOPODOBIENSTWA ZDATNOŚCI

Przy przyjętych powyżej założeniach gęstość prawdopodobieństwa wystąpienia rzeczywistej

wartości sygnału badanego dla pojedynczego pomiaru wyraża się wzorem:

$$p(x, x_o, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-x_o)^2}{2\sigma^2}}$$

gdzie:

x – wartość rzeczywista sygnału badanego;

x_o – wartość zmierzona sygnału badanego;

σ – niepewność standardowa pomiaru;

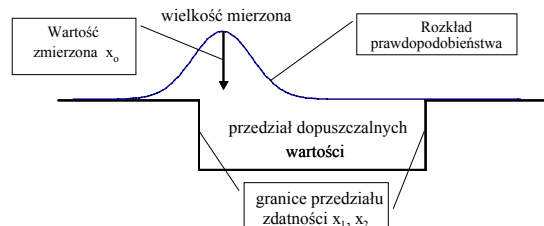
Prawdopodobieństwo zdatności będzie natomiast zależne od dopuszczalnych wartości wielkości mierzonej, czyli od charakteru przedziału zdatności. Dla przedziału zdatności ograniczonego dwustronnie prawdopodobieństwo zdatności wyraża się wzorem:

$$P(x, x_o, \sigma) = \int_{x_1}^{x_2} p(x, x_o, \sigma) dx$$

gdzie:

x_1, x_2 – wartości graniczne przedziału dopuszczalnych wartości sygnału badanego (przedziału zdatności);

Powyższy przypadek można zilustrować graficznie jak na rys. 4.

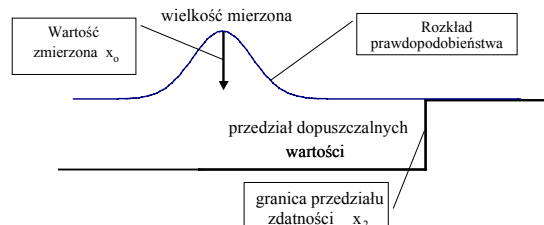


Rys. 4. Ilustracja wpływu niepewności pomiaru na prawdopodobieństwo zdatności dla przedziału zdatności ograniczonego dwustronnie

Dla przedziału zdatności ograniczonego jednostronnie od góry prawdopodobieństwo zdatności wyraża się wzorem:

$$P(x, x_o, \sigma) = \int_{-\infty}^{x_2} p(x, x_o, \sigma) dx$$

Powyższy przypadek można zilustrować graficznie jak na rys. 5.

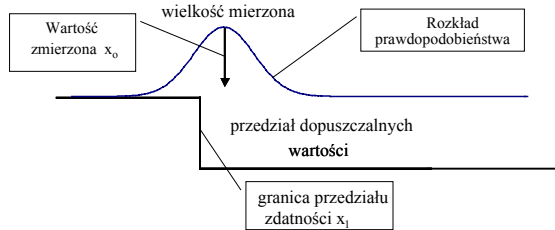


Rys. 5. Ilustracja wpływu niepewności pomiaru na prawdopodobieństwo zdatności dla przedziału zdatności ograniczonego od góry

Dla przedziału zdatności ograniczonego jednostronnie od dołu prawdopodobieństwo zdatności wyraża się wzorem:

$$P(x, x_o, \sigma) = \int_{x_1}^{\infty} p(x, x_o, \sigma) dx$$

Powyższy przypadek można zilustrować graficznie jak na rys. 6.



Rys. 6. Ilustracja wpływu niepewności pomiaru na prawdopodobieństwo zdatności dla przedziału zdatności ograniczonego od dołu

4. PRAWDOPODOBIEŃSTWO ZDATNOŚCI DLA PRZEDZIAŁU ZDATNOŚCI OGRANICZONEGO DWUSTRONNIE

Dla przedziału zdatności o wartościach granicznych x_1 i x_2 możemy określić następujące wielkości:

- wartość środkową przedziału zdatności

$$x_s = \frac{x_1 + x_2}{2},$$

- względne odchylenie od wartości środkowej przedziału zdatności zdefiniowane jako stosunek wartości bezwzględnej odchylenia do połowy długości przedziału zdatności

$$d_x = 2 \frac{x_o - x_s}{x_2 - x_1},$$

- względną niepewność pomiaru jako stosunek bezwzględnej niepewności pomiaru do połowy długości przedziału zdatności

$$\sigma_r = 2 \frac{\sigma}{x_2 - x_1}.$$

Wówczas gęstość prawdopodobieństwa zdatności wyraża się wzorem:

$$p(d_x, \sigma_r) = \frac{1}{\sqrt{2\pi}\sigma_r} e^{-\left(\frac{d_x^2}{2\sigma_r^2}\right)}$$

a prawdopodobieństwo zdatności wyraża się wzorem:

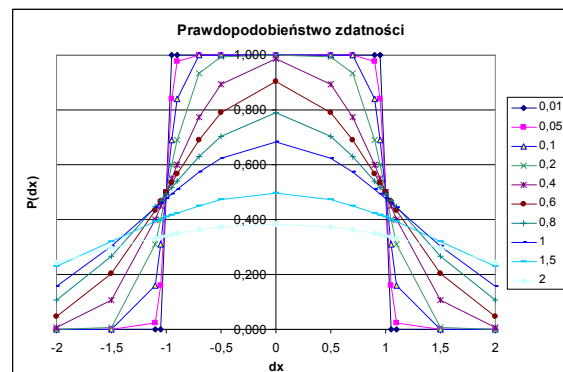
$$P(d_x, \sigma_r) = \int_{x_1}^{x_2} p(d_x, \sigma_r) dd_x$$

Korzystając z przedstawionych wzorów można obliczyć prawdopodobieństwo zdatności dla poszczególnych przypadków względnego odchylenia pomierzonej wartości od środka przedziału zdatności i względnej niepewności pomiaru. Wyniki symulacji cyfrowej przedstawiono w poniższej tabeli.

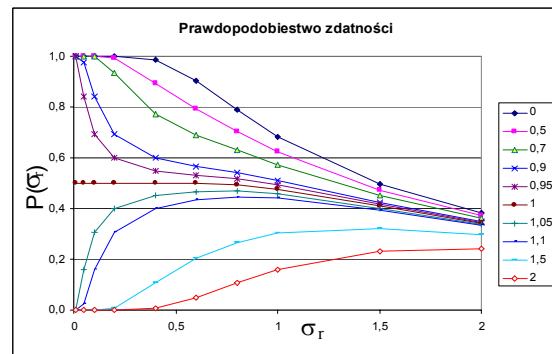
Tabela 1. Prawdopodobieństwo zdatności

		względna niepewność pomiaru - σ_r						
		0,01	0,05	0,1	0,2	0,4	0,6	0,8
względne odchylenie od wartości środkowej - d_x	-2	0,000	0,000	0,000	0,000	0,006	0,048	0,106
	-1,5	0,000	0,000	0,000	0,006	0,106	0,202	0,265
	-1,1	0,000	0,023	0,159	0,309	0,401	0,434	0,446
	-1,05	0,000	0,159	0,309	0,401	0,450	0,466	0,470
	-1	0,000	0,000	0,000	0,000	0,000	0,000	0,006
	-0,95	0,000	0,159	0,309	0,401	0,450	0,467	0,482
	-0,9	0,000	0,023	0,159	0,309	0,401	0,435	0,459
	-0,7	0,000	0,000	0,001	0,067	0,227	0,311	0,371
	-0,5	0,000	0,000	0,000	0,006	0,106	0,209	0,296
	0	0,000	0,000	0,000	0,000	0,012	0,096	0,211
	0,5	0,000	0,000	0,000	0,006	0,106	0,209	0,296
	0,7	0,000	0,000	0,001	0,067	0,227	0,311	0,371
	0,9	0,000	0,023	0,159	0,309	0,401	0,435	0,459
	0,95	0,000	0,159	0,309	0,401	0,450	0,467	0,482
	1	0,000	0,000	0,000	0,000	0,000	0,000	0,006
	1,05	0,000	0,159	0,309	0,401	0,450	0,466	0,470
	1,1	0,000	0,023	0,159	0,309	0,401	0,434	0,446
	1,5	0,000	0,000	0,000	0,006	0,106	0,202	0,265
	2	0,000	0,000	0,000	0,000	0,006	0,048	0,106

Natomiast wykresy prawdopodobieństwa zdatności jako funkcji względnego odchylenia od środka przedziału zdatności oraz względnej niepewności pomiaru przedstawiają odpowiednio wykresy na rys. 7 i 8.



Rys. 7. Wykres prawdopodobieństwa zdatności w funkcji względnego odchylenia od środka przedziału zdatności dla różnych wartości względnej niepewności pomiaru



Rys. 8. Wykres prawdopodobieństwa zdatności w funkcji względnej niepewności pomiaru dla różnych wartości względnego odchylenia od środka przedziału zdatności

Można założyć, że obiekt diagnozowany jest zdalny, gdy wartość wielkości badanej znajduje się wewnątrz przedziału zdalności oraz niezdalny, gdy na zewnątrz przedziału. Natomiast w przypadku, gdy wartość wielkości badanej przyjmuje dokładnie wartość graniczną przedziału zdalności można przyjąć, że prawdopodobieństwo zdalności obiektu diagnozowanego wynosi $\frac{1}{2}$. Czyli oczekiwane prawdopodobieństwo zdalności P_0 bez uwzględniania wpływu niepewności pomiaru, można przedstawić w postaci wzoru:

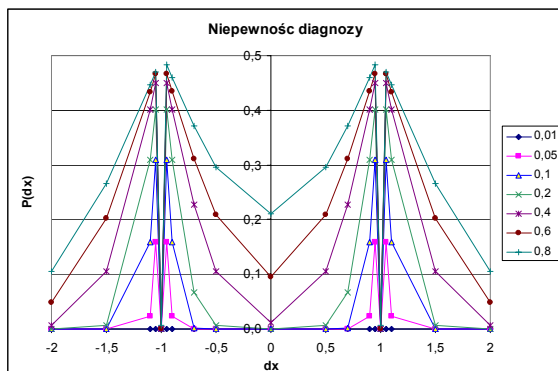
$$P_o(x) = \begin{cases} 1 & \text{dla } x_1 < x < x_2 \\ \frac{1}{2} & \text{dla } x = x_1 \text{ lub } x = x_2 \\ 0 & \text{dla } x < x_1 \text{ lub } x > x_2 \end{cases} \quad \text{lub}$$

$$P_o(d_x) = \begin{cases} 1 & \text{dla } -1 < d_x < 1 \\ \frac{1}{2} & \text{dla } d_x = -1 \text{ lub } d_x = 1 \\ 0 & \text{dla } d_x < -1 \text{ lub } d_x > 1 \end{cases}$$

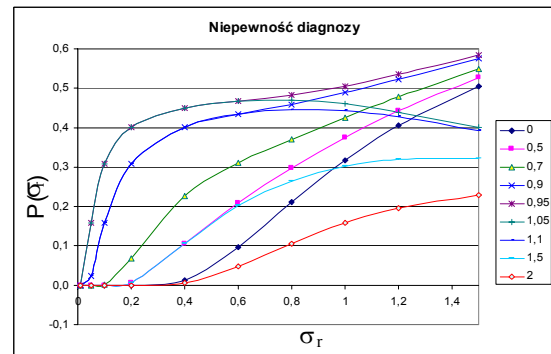
Natomiast niepewność diagnozy P_σ , rozumiana jako prawdopodobieństwo niewłaściwej diagnozy, może być reprezentowana przez wartość bezwzględną różnicy prawdopodobieństwa zdalności bez uwzględniania niepewności pomiaru i prawdopodobieństwa zdalności obiektu z uwzględnieniem niepewności pomiaru

$$P_\sigma = |P_o(d_x) - P(d_x, \sigma_r)|$$

Stąd prawdopodobieństwo niepewności diagnozy w funkcji niepewności pomiaru i względnej wartości wielkości mierzonej przedstawiono na wykresach rys. 9. i rys. 10.



Rys. 9. Wykres prawdopodobieństwa niepewności diagnozy P_σ w funkcji względnego odchylenia od środka przedziału zdalności



Rys. 10. Wykres prawdopodobieństwa niepewności diagnozy P_σ w funkcji względnej niepewności pomiaru

Analogicznie można dokonać analizy niepewności diagnozy dla przypadków, gdy przedział zdalności jest tylko jednostronnie ograniczony.

5. WNIOSKI

Na podstawie przedstawionych wykresów można zauważyć, że:

- największa niepewność diagnozy występuje w przypadku, gdy wartość pomierzona sygnału diagnostycznego jest bliska wartości granicznej przedziału zdalności;
- dla niepewności pomiaru przekraczającej 10% przedziału zdalności niepewność diagnozy jest zbyt duża, by mogła stanowić podstawę do oceny zdalności procesów krytycznych.

LITERATURA

- [1] Dąbrowski T.: *Metrologiczno-informacyjne aspekty badań diagnostycznych*. Diagnostyka, PTDT, Vol. 26, 2002, ss. 101-106.
- [2] Będowski L., Dąbrowski T.: *Diagnozowanie na podstawie niepewnych syndromów stanu obiektu*. Diagnostyka, PTDT, Vol. 37, 2006. ss. 55÷60.



Wiktor OLCHOWIK ukończył studia na Wydziale Elektroniki WAT w 1985 r. Od tego czasu jest pracownikiem naukowo-dydaktycznym w Instytucie Systemów Elektronicznych WAT. W 1999 r. uzyskał tytuł doktora nauk technicznych. Jego zainteresowania zawodowe obejmują obszary: automatyka, diagnostyka techniczna, komputerowe systemy pomiarowe i diagnostyczne, cyfrowe przetwarzanie sygnałów.

DIAGNOSTYKA IZOLOWANEGO SYSTEMU ELEKTROENERGETYCZNEGO NA PRZYKŁADZIE JEDNOSTKI PŁYWAJĄCEJ

Janusz MINDYKOWSKI¹, Tomasz TARASIUK¹, Edward SZMIT², Daniel CZARKOWSKI²

¹ – Akademia Morska w Gdyni, Katedra Elektroenergetyki Okrętowej;

² – Polski Rejestr Statków S.A., Inspektorat Elektrotechniki i Automatyki,
al. gen. Józefa Hallera 126, fax 346 03 92, E.Szmit@prs.pl, D.Czarkowski@prs.pl

Streszczenie

Publikacja omawia zaawansowane techniki diagnostyczne wykorzystywane do pomiarów jakości energii elektrycznej w izolowanych systemach elektroenergetycznych stosowanych na jednostkach pływających, małych wyspach, terenach zajętych przez strony walczące itp. W publikacji zawarto obserwacje poczynione podczas dotychczasowej współpracy oraz badań prowadzonych na trzech różnych obiektach pływających, z których jeden był tzw. *all electric ship*, gdzie podstawową część odbiorników energii elektrycznej stanowiły obwody nieliniowe, posiadające zasadniczy wpływ na parametry jakości energii elektrycznej.

Słowa kluczowe: izolowany system elektroenergetyczny, jakość energii elektrycznej, zniekształcenia sygnałów.

DIAGNOSTICS OF ISOLATED ELECTRICAL POWER SYSTEM ON EXAMPLE OF THE SHIP

Summary

Publication treats of advanced diagnostic technologies used for electrical power quality measurements in isolated electrical power systems, applied on ships, small islands, territories occupied by fighting sides, etc. Publication includes observations made during the present cooperation and tests carried out onboard three different ships, which one of them was all electric ship, where main part of power consumers were non-linear circuits having principle influence on the parameters of electrical power quality.

Keywords: isolated electrical power system, electrical power quality, waveform distortions.

1. WSTĘP

Wraz z rozwojem techniki człowiekowi przestają wystarczać proste metody diagnostyczne. Jak kiedyś zegar słoneczny stał się niewystarczający do pomiaru czasu, tak dziś oceny jakości energii elektrycznej nie zapewniają takie urządzenia jak woltomierz, częstotliwościomierz lub amperomierz. Obecnie, w dobie wysokiego nasycenia okrętu elektroniką i energoelektroniką, potrzeba bardziej skomplikowanych technik i metod pomiarowych, aby ocenić właściwie parametry jakościowe energii elektrycznej. Wartości dopuszczalne tych parametrów określają dla jednostek pływających:

- cywilnych: publikacja normatywna IEC 60092-101 *Electrical installations in ships Part 101: Definitions and general requirements*
- wojennych: porozumienie standaryzacyjne STANAG 1008 *Characteristics of Shipboard Electrical Power Systems in Warships of the North Atlantic Treaty Navies*

Tablica 1 przedstawia porównanie dopuszczalnych wartości wybranych parametrów

jakości energii zawartych w wymienionych dokumentach.

Tablica 1

Parametr	60092-101	STANAG 1008
Tolerancja napięcia (ciągła)	+6, -10%	±5% (±6%)
Asymetria napięcia	3%	2%
Składowe przejściowe napięcia	±20%	±16% (±20%)
Czas powrotu składowych przejściowych napięcia	1,5 s	2 s
THD	5%	5%
Pojedyncza harmoniczna	3%	3%
Tolerancja częstotliwości	±5%	±3%
Składowe przejściowe częstotliwości	±10%	±4% (±5,5%)
Czas powrotu składowych przejściowych częstotliwości	5 s	2 s

W kolumnie STANAG 1008 podano w nawiasach wyższe wartości kilku parametrów, dopuszczone pod pewnymi warunkami. W przypadku parametrów związanych z napięciem nie ma tak dużych różnic między oboma aktami normatywnymi, jakie występują w związku z częstotliwością.

Tablica 2 przedstawia parametry jakości energii związane z rozpięciem obciążeń. Tablicę 2 opracowano na podstawie jednakowych wymagań określonych zarówno w *Przepisach Klasyfikacji i Budowy Statków Morskich*, jak i *Przepisach Nadzoru i Budowy Okrętów RP* opracowanych przez Polski Rejestr Statków S.A. [1] [2].

Tablica 2

Parametr	Wartość
Rozpięcie obciążeń czynnych	15%
Rozpięcie obciążeń biernych	10%
Rozpięcie prądów	15%

Prace nad wyposażeniem umożliwiającym szczegółową ocenę jakości energii i wyeliminowanie ewentualnych zagrożeń prowadzi Katedra Elektroenergetyki Okrętowej Akademii Morskiej w Gdyni przy współudziale Polskiego Rejestru Statków S.A. w Gdańsku. W ramach wspólnie prowadzonych pomiarów na kilku budowanych obiektach, zarówno okrętach Marynarki Wojennej RP, jak i statkach cywilnych, stwierdzono, że chociaż znaczne odchylenia parametrów energii elektrycznej od ich wartości nominalnych nie występują zbyt często, to jednak pomiary takie mogą już teraz przyczyniać się do poprawy mniej widocznych, ale równie istotnych szczegółów. Warto jednak zauważyć, że wraz z postępującym wzrostem zastosowań nieliniowych odbiorników wielkiej mocy, jakimi są okrętowe napędy główne oraz stery strumieniowe, wzrośnie znaczenie właściwej jakości parametrów energii jako czynnika podstawowego dla bezpieczeństwa jednostki pływającej.

2. OD ZABURZEŃ ENERGII DO KATASTROF MORSKICH

2.1 Zniekształcenia przyczyną uszkodzeń w sieciach energetycznych

Zniekształcenia od składowych wyższych częstotliwości wywołują następujące typowe uszkodzenia większości elementów i podzespołów okrętowych instalacji elektrycznych:

- **źródeł energii elektrycznej:**
 - przegrzanie oraz trwałe uszkodzenia łożysk, izolacji blach i uzwojeń prądnic, spowodowane m.in. starzeniem termicznym materiałów elektroizolacyjnych,
- **odbiorników energii elektrycznej:**
 - przegrzanie stojana i wirnika silników elektrycznych o stałej prędkości obrotowej, ryzyko uszkodzenia łożysk zależne od stopnia nagrzania wirnika, dodatkowe przyrosty temperatury izolacji i jej przyspieszone starzenie termiczne. Szczególne zagrożenie występuje w przypadku silników w wykonaniu przeciwwybuchowym pracujących w strefach zagrożonych wybuchem.

- niezamierzone wyzwolenie wyłączników, zakłócenia w całym wyposażeniu pokładowych systemów elektrycznych, elektronicznych i sterowniczych, w tym komputerów nawigacyjnych, urządzeń radionawigacyjnych, oświetlenia, itp.
- **sieci przesyłowych energii elektrycznej:**
 - przegrzanie kabli oraz dodatkowe zagrożenie wystąpienia uszkodzeń z powodu zjawiska rezonansu, obniżona zdolność przewodzenia prądów znamionowych z powodu „zjawiska naskórkowości”, zmniejszającego skuteczną wartość przekroju kabli,
 - przegrzanie oraz przedwczesne starzenie izolacji blach i uzwojeń transformatorów [3].

2.2 Synergia pogłębia efekt

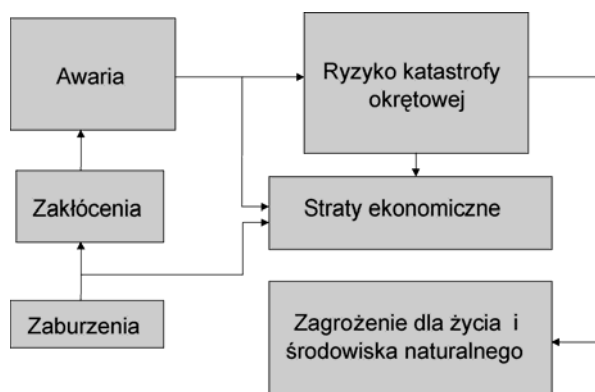
Wzajemne oddziaływanie kilku zaburzeń może spotęgować poziom zakłóceń występujących w sieci elektroenergetycznej. W tym zakresie baczna uwaga należy zwracać na szczególnie powszechne we współczesnych układach elektroenergetycznych następujące zaburzenia:

- **zniekształcenia krzywej przebiegu w czasie napięcia,**
- **odchylenia od wartości skutecznej napięcia,**
- **niesymetrię napięcia.**

Znacznie wyższy poziom zakłóceń prowadzi w prosty sposób do degradacji izolacji urządzeń zainstalowanych w systemie zasilania a w rezultacie powoduje ich przyspieszone zużycie. W wyniku zakłóceń mogą następować też uszkodzenia lub wręcz poważne awarie systemu elektroenergetycznego [4]. Badania wykonane w Katedrze Elektroenergetyki Okrętowej wykazały, że zasilanie typowych silników indukcyjnych napięciem o parametrach granicznych określonych w tablicy 1 może doprowadzić w stosunkowo krótkim czasie do uszkodzenia tych silników. Przykładowo dla zaburzeń o parametrach $f=103,7\%$, $f_n=51,85\text{ Hz}$, $U_{rms}=99,2\%$, $U_n=377\text{ V}$ i $THD=12,45\%$ (dominują harmoniczne nieparzyste niepodzielne przez 3, zwłaszcza piątą) zmierzono dodatkowy przyrost temperatury o 11 K [5], podczas gdy dla klasy izolacji F dodatkowy przyrost temperatury o 9,3 K powoduje dwukrotne przyspieszenie starzenia termicznego izolacji [9]. Warto dodać, że przyjęte poziomy zaburzeń były wcześniej obserwowane w okrętowych systemach elektroenergetycznych.

2.3 Od zaburzeń do strat

Zaburzenia często powodujące zakłócenia w wielu wypadkach prowadzą do awarii. Od awarii na morzu prowadzi już prosta droga do katastrofy okrętowej. Co oznaczają niewłaściwa jakość energii elektrycznej i związane z nią zakłócenia, przedstawia rys. 1 [5].



Rys. 1. Droga od zaburzeń do katastrofy okrętowej i strat z nią związanych

3. BADANIA JAKOŚCI ENERGII ELEKTRYCZNEJ NA JEDNOSTKACH PŁYWAJĄCYCH

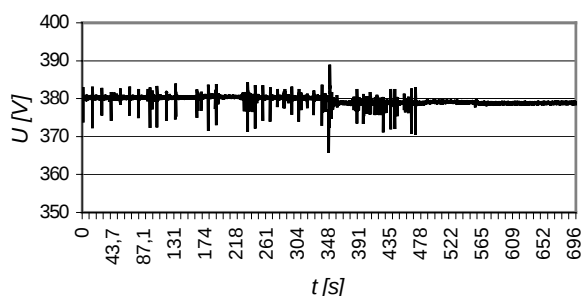
W ciągu kilkuletniej współpracy przedstawiciele Katedry Elektroenergetyki Okrętowej Akademii Morskiej w Gdyni oraz Inspektoratu Elektrycznego i Automatyki Polskiego Rejestru Statków S.A. w Gdańsku przeprowadzili pomiary jakości energii elektrycznej na jednostkach sklasyfikowanych przez PRS S.A. Niezależnie pracownicy naukowcy Akademii Morskiej w Gdyni przeprowadzili badania na wielu statkach klasyfikowanych przez inne instytucje klasyfikacyjne. W rozdziale tym przedstawiono wybrane wyniki pomiarów podstawowych parametrów jakości energii na dwóch jednostkach sklasyfikowanych przez PRS S.A. Zaprezentowano również ciekawsze wyniki pomiarów uzyskane przez naukowców z Akademii Morskiej w Gdyni na innych statkach.

3.1. ORP „Kontradmirał X. Czernicki”

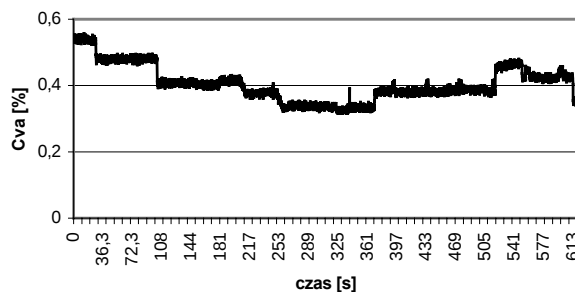
Jedną z pierwszych jednostek poddanych pomiarom jakości energii elektrycznej był ORP „Kontradmirał X. Czernicki” – okręt, na którym nie występują odbiorniki nieliniowe o większej mocy. W czasie pomiarów nie stwierdzono zaburzeń napięcia zasilającego o poziomach przekraczających dopuszczalne granice. Okresem o największej intensywności zmian wartości skutecznej napięcia były manewry wyjściowe z portu i wejściowe do portu. Przykładowe zmiany wartości skutecznej napięcia w czasie manewrów wyjściowych okrętu (początkowe 470 s) oraz okres bezpośrednio po manewrach (końcowe 230 s) przedstawiono na rys. 2.

Obserwowane zmiany napięcia mieściły się w dopuszczalnych granicach. Dodatkowo na badanym okręcie nie stwierdzono znaczącej niesymetrii napięcia, zarówno w sieci 380 V, jak i 220 V. Wskaźnik niesymetrii napięciowej C_{va} , opisujący maksymalne odchylenie napięcia

międzyfazowego od wartości średniej napięć międzyfazowych, przyjmował większe wartości w przypadku sieci 220 V. Przykładowe zmiany tego wskaźnika pokazano na rys. 3.



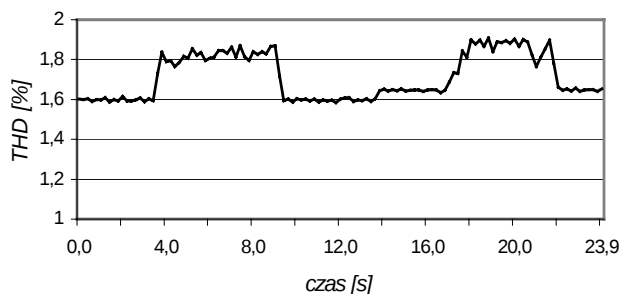
Rys. 2. Zmiany wartości skutecznej napięcia w czasie manewrów wyjściowych i wejściowych z portu



Rys. 3. Zmiany wskaźnika niesymetrii napięciowej C_{va} przy zmianach obciążenia w sieci 220V

W czasie badań wykonano pomiary wskaźnika THD oraz zawartości harmoniczných i interharmoniczných w paśmie do 3,5 kHz. Jednak nie stwierdzono znaczących wartości zniekształceń kształtu krzywej przebiegu czasowego napięć w badanym systemie 380 i 220V.

Analiza zniekształceń napięcia w sieci 220 V wykazała podobny poziom zniekształceń, zmieniający się w zakresie 1,45–1,9% w zależności od obciążenia. Przykładowe zmiany współczynnika zniekształceń THD przy zmianach obciążenia pokazano na rys. 4.

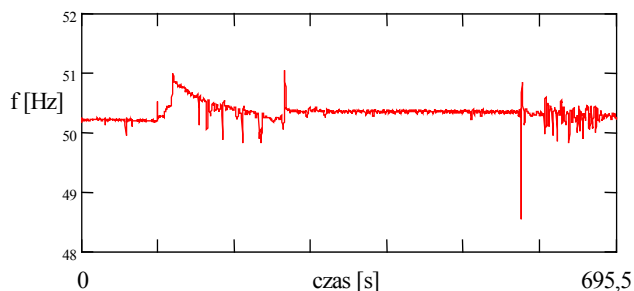


Rys. 4. Zmiany współczynnika THD w czasie zmian obciążenia w sieci 220 V

Średnia wartość częstotliwości na badanym obiekcie nieznacznie przekraczała 50,2 Hz. Zmiany częstotliwości obserwowane są jedynie w czasie

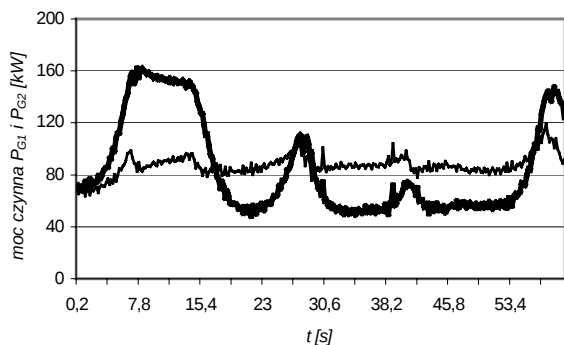
zmian obciążenia elektrowni okrętowej. Zmiana konfiguracji elektrowni posiadała pomijalnie mały wpływ na średnią wartość częstotliwości.

Największe obserwowane zmiany częstotliwości wystąpiły w czasie przejmowania obciążenia z prądnicy awaryjnej przez prądnicę wolnostojącą i odstawienia prądnicy awaryjnej (rys. 5). Wówczas zaobserwowano średnią wartość częstotliwości 50,88 Hz w czasie 5 s. Podobnie rozpląt obciążenia był w granicach określonych wymienioną normą.



Rys. 5. Zmiany częstotliwości w systemie elektroenergetycznym w czasie zmiany konfiguracji elektrowni okrętowej

Wydaje się że na badanym okręcie problem zaniżonej jakości energii nie występuje. Jednakże stwierdzono kilka niepokojących zjawisk związanych z rozplątem obciążenia pomiędzy równolegle pracujące zespoły prądowców. Zaobserwowany rozpląt mocy biernych był na poziomie zbliżonym do dopuszczalnego, tj. 10%. Natomiast największym problemem było zjawisko początkowego nierównomiernego obciążenia mocą czynną zespołów prądowców w stanach dynamicznych. Stwierdzono, że w czasie gwałtownych zmian obciążenia czas reakcji jest znacząco różny dla każdej ze współpracujących prądnic. Proces ten pokazano na rys. 6, linią pogrubioną oznaczono obciążenie czynne P_{G2} zespołu prądowcowego nr 2.

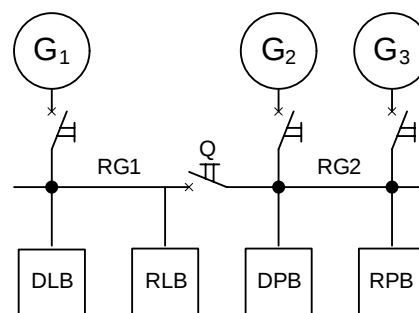


Rys. 6. Przykładowe zmiany obciążenia mocą czynną zespołów prądowców nr 1 i nr 2 w czasie manewrów wejściowych do portu

Ostatecznie zalecono sprawdzenie funkcjonowania regulatorów prędkości obrotowej silników napędowych prądnic [6].

3.2. e-m/v IMOR

Kolejną jednostką poddaną badaniom był tzw. *all electric ship* – całkowicie zelektryfikowany statek naukowo-badawczy Instytutu Morskiego w Gdańsku „IMOR”. Napęd statku stanowią dwa silniki elektryczne zasilane za pośrednictwem przekształtników energoelektronicznych CXI firmy VACON o mocy 315 kW każdy. Silniki te są umieszczone na rufie i oznaczone na rys. 7 jako *RLB* dla pędnika umieszczonego na lewej burcie oraz *RPB* dla pędnika na prawej burcie. Za pomocą pojedynczego bloku oznaczono układ falownik – pędnik azymutalny.



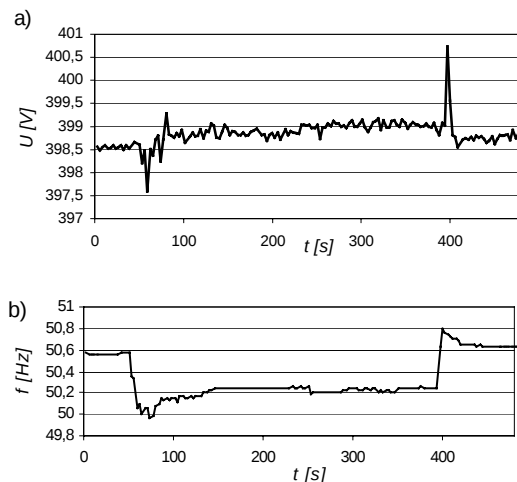
Rys. 7. Schemat blokowy układu elektroenergetycznego na jednostce IMOR

W celu zwiększenia manewrowości statku dodatkowo katamaran jest wyposażony w dwa pędniki o mocy 72 kW każdy umieszczone na dziobie – na schemacie oznaczone jako *DLB* oraz *DPB*. Podstawowym źródłem energii są trzy generatory, z czego dwa o mocy 425 kVA (G_1 i G_3) oraz dodatkowo jeden o mocy 200 kVA (G_2). Charakterystyczną cechą dla katamaranów są dwa pływaki z własnymi rozdzielnicami głównymi RG_1 oraz RG_2 , które są rozdzielone wyłącznikiem Q . Dzięki takiemu układowi zostaje zwiększona niezawodność jednostki. W przypadku załania jednego pływaka, wyłączenia z eksploatacji jednej z rozdzielnic, statek wciąż jest manewrowalny.

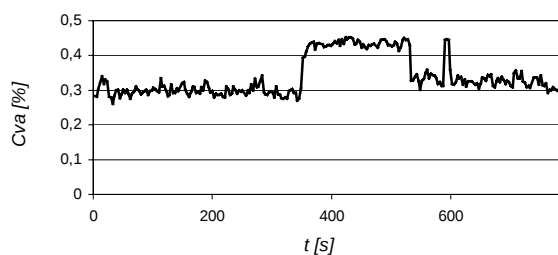
Szczególnie interesujące wyniki spodziewano się uzyskać w czasie działania napędów elektrycznych statku, wprowadzających do sieci jako obwody nieliniowe pewne zaburzenia elektromagnetyczne. W trakcie badań okazało się jednak, że generalnie jakość energii elektrycznej na badanej jednostce była poprawna. Największe zmiany napięcia i częstotliwości zaobserwowano w sieci 400 V, podczas jazdy morskiej w czasie zmiany prędkości statku, co zostało pokazane na rys. 8.

Na badanej jednostce nie stwierdzono znaczącej niesymetrii napięcia, zarówno w sieci 400 V, jak i 230 V. W przypadku sieci 230 V

wskaźnik C_{va} zmieniał się w przedziale 0,15-0,48, w zależności od obciążenia (rys. 9), natomiast w sieci 400 V ten wskaźnik przyjmował podobne wartości 0,21-0,42.



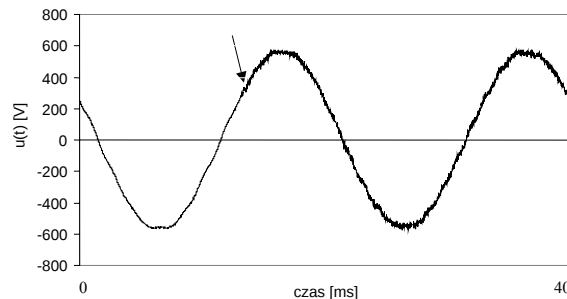
Rys. 8. Zmiany napięcia U (a) i częstotliwości f (b) w sieci 400 V w czasie zmiany prędkości statku (jazda w morzu)



Rys. 9. Zmiana wskaźnika niesymetrii napięciowej C_{va} w czasie jazdy w morzu przy zmianie obciążenia w sieci 230 V

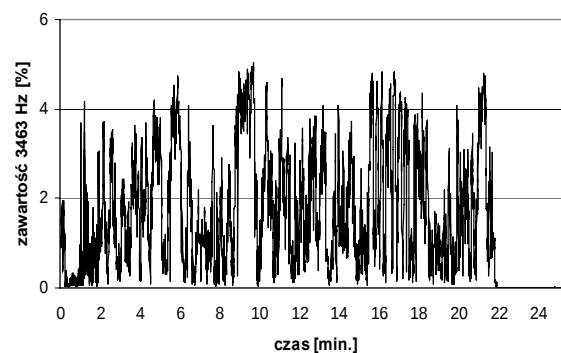
Ze względu na charakter zainstalowanych odbiorników szczególnie interesująca była ocena zniekształceń prądów i napięć na badanym statku. Moc zainstalowanych przekształtników była niewiele niższa od całkowitej mocy elektrowni okrętowej. Ocenę zniekształceń napięcia zasilającego przeprowadzono przede wszystkim w zakresie zaburzeń przewodzonych niskiej częstotliwości. Zgodnie z *IEC60533 Electrical installations in ships – Electromagnetic compatibility* za górną granicę zakresu niskich częstotliwości przyjęto 10 kHz. Pomiary przeprowadzono, w czasie przełączania napędów, dla napięć 400 V i 230 V na szynach rozdzielnic głównej oraz dla napięcia zasilającego 230 V urządzenia na mostku nawigacyjnym, ze względu na największe ryzyko wystąpienia zakłóceń (koncentracja urządzeń wrażliwych na zaburzenia). Jak wcześniej wspomniano, zasadniczym źródłem zaburzeń na badanej jednostce były napędy dużych mocy (głównie pędniki rufowe) zasilane z przekształtników energoelektronicznych. Na rys. 10 pokazano zmianę kształtu przebiegu

czasowego napięcia 400 V w chwili załączenia napędu rufowego z lewej burty. Wyraźnie widoczny jest wpływ pracy przekształtnika na kształt krzywej napięcia w rozważanym systemie. Należy jeszcze dodać, że w analizowanym przypadku napęd był nieobciążony.



Rys. 10. Przykładowy przebieg czasowy napięcia 400 V bezpośrednio przed i po załączeniu napędu rufowego (chwile załączenia oznaczono strzałką)

Pewien niepokój mogą budzić jedynie składowe w paśmie 3,2-3,7 kHz oraz niektóre pasma powyżej częstotliwości 10 kHz, przy czym zaburzenia te są niesymetryczne. Rys. 11 ukazuje zmiany składowej o częstotliwości ok. 3463 Hz w czasie jazdy morskiej i manewrów wejściowych statku do portu. Pomiary wykonano na szynach rozdzielnic na mostku nawigacyjnym dla napięcia 230 V. Maksymalna obserwowana wartość składowej 3463 Hz, zmierzona dla okna pomiarowego równego 10 okresom analizowanego napięcia, wyniosła 5,03%. Zaburzenia w paśmie 3,2-3,7 kHz były spowodowane pracą przekształtników o częstotliwości przełączania 3,6 kHz (według danych producenta).

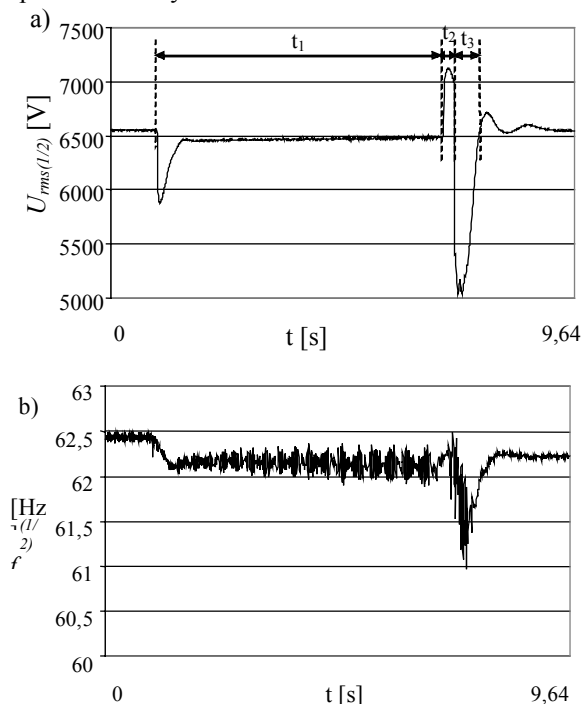


Rys. 11 Zmiany zawartości składowej o częstotliwości ok. 3463 Hz w napięciu zasilającym urządzenia na mostku nawigacyjnym

Po przeprowadzeniu badań zalecono m.in. zastosowanie filtrów RFI (dla składowych powyżej 10 kHz) oraz rozważenie zastosowania różnych częstotliwości przełączania dla napędów rufowych [7].

3.3. m/s WAWEL

Trzecia jednostka poddana badaniom to prom m/s Wawel z systemem o napięciu nominalnym 6,6 kV i częstotliwości 60 Hz. Największe zmiany wartości napięcia i częstotliwości, związane z procesem rozruchu sterów strumieniowych, zaobserwowano w czasie manewrów wejściowych do portów. Najgorszy przypadek zmian U_{rms} pokazano na rys. 12.



Rys. 12. Zmiany napięcia $U_{rms(1/2)}$ (a) i częstotliwości $f_{(1/2)}$ (b) w czasie wybranego procesu łączeniowego – rozruch silnika napędowego steru strumieniowego; $t_1 = 5970$ ms, $t_2 = 225$ ms, $t_3 = 542$ ms

Przedstawiony zapad napięcia przyjmował wartości przekraczające te określone stosownymi normami (patrz Tablica 1) i wyniósł 23%.

Ustalono iż wskazana jest analiza rozruchu, zwłaszcza należy sprawdzić dobór czasów poszczególnych faz rozruchu do charakteru procesu [8].

4. PODSUMOWANIE

Inspektorat Elektryczny i Automatyki Polskiego Rejestru Statków w ramach współpracy z Katedrą Elektroenergetyki Okrętowej Akademii Morskiej w Gdyni zamierza wprowadzić obowiązek monitorowania sieci elektroenergetycznych jednostek pływających. Monitorowanie jakości energii elektrycznej zapewni niezawodność systemu zasilania. Kontrola, rejestracja i analiza parametrów jakości energii niezbędnych do prawidłowego działania systemu elektroenergetycznego mogą zapobiegać awariom,

a ponadto powinny uprościć planowanie remontów i przyczynić się do oszczędności paliwowych.

Jednocześnie rozwój nowoczesnych technik pomiarowych i dostępność coraz bardziej zaawansowanych technologicznie narzędzi pomiarowych sprawiają, że tego rodzaju monitoring jest obecnie możliwy i relatywnie tani.

LITERATURA

- [1] PRS S.A., *Przepisy Klasyfikacji i Budowy Statków Morskich Cz. VIII*; PRS, Gdańsk 2002.
- [2] PRS S.A., *Przepisy Nadzoru i Budowy Okrętów RP Cz. VIII*; MON Warszawa 2006.
- [3] E. Szmit, J. Mindykowski, T. Tarasiuk, *Jakość energii elektrycznej na statkach wspólnym problemem armatorów, stoczni, uczelni morskich i towarzystw klasyfikacyjnych*, Jakość energii elektrycznej, SEP, Gdańsk 2004, s. 23-30/
- [4] J. Mindykowski, E. Szmit, T. Tarasiuk, *Electric Power Quality and Ship's Safety*; Polish Academy of Sciences, Branch in Gdańsk Marine Technology Transactions, Vol. 15, Gdańsk 2004, s. 351-360.
- [5] P. Gnaciński, J. Mindykowski, T. Tarasiuk, *Thermal Phenomena in Electric Machines in the Wake of Electric Power Quality in Ships' Networks*. 8th International Conference „Electrical Power Quality and Utilization”, Cracow September 21-23, 2005, s. 399-404.
- [6] J. Mindykowski, T. Tarasiuk, M. Szweđa, E. Szmit, *Badania jakości energii elektrycznej na wybranym statku*, PRS Raport Techniczny Nr 48, Gdańsk 2004.
- [7] J. Mindykowski, T. Tarasiuk, M. Szweđa: *Raport z badań jakości energii elektrycznej na statku „IMOR”*; Akademia Morska w Gdyni, Gdynia 2005.
- [8] J. Mindykowski, T. Tarasiuk, M. Szweđa: *Analiza zmienności napięcia i częstotliwości na statku m/s WAWEL*; Akademia Morska w Gdyni, Gdynia 2005.
- [9] E. L. Brancato, *Estimation of lifetime expectationcies of motor*, IEEE Electrical Insulation Magazine, May/June 1992, Vol 8 no 3, s. 5-13.

Janusz MINDYKOWSKI – prof. dr hab. inż. w dyscyplinie elektrotechnika, obecnie dziekan Wydziału Elektrycznego Akademii Morskiej w Gdyni i kierownik Katedry Elektroenergetyki Okrętowej

Tomasz TARASIUK – dr inż. w dyscyplinie elektrotechnika, adiunkt w Katedrze Elektroenergetyki Okrętowej

Edward SZMIT – inż. elektryk, obecnie kierownik Inspektoratu Elektrycznego i Automatyki PRS, w przemyśle okrętowym od 1965 roku.

Daniel CZARKOWSKI – mgr inż., inspektor Polskiego Rejestru Statków S.A.

ANALIZA SYGNAŁU ŚWIETLNEGO ODBITEGO OD POWIERZCHNI W DIAGNOSTYCE OBIEKTÓW TECHNICZNYCH

Mariusz BOGDAN*, Józef BŁACHNIO**

*Politechnika Białostocka, Wydział Mechaniczny, Katedra Automatyki i Robotyki
ul. Wiejska 45c, 15-351 Białystok, marbog@pb.edu.pl

**Politechnika Białostocka, Wydział Mechaniczny, Katedra Automatyki i Robotyki
ul. Wiejska 45c, 15-351 Białystok

Streszczenie

W artykule przedstawiono warunki badań dotyczące oddziaływania obciążeń cieplnych o wysokiej temperaturze na próbki ze stali 2H13. Starano się powiązać temperaturę wygrzewania (zmianę barwy badanych powierzchni) ze zmianami mikrostruktury warstwy wierzchniej badanych próbek. Zatem z punktu widzenia diagnostyki obiektów technicznych, w sposób bezinwazyjny starano się uzyskać ważną informację diagnostyczną, która w praktyce może posłużyć do oceny zmian mikrostruktury lub też stanu przegrzania materiału w eksploatacji elementów maszyn.

Słowa kluczowe: łopátka turbiny gazowej, rozkład widmowy RGB, stali 2H13, akwizycja obrazów, przegrzanie struktury.

ANALYSIS OF LIGHT SIGNAL REFLECTED OF A SURFACE IN TECHNICAL SYSTEM DIAGNOSTICS

Summary

In this article, main emphasis was put on determining the influence of high temperature on steel 2H13 samples. Attempted was linking the high temperature of heating with change of tested surface color and changes in microstructure of sample's surface layer. In terms of technical system diagnostics, in a non-invasive manner obtaining important diagnostic information was attempted, which in practice can be used for evaluation of microstructure changes as well as the state of overheating in machine elements operation.

Keywords: gas turbine scoop, RGB spectrum distribution, steel 2H13, canvassing and image analysis, structure overheating.

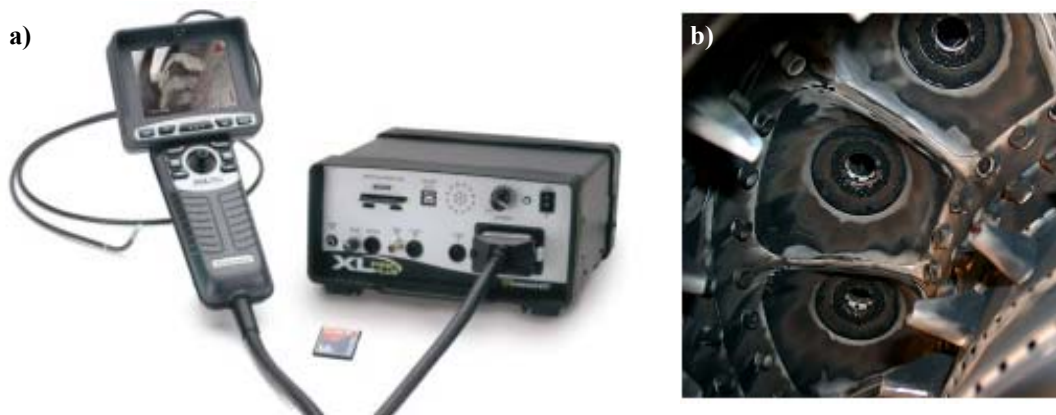
1. ISTOTA PODJĘTEGO PROBLEMU DIAGNOSTYCZNEGO

Badania diagnostyczne, w dużym uogólnieniu, mają na celu określenie (weryfikację) stanu technicznego badanego obiektu. W przypadku złożonych systemów (struktur) diagnostyka odgrywa znaczącą rolę w ocenie rzeczywistego czasu bezawaryjnej pracy. Bardzo ważnym czynnikiem wpływającym na parametry eksploatacyjne są zmiennie-skrajne warunki pracy poszczególnych elementów i podzespołów obiektów technicznych. To one w głównej mierze przyczyniają się do zużycia poszczególnych elementów i podzespołów obiektów technicznych. Zwiększenie niezawodności, określane mianem trwałości eksploatacyjnej lub żywotności, zależy od wielu czynników ale podstawowym jest kryterium materiałowe [1][2].

W procesie eksploatacji silników turbinowych tak lotniczych, trakcyjnych jak i morskich występują

różnego rodzaju uszkodzenia ich zespołów turbinowych. Najczęstszymi przypadkami uszkodzeń są przegrzania materiału, a także zmęczenie cieplne łopatek aparatu dyszowego jak i wirnika. Uszkodzenia tego typu prowadzą do wadliwej pracy silnika, a niekiedy jak w przypadku silników lotniczych do tragicznego w skutkach wypadku. Usunięcie tego typu uszkodzeń zawsze realizowane jest przez naprawę główną silnika, co skutkuje ogromnymi kosztami, w lotnictwie rzędu 5-10 mln zł w zależności od typu silnika. Zatem koszty naprawy głównej silnika w porównaniu z wartością np. łopatki wirnika turbiny są od 500-1000 razy wyższe.

Decyzję o konieczności naprawy silnika podejmuje diagnosta, który posługując się metodą wizualną mając do dyspozycji np. wideoskop (rys. 1) może diagnozować stan elementów trudnodostępnych silnika w tym również turbiny.



Rys. 1. a) wideoskop przemysłowy [3], b) komora spalania silnika odrzutowego na ekranie wideoskopu [4]

Ocenę stanu dokonuje się na podstawie zarejestrowanego obrazu powierzchni diagnozowanego elementu i porównanie tego obrazu z wzorcowymi obrazami powierzchni zdalnych i niezdatnych analogicznych elementów np. łopatek turbiny tabela 1.

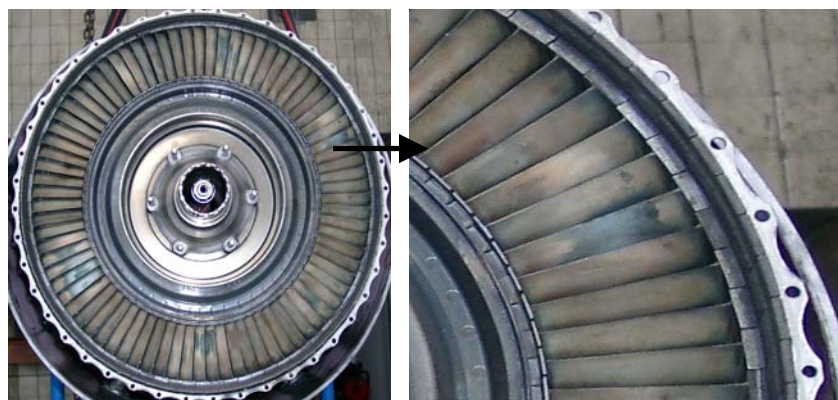
Takie kryteria oceny stanu są bardzo subiektywne gdyż zależą od wiedzy i stanu wzroku diagnosty. Weryfikację decyzji diagnosty realizuje się metodą niszczącą. Badany element poddawany jest analizie mikrostruktury na zgładzie metalograficznym. Pomyłki subiektywnej oceny diagnosty prowadzić mogą do uznania przegrzanej łopatki za zdatną lub nieprzegrzanej za niezdatną.

Wynikiem tego w pierwszym przypadku dochodzi w krótkim czasie pracy silnika do wypadku lotniczego, w drugim przypadku zaś do niepotrzebnych kosztów naprawy głównej silnika.

W dalszym ciągu brakuje obiektywnych kryteriów umożliwiających jednoznaczną ocenę stopnia przegrzania materiału łopatek. W przypadku przedstawionym na rys. 2 nie można jednoznacznie ocenić czy powierzchnia chociażby jednej z przedstawionych łopatek wykazuje przegrzanie materiału, a już w ogóle według dotychczasowych kryteriów nie można się wypowiedzieć o stopniu

Tabela 1. Kolory warstwy tlenków i odpowiadające im temperatury na przelomie łopatki w przypadku chłodzenia na powietrzu [5]

Temperatura w °C	Kolor warstwy tlenków na przelomie łopatki
400	jasno szary
500	jasno szary ze słabym żółtym kolorem
600	jasno żółty
650	żółty, ciemno żółty
700	żółto brązowy
750	żółto brązowy z odcieniem fioletowym
800	ciemno fioletowy
850	błękitny, granatowy



Rys. 2. Silnik turbinowy z uwidocznionymi zmianami barwy na powierzchniach łopatek [4]

przegrzania. Jak dotąd nie opracowano nieniszczącej metody badań stopnia przegrzania materiału łopatek opartej na obiektywnych kryteriach [5].

W związku z powyższym istnieje konieczność opracowania nieniszczącej metody obiektywnej oceny stopnia przegrzania materiału łopatek turbiny, zapewniającej odpowiednią wiarygodność rozumianą jako odpowiednią czułość wykrywania zmian struktury materiału oraz niezawodność zapewniającą minimalny błąd diagnozowania.

2. METODY ANALIZY SYGNAŁÓW ŚWIETLNYCH ODBITYCH OD POWIERZCHNI OBIEKTU METALOWEGO

Ludzkie oko postrzega mały wycinek widma elektromagnetycznego, jednak mimo to jest w stanie rozróżnić całą tęczę barw (kolory w zakresie od 400 do 700 nanometrów, od fioletu do głębokiej czerwieni). Światło padające zostaje reemitowane z powierzchni dzięki czemu możemy obserwować i rozróżniać kształty i barwy badanych powierzchni obiektów metalowych [6][7]. Współczesny rozwój techniki akwizycji obrazów zarówno monochromatycznych jak i kolorowych umożliwia coraz to szersze i prostsze wykorzystywanie informacji zawartych w rejestrowanych obrazach w szeroko pojętej diagnostyce. Istotnym zagadnieniem z punktu widzenia diagnostyki w wielu przypadkach jest prawidłowa ocena barwy [8][9][10]. Jest ona często wykorzystywana w wielu procedurach diagnostycznych w takich obszarach jak przemysł spożywczy, włókienniczy, chemiczny, czy też lotnictwo. Decyzja podejmowana jest na podstawie porównania barwy wzorca z zarejestrowanymi obrazami. W pracy [8] zaproponowano budowę systemu doradczego wspomagającego zobiektywizowanie oceny barwy, jako metody służącej do określenia produktów naftowych na korozję metali. Zaproponowany system wykorzystuje metody logiki rozmytej w celu oceny i klasyfikacji barw [9][10]. Autor przedstawia także metody wyliczenia funkcji przynależności analizowanego punktu obrazu barwnego do określenia barwy (obszaru) z przyjętą klasyfikacją barw opartą na trójkącie chromatyczności CIE.

Na potrzeby współczesnej diagnostyki w coraz to w większym stopniu wykorzystuje się scalone analizatory obrazu (matryce obrazowe) do różnych technik rejestracji i analizy obrazów w aspekcie pozyskiwania informacji ilościowej i jakościowej o badanych stanach i zjawiskach odwzorowanych za pomocą obrazów. Do tych metod wspomagających należy w szczególności zaliczyć morfologiczne metody przekształcania obrazu. Są one jednymi z najważniejszych operacji w komputerowej analizie obrazu ponieważ są punktem wyjściowym do tworzenia bardziej złożonych operacji, związanych z analizą kształtu obiektów i ich wzajemnego rozmieszczenia [12].

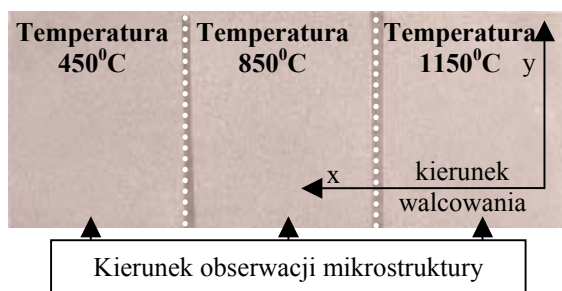
Charakterystyczną grupę stanowią obrazy o jednolitej strukturze nazywane potocznie teksturami. Mają one duże znaczenie w wielu dziedzinach nauki i techniki jak choćby: metalografia, krystalografia i tribologia (obrazy powierzchni trących i cząstek zużyciowych) [13]. Ekstrakcja ich cech (np. parametrów statystycznych, parametrów Haralick'a) pozwala klasyfikować je i na ich podstawie wnioskować o cechach materiałów, obiektów i procesów odwzorowanych obrazami tekturowymi [14].

Interesującą nieniszczącą metodą oceny (diagnozowania) stanu niewirujących łopatek wirnikowych jest metoda wykorzystująca technikę laserową. Promieniowanie laserowe charakteryzuje się wysoką spójnością, monochromatycznością i kierunkowością rozchodzenia się wiązki („źródło światła” - wielka koncentracja energii w wiązce o małej rozbieżności). W badaniach wykorzystuje się przedstawione powyżej własności jako parametry opisujące źródło promieniowania, w dużym uproszczeniu bada się różnice pomiędzy promieniowaniem padającym (energiją źródła) a promieniowaniem odbitym (strata energii wywołana pewnym rozproszeniem wiązki na badanej powierzchni). Przy czym należy wspomnieć, że różne typy laserów emitują promieniowanie o określonej spójności (mniejszej lub większej) ponad to, o zróżnicowanych parametrach wiązek (długości fal i energii). Badania diagnostyczne z zastosowaniem techniki laserowej pozwalają na zobrazowanie pól przemieszczeń i deformacji przy różnych wymuszeniach statycznych, jak i dynamicznych. Pomiar deformacji łopatek przy wymuszeniu statycznym pozwalają oszacować wady łopatek w postaci pęknięć łopatek na krawędziach (kilkumilimetrowych). Natomiast badania przy wymuszeniu dynamicznym pozwalają uzyskać interferogramy dla określenia postaci drgań łopatek przy różnych częstotliwościach rezonansowych, co stanowi źródło informacji o dynamice łopatki, a zatem o jej własnościach mechanicznych, stanie konstrukcyjnym itp. [15].

3. OPIS EKSPERYMENTU

3.1. Opis próbek wykorzystywanych w doświadczeniu

Eksperyment badawczy przeprowadzono z wykorzystaniem trzech próbek wykonanych ze stali 2H13 (rys. 3). Próbkę wycięto wzdłuż wyrobu hutniczego i powierzchnię szlifowano w kierunku wzdłużnym z chłodzeniem w celu uniknięcia przypałów szlifierskich. Dodatkowo w celu zwiększenia parametrów eksploatacyjnych badanego materiału jego warstwa wierzchnia została poddana technologii kulowania.



Rys. 3. Schemat próbek ze stali 2H13; x - kierunek walcowania, y - kierunek obserwacji mikrostruktury

3.2 Metodyka akwizycji obrazów – rejestracji luminancji (jasności) i chrominancji (barwy) powierzchni badanych próbek

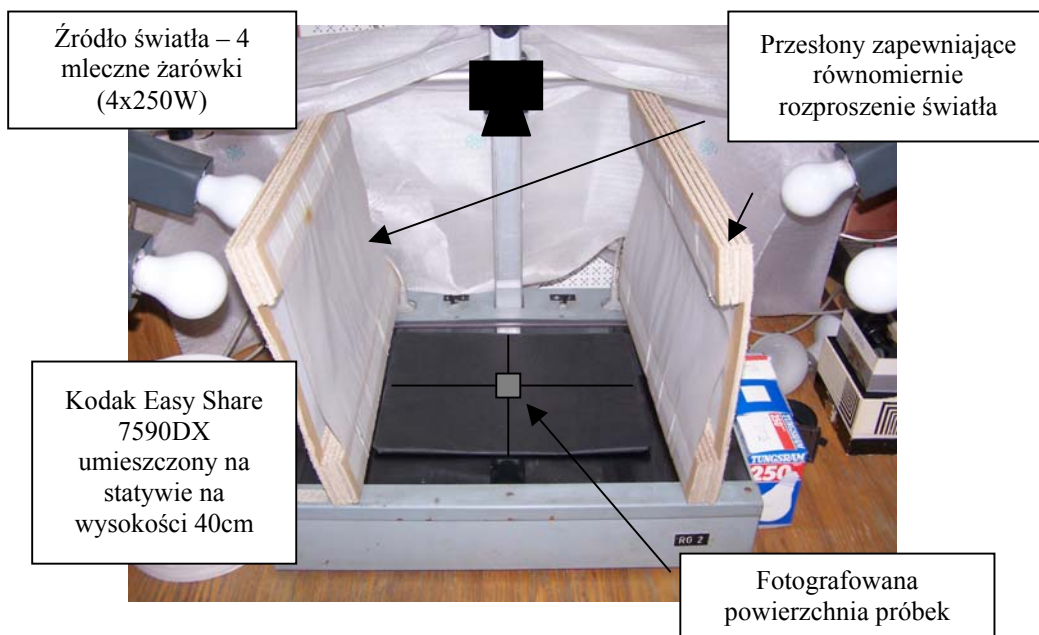
Barwa i natężenie to cechy światła – promieniowania widzialnego (część widma elektromagnetycznego), które są rejestrowane przez nasz zmysł wzroku. W przypadku aparatu cyfrowego rolę „narządu wzroku” pełni wbudowana optyka skupiająca promienie świetlne oraz elektronika będąca światłoczułym sensorem. Obraz znajdujący się przed obiektywem aparatu zostaje odwzorowany na powierzchni matrycy zaopatrzonej w detektory wyposażone w trójkolorowe filtry RGB (trzy główne składowe barwy, kolor czerwony – R, zielony – G i niebieski – B) – umożliwia to detekcję określonych barw. W dużym uproszczeniu, pozyskana informacja o intensywności rozkładu barw w przypadku obrazów kolorowych lub odcieni szarości w przypadku obrazów monochromatycznych jest zapisywana na nośnik pamięci w postaci punktów zwanych pikselami które tworzą obraz [7].

Współczesne metody analizy obrazów znalazły szerokie zastosowanie w szeroko pojętej diagnostyce technicznej [8][9][10][13][15]. Do podstawowych zalet tego typu diagnozowania należy wymienić bezinwazyjną metodę pozyskiwania informacji o stanie technicznym danego obiektu.

Obraz w naszym przypadku opisuje stan danej powierzchni – warstwy wierzchniej. Postrzeganej jako luminancję (jasność) i chrominancję (barwę) docierającą do urządzenia rejestrującego – aparatu fotograficznego i sposobie zapisu tego zjawiska jako reprezentacji danych w postaci „cyfrowego obrazu”.

Zdjęć powierzchni próbek dokonano przed i po wygrzaniu ich w piecu. Rejestrację realizowano na stanowisku badawczym – rys. 4 za pomocą aparatu cyfrowego Kodak Easy Share 7590DX. Powtarzalność otrzymywanych wyników zapewniono poprzez fotografię próbek w tych samych warunkach przy manualnym ustawieniu wartości parametrów cyfrowego aparatu fotograficznego. Zapewniono, także równomierne odbicie światła od powierzchni metalu poprzez zastosowanie przesłon dzięki czemu fotografowana powierzchnia próbek pozbawiona została odbić światła powodującego prześwieślenie uzyskanych obrazów. Identyczne warunki fotografowania pozwalają na porównanie zmian wywołanych oddziaływaniem określonej temperatury na barwę powierzchni.

Formatem zapisu zdjęć był format JPEG, dostosowany do potrzeb kompresji obrazu bazujący na standardach Exif 2.2. W celu zachowania jak największej ilości informacji o rejestrowanym obrazie kompresję ustawiono na poziomie najniższym z możliwych.

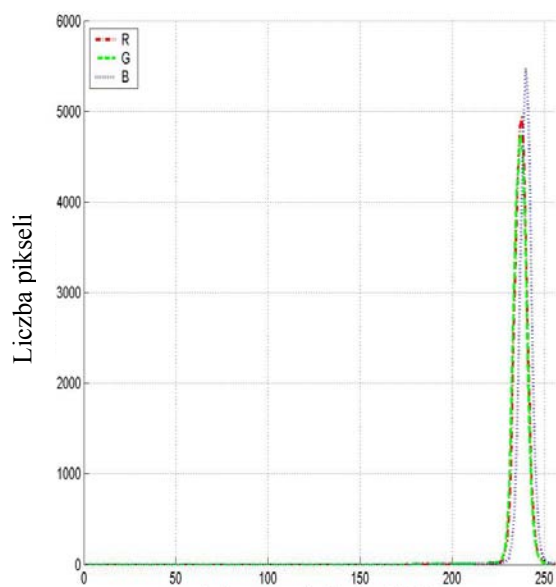


Rys. 4. Stanowisko do akwizycji obrazów powierzchni obiektów badanych (próbek)

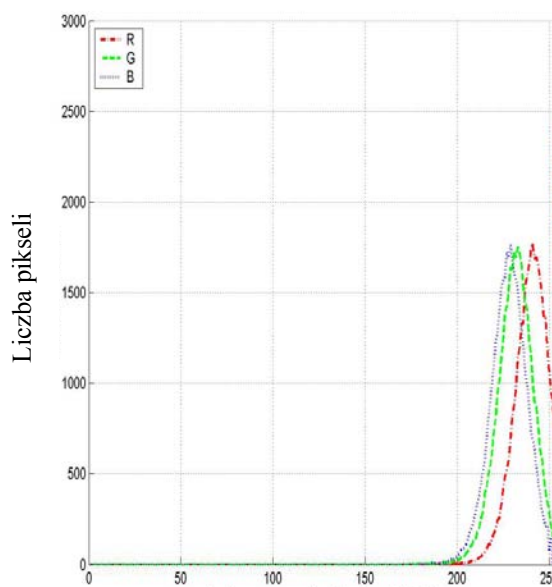
4. WYNIKI ANALIZY OBRAZÓW POWIERZCHNI PRÓBEK

W wyniku wygrzewania próbek w piecu zaobserwowano zmianę barw ich powierzchni. Z czterech fotografii powierzchni próbek ze stali 2H13 wykonanych przed i po wygrzaniu wyodrębniono obszary zainteresowania (Region of Interest - ROI) z 24 bitową głębokością koloru o wielkości 200x200 pikseli. W celu ustalenia

parametrów umożliwiających opisanie stopnia przegrzania badanych powierzchni, posłużono się analizą obrazu w zakresie rozkładu barw podstawowych tzn. czerwonej (Red), zielonej (Green) i niebieskiej (Blue) - RGB (rys. 5, 7, 9) oraz od stopnia szarości (rys. 6, 8) poszczególnych powierzchni (parametry statystyczne pierwszego rzędu - składowe RGB, odcienie szarości - opis parametryczny histogramów).



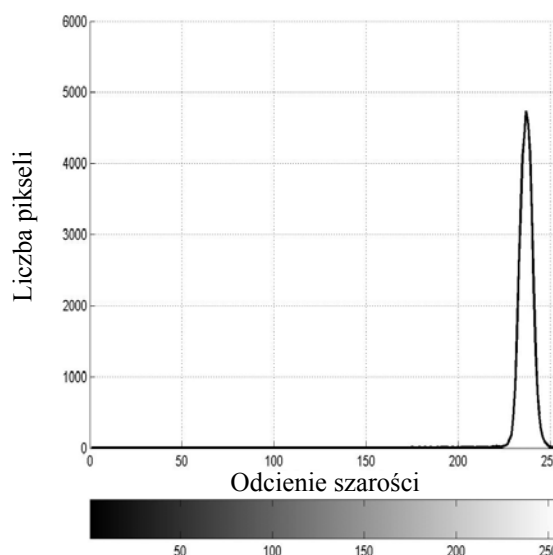
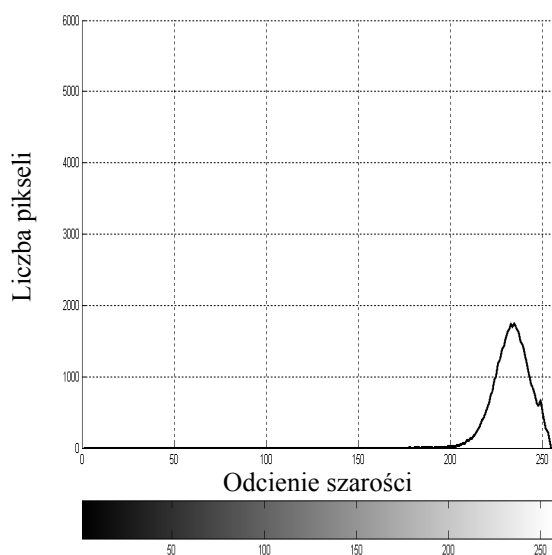
Poziom nasycenia składowych RGB



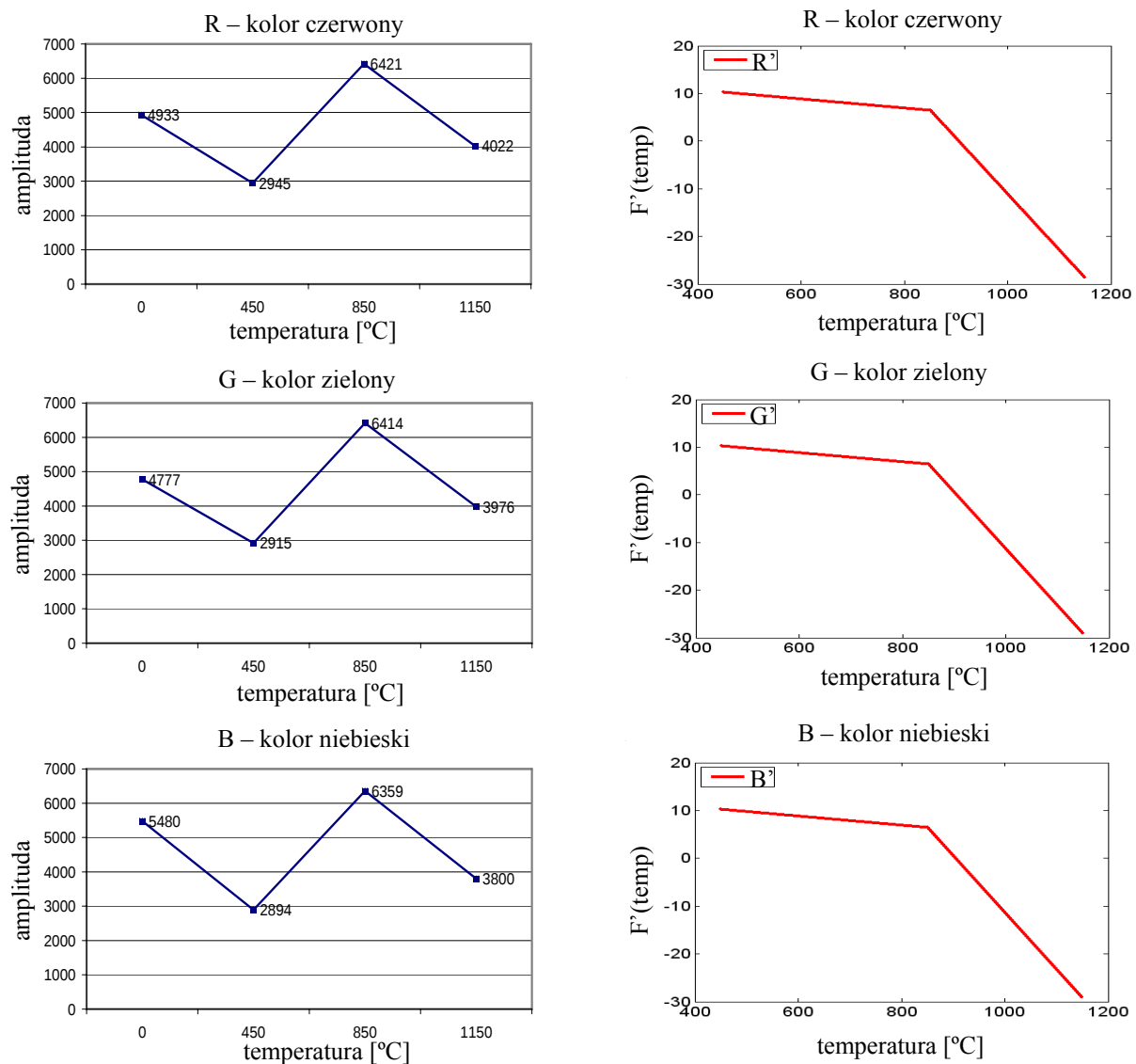
Poziom nasycenia składowych RGB

Rys. 5. Poziom nasycenia składowych RGB obrazu powierzchni próbki ze stali 2H13

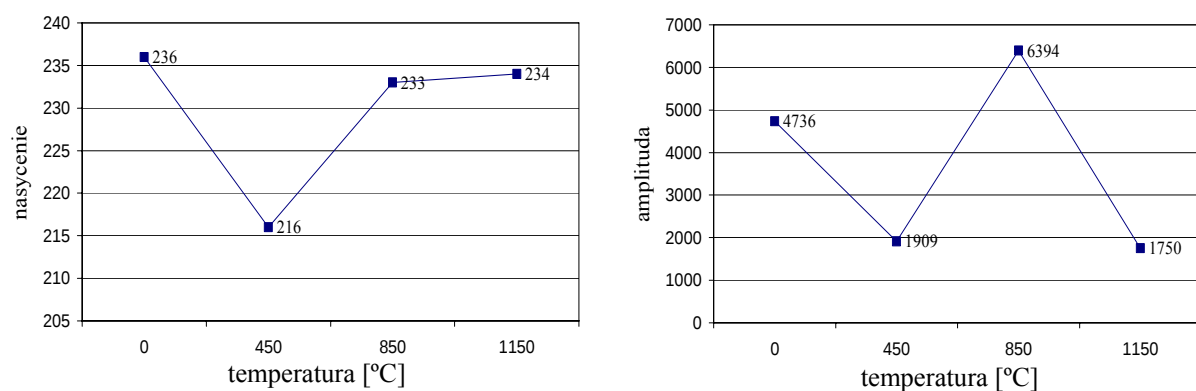
a) przed piecem; b) po wygrzewaniu w temperaturze 1150°C



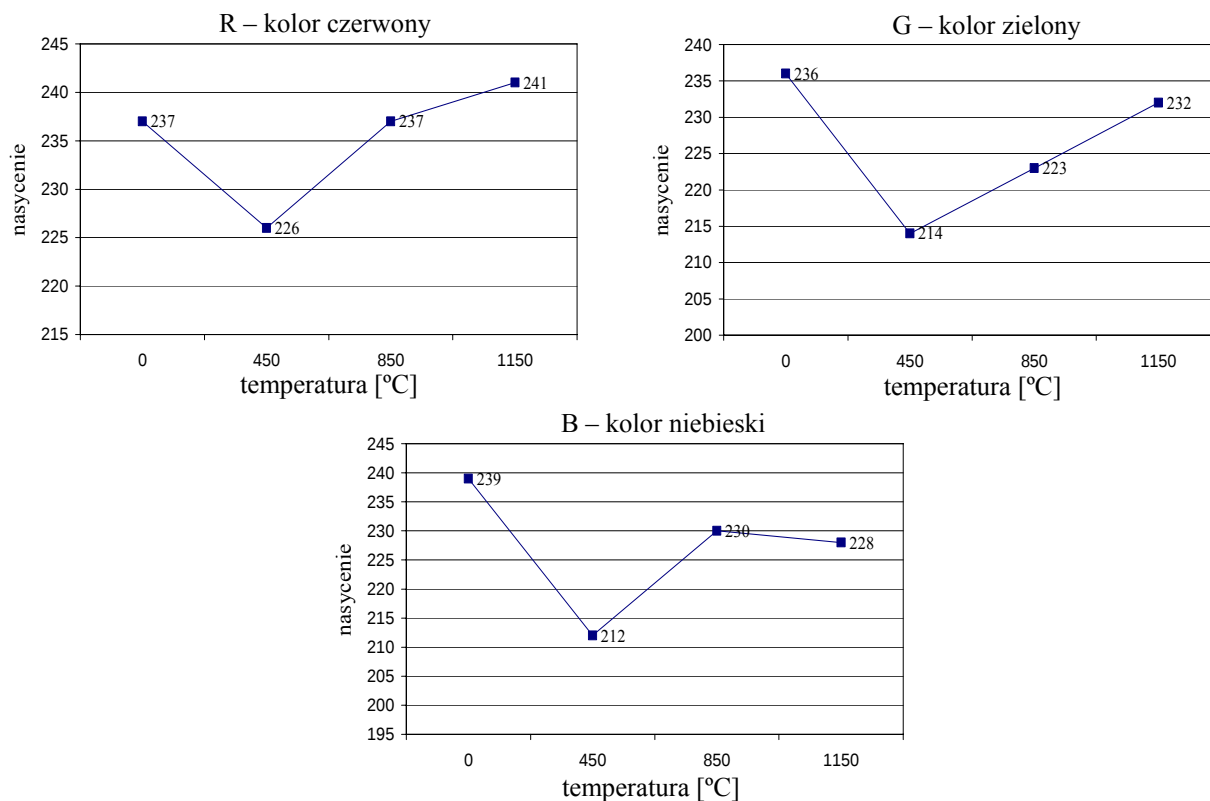
Rys. 6. Odcienie szarości obrazu monochromatycznego obrazu powierzchni próbki ze stali 2H13 a) przed piecem; b) po wygrzewaniu w temp. 1150°C



Rys. 7. Zmiany wartości maksymalnej amplitudy (ilość pikseli) nasycenia obrazu składowymi RGB dla różnych temperatur wygrzewania próbek oraz pochodna liczona w punktach temperaturowych



Rys. 8. Odcienie szarości w obrazie monochromatycznym w uzależnieniu od temperatury wygrzewania: a) wartość maksymalnej amplitudy, b) wartość nasycenia (odcienie szarości)

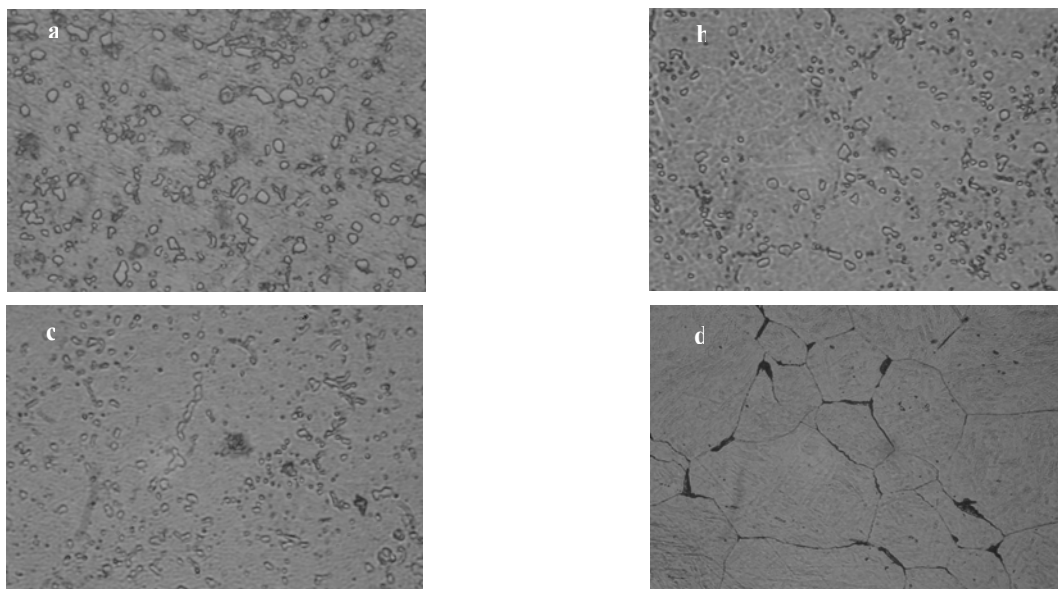


Rys. 9. Zmiany wartości położenie środka maksymalnej amplitudy nasycenia obrazu składowymi RGB dla różnych temperatur wygrzewania próbek

5. STRUKTURA METALOGRAFICZNA PRÓBEK ZE STALI 2H13 PRZED I PO ODDZIAŁYWANIU CIEPŁA O WYSOKIEJ TEMPERATURZE

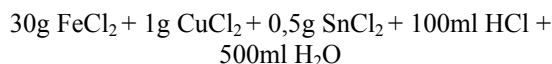
Własności wytrzymałościowe i technologiczne stali związane są z jej mikrostrukturą zależną w dużym stopniu od rodzaju obróbki cieplnej [17][18]. Pod wpływem wielokrotnych zmian

temperatury ulega ona zmęczeniu cieplnemu [2]. W czasie nagrzewania następuje całkowita przemiana ferrytu w austenit dzięki czemu możliwe jest hartowanie i powstawanie struktury martenzytycznej. Stal ta hartuje się nawet w czasie chłodzenia na powietrzu [19]. Na rysunku 10 pokazano mikrostrukturę badanych próbek ze stali 2H13 przed i po wygrzewaniu w różnej temperaturze.



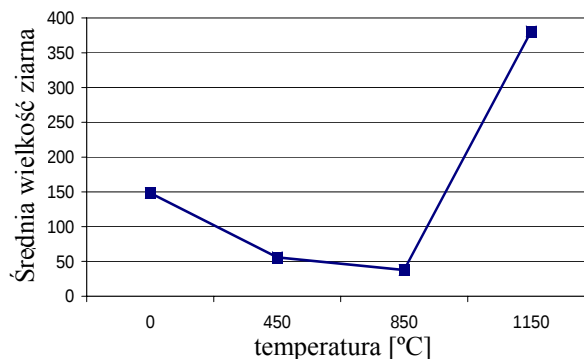
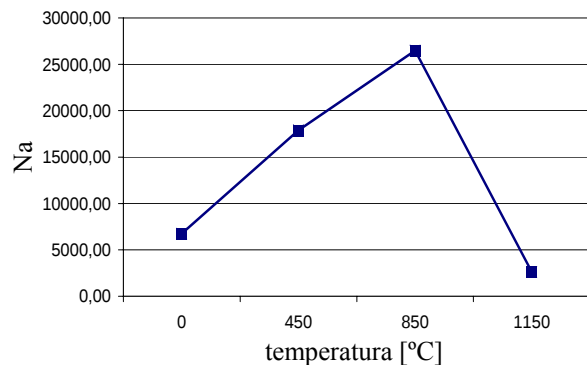
Rys. 10. Struktura metalograficzna, próbka ze stali 2H13 poddana działaniu ciepła: a) temp. otoczenia 21 °C, b) temp. 450 °C, c) temp. 850 °C,; powiększenie x 800, d) temp. 1150 °C, powiększenie x 400

Zgłady metalograficzne próbek przygotowano metodami standartowymi i następnie trawiono odczynnikami o składzie chemicznym:



Mikrostrukturę obserwowano na mikroskopie optycznym w powiększeniu 800x (rys. 10) w kierunku równoległym do walcowania hutniczego (rys. 1). W wyniku wygrzewania uzyskano zmianę wielkości ziarna (rys. 11). Początkowo średnia

wielkość ziarna wynosiła $148 \mu\text{m}^2$, po wygrzewaniu osiągnięto następujące średnie wielkości ziarna: po 450°C – $55,91 \mu\text{m}^2$, po 850°C – $37,76 \mu\text{m}^2$, po 1150°C – $380,74 \mu\text{m}^2$. Jak wynika z badań mikrostruktury wielkość ziarna zmienia się pod wpływem działania temperatury, co przedstawiona wykreślić na rysunku 11. Średnią wielkość ziarna wyznaczono na podstawie obliczeń największych powierzchni ziaren uwidocznionych na fotografiach rys. 11. Zmiany strukturalne najlepiej uwidacznia rysunek 10a, przedstawiający zmianę wielkości ziarna – wyraźne granice ziaren.



Rys. 11. Wpływ działania temperatury na mikrostrukturę próbek ze stali 2H13

- a) Zmiana ilości ziaren (Na) przypadająca na elementarną jednostkę powierzchni - $\frac{\text{Szt.}}{\text{mm}^2}$
 b) Średnia wielkość ziarna (μm^2)

6. PODSUMOWANIE

Na podstawie przeprowadzonych badań i obserwacji w fazie akwizycji obrazów - „danych pomiarowych” zauważono, że szczególną uwagę należy zwrócić na wartość natężenia oświetlenia, a także na jego rodzaj i charakter. Jest to parametr określający warunki obserwacji, a tym samym ma duży wpływ na jakość uzyskiwanych wyników pomiaru. Decyduje w głównej mierze o jakości otrzymywanych zdjęć - z punktu widzenia przeprowadzonych badań o rzeczywistym rozkładzie barw (światło białe rzetelniej oddaje barwę fotografowanego obiektu).

W poniższym artykule skupiono się głównie na określeniu wpływu oddziaływania wysokiej temperatury na materiał. W tym celu przeprowadzono badania w warunkach laboratoryjnych z wykorzystaniem próbek ze stali 2H13. Zaobserwowana zmiana barwy powierzchni próbek była pretekstem do analizy barwnej sfotografowanych powierzchni próbek. Posłużono się analizą statystyczną pierwszego rzędu (składowe RGB, odcienie szarości - opis parametryczny histogramów). Wykazano korelację pomiędzy temperaturą wygrzewania (zmiana barwy badanych powierzchni) a zmianami mikrostruktury warstwy wierzchniej badanych próbek. Zatem z punktu widzenia diagnostyki obiektów technicznych

w sposób bezinwazyjny starano się uzyskać ważną informację eksploatacyjną, która w praktyce może posłużyć do oceny zmian mikrostruktury, stanu przegrzania, a także zmęczenia cieplnego elementów i podzespołów obiektów technicznych poddanych zmiennym obciążeniom cieplnym o wysokiej temperaturze.

LITERATURA

- [1] Lewińska - Romicka A.: *Badania nieniszczące: badania defektoskopowe*, WNT, Warszawa 2001.
- [2] Chodorowski J., Ciszewski A., Radomski T.: *Materiałoznawstwo lotnicze*, Warszawa Oficyna Wydawnicza Politechniki Warszawskiej, 1996.
- [3] http://www.wme.pwr.wroc.pl/pl_badnieni.htm.
- [4] Błachnio J.: *Sprawozdanie z pracy - Nr 15/34/2002, ITWL*, Warszawa, 2002.
- [5] Andruszkiewicz S., Borgoń J.: *Metody określenia zakresu pracy silników typu 37F stosowane podczas badania wypadków lotniczych*, Instytut Techniczny Wojsk Lotniczych, Warszawa, 1982.
- [6] Kolek Z.: *Właściwości optyczne a wrażenia barwne*, Materiały VII Krajowego Sympozjum Kolorymetrycznego, Wyd. Politechnika Białostocka, Kielce-Amelinówka, 2005.
- [7] Long B.: *Fotografia cyfrowa*, Wydawnictwo Helion, Gliwice, 2002.

- [8] Wojutyński J., Wiaderek K.: *Komputerowe diagnozowanie stopnia skorodowania na podstawie oceny barwy*, Materiały IV Krajowej Konferencji Diagnostyka Techniczna Urządzeń i Systemów Diag '98, Gdańsk, 1998.
- [9] Wiaderek K., Wojutyński J.: *Sposoby fuzyfikacji określeń barw na diagramie C.I.E.* Problemy Eksploatacji 1'99, s. 117-128.
- [10] Wiaderek K.: *Wykorzystanie trójkąta chromatyczności CIE do ograniczenia liczby barw w cyfrowym obrazie barwnym.* Problemy Eksploatacji, Nr 2, ITE, Radom, 2002, s. 213-225.
- [11] Wiaderek K.: *Badanie funkcji przynależności w analizie obszarów barwnych.* Problemy Eksploatacji, Nr 1, ITE, Radom, 2000, s. 207-219.
- [12] Tadeusiewicz, R., *Komputerowa analiza i przetwarzanie obrazów.* Wydawnictwo Fundacji Postępu Telekomunikacji, Kraków, 1997.
- [13] Rudnicki Z., Oleksiak Z.: *Próba oceny zmiany filmu nanoszonego podczas tarcia ślizgowego kompozytów PFe przy pomocy analizy sekwencji obrazów*, Materiały z X Seminarium „Tworzyw Sztucznych w Budowie Maszyn”, Kraków 2003.
- [14] Rudnicki Z.: *Wybrane metody przetwarzania i analizy obrazów teksturowych*, Sprawozdanie z grantu KBN nr T07C 03318.
- [15] Lewitowicz J., Borgoń J., Ząbkowicz W.: *Problemy badań i eksploatacji techniki lotniczej*, tom 2, Wyd. ITWL, Warszawa, 1993.
- [16] www.kodak.com – oficjalna strona producenta.
- [17] Rudnik S.: *Metaloznawstwo*, Wydaw. Naukowe PWN, Warszawa 1996.
- [18] Przybyłowicz K.: *Metaloznawstwo*. Warszawa, Wydawnictwa Naukowo-Techniczne, 2003.
- [19] Przybyłowicz K., Przybyłowicz J.: *Obróbki cieplne i powierzchniowe*, Kielce, Wydawnictwo Politechniki Świętokrzyskiej, 2004.



Mgr inż. Mariusz BOGDAN Asystent na Wydziale Mechanicznym Politechniki Białostockiej. Reprezentowane dziedziny, dyscypliny i specjalności: analiza i przetwarzanie obrazów, diagnostyka techniczna. W przygotowaniu do otwarcia przewodu doktorskiego z zakresu budowy i eksploatacji, temat rozprawy doktorskiej: „Diagnozowanie stanu łopatek turbiny gazowej metodą cyfrowego przetwarzania obrazów powierzchni w świetle widzialnym”.



Dr hab. inż. Józef BŁACHNIO profesor nadzwyczajny na Wydziale Mechanicznym Politechniki Białostockiej. Reprezentowane dziedziny, dyscypliny i specjalności: nauki techniczne, budowa eksploatacja i diagnostyki maszyn: samoloty i śmigłowce, silniki lotnicze, inżynieria materiałowa, inżynieria powierzchni.

BADANIA DIAGNOSTYCZNE AMORTYZATORÓW

Janusz GARDULSKI

Politechnika Śląska, Wydział Transportu, Katedra Budowy Pojazdów Samochodowych
40-019 Katowice ul. Krasińskiego 8, email: janusz.gardulski@polsl.pl

Streszczenie

W referacie przedstawiono powszechnie stosowane metody diagnozowania amortyzatorów samochodów osobowych oraz zaproponowano nowe metody oparte na wibroakustycznej analizie sygnałów drganiowych

Słowa kluczowe: amortyzatorów samochodowy, diagnostyka.

DIAGNOSTIC OF SHOCK ABSORBERS

Summary

The paper presents generally used methods of passenger car shock-absorbers diagnosis and new methods based on vibroacoustic analysis of vibration signals was proposed.

Keywords: car shock absorber, diagnostic.

1. WPROWADZENIE

Jednym z podzespołów samochodowych decydujących o bezpieczeństwie, komforcie jazdy, kierowności pojazdu jest jego zawieszenie. Obecnie w samochodach osobowych stosowane są wyłącznie zawieszenia niezależne składające się z:

- elementów sprężystych,
- elementów tłumiących,
- zespołu łączników, łączących osie kół lub koła z ramą względnie nadwoziem samochodu za pomocą połączeń sworzniowych,
- opon charakteryzujących się określoną sprężystością i tłumieniem.

Ich łączne masy są masami nieresorowanymi pojazdu, które pobudzane są do drgań podczas jazdy zwłaszcza przez nierówności drogi. Drgania to po odpowiednim wzmocnieniu przez elementy sprężyste przenoszone są na masy resorowane, a więc nadwozie i znajdujących się nim kierowcę i pasażerów. Przed konstruktorami zawieszeń samochodów osobowych zawsze stoi trudne zadanie odpowiedniego doboru nie tylko rodzaju zawieszenia, ale również charakterystyk ich elementów sprężystych i tłumiących. Aktualnie w samochodach osobowych elementy te posiadają charakterystyki nieliniowe.

W samochodach osobowych klas średnich i niższych powszechnie elementami sprężystymi są stalowe sprężyny śrubowe, charakteryzujące się prostotą konstrukcyjną, łatwością kształtowania, małą masą i niską ceną. Ich nieliniową charakterystykę sztywności uzyskuje się poprzez zmianę skoku, średnicy zewnętrznej, względnie średnicy drutu. Elementami tłumiącymi drgania w samochodach osobowych są amortyzatory teleskopowe. Dobór ich charakterystyki tłumienia jest trudny. Mogą one mieć charakterystykę

sportową tzw. twardą, lub komfortową (miękką). Najczęściej w samochodach osobowych konstruktorzy dobierając typ amortyzatora przyjmują rozwiązanie kompromisowe. Charakterystyki stosowanych amortyzatorów są nieliniowe, niesymetryczne – najczęściej progresywne lub progresywno-degresywne. Należy przy tym pamiętać, że na tłumienie drgań, poza amortyzatorami mają wpływ opony, sprężyny, przeguby kuliste, tuleje metalowo – gumowe i stabilizatory. W większości przypadków o wartości tłumienia decydują jednak amortyzatory. Charakterystyki elementów sprężystych i tłumiących w trakcie eksploatacji ulegają zmianie. W zawieszeniu pojawiają się nowe nieliniowości (luzy) wpływające na bezpieczeństwo jazdy. Efektem niesprawności amortyzatorów są:

- duże amplitudy przyspieszeń drgań nadwozia,
- powiększone przechyły poprzeczne i podłużne samochodu decydujące o komforcie jazdy i stabilności ruchu pojazdu,
- odrywanie się kół od nawierzchni podczas jazdy i hamowania,
- wydłużenie drogi hamowania.

Ponadto występujące podczas jazdy w zawieszeniu nadmierne siły dynamiczne powodują szybsze, zmęczeniowe zużycie elementów pojazdu oraz niszczenie nawierzchni drogi. Niezbędne są zatem badania okresowe zawieszeń celem których jest identyfikacji ich stanu technicznego.

Badania diagnostyczne zawieszeń poza pomiarami geometrii, podzielić można na dwie grupy obejmujące:

- identyfikację luzów w połączeniach sworzniowych,
- ocenę stanu technicznego amortyzatorów.

Diagnozowanie luzów w połączeniach elementów zawieszonych jest stosunkowo proste. Do tego celu stosuje się urządzenia do wymuszenia szarpnięć kołami jezdny. Urządzenia te składają się z dwóch płyt najazdowych przesuwanych hydraulicznie w czterech kierunkach. Badanie organoleptyczne polega na obserwacji sprawdzanych połączeń elementów oświetlanych silnym strumieniem lampy sterującej, uwidaczniającej luzy. Pomiary diagnostyczne amortyzatorów są bardziej złożone. Prawidłowa praca zawieszonych pojazdów samochodowych wpływa bezpośrednio na bezpieczeństwo ruchu samochodowego oraz stan nawierzchni dróg. W trakcie badań okresowych prowadzonych w Stacjach Kontroli Pojazdów (SKP) szczególną uwagę zwraca się na stan techniczny zawieszonych. Amortyzatory, jako elementy odpowiedzialne za prawidłowe tłumienie drgań nadwozia, badane są w SKP metodą bezdemontażową, wg kryteriów ustalonych przez Europejskie Stowarzyszenie Producentów Amortyzatorów (EUSAMA). Głównym mankamentem tej metody jest brak informacji o rodzaju powstającego uszkodzenia i stopniu jego zaawansowania w trakcie eksploatacji pojazdu. Najczęściej występującymi usterkami amortyzatorów są wycieki płynu i ubytek uszczelnienia tłoczka. Uszkodzenia te są trudno diagnozowalne pod względem ilościowym. Powyższe zagadnienia są przedmiotem przeprowadzonych badań.

Jako nośnik informacji diagnostycznej wykorzystano sygnały wibroakustyczne generowane przez masy pojazdu pobudzone do drgań.

2. METODY DIAGNOZOWANIA AMORTYZATORÓW

Z pośród metod badawczych stanu technicznego amortyzatorów należy wyróżnić się badania stanowiskowe. Można je podzielić na dwie grupy:

- badanie amortyzatorów zabudowanych w pojeździe,
- badanie amortyzatorów wybudowanych z pojazdu.

2.1. BADANIA AMORTYZATORÓW ZABUDOWANYCH W POJEŹDZIE

Wśród metod badawczych amortyzatorów zabudowanych w samochodzie rozróżnić należy metody:

- drgań wymuszonych,
- drgań swobodnych,
- płyt najazdowych.

Metoda trzecia ze względu na małą dokładność praktycznie nie jest stosowana.

Grupa metod drgań wymuszonych sprowadza się do testerów wibracyjnych:

- o zmiennej amplitudzie drgań,
- o stałej amplitudzie drgań.

Istotą obu metod jest wymuszanie drgań pionowych badanego koła jezdno powyżej częstotliwości rezonansowej. W trakcie badań koło spoczywa na płycie włączonego wzbudnika drgań. Po przekroczeniu częstotliwości rezonansowej następuje wyłączenie siły wymuszającej. Na skutek tłumienia drgań przez amortyzator, elementy zawieszenia i opony – następuje obniżanie częstotliwości drgań. Wskaźnikiem stanu technicznego amortyzatora jest amplituda występująca przy częstotliwości rezonansowej zawieszenia. Efektywność tłumienia można określić na podstawie analizy drgań:

- w funkcji nacisku koła na podłoże (metoda EUSAMA),
- w funkcji czasu (metoda Boge).

2.1.1. METODA EUSAMA

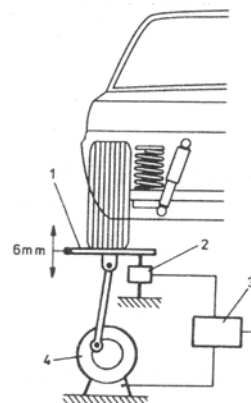
Europejskie Stowarzyszenie Producentów Amortyzatorów (EUSAMA) opracowało metodę oceny efektywności tłumienia. Zasada pomiaru polega na procentowym określeniu siły przylegania koła do podłoża. Ocena skuteczności tłumienia amortyzatora określa wskaźnik Eusama opisany zależnością:

$$WE = \frac{W_{\min}}{W_{st}} \cdot 100\%$$

gdzie:

- $W_{\min}$ – zmierzona minimalna siła dynamiczna przylegania opony do podłoża,
- W_{st} – statyczna siła przylegania opony do podłoża (spoczynkowa).

Schemat stanowiska badawczego przedstawia rys. 1.

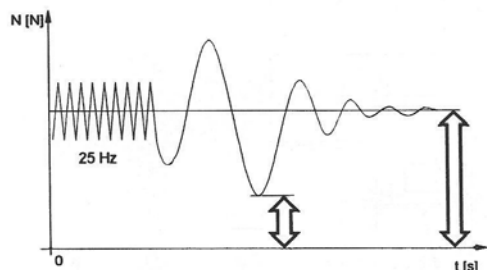


Rys. 1. Schemat ideowy stanowiska badawczego badań amortyzatorów wg metody EUSAMA: 1) płyta najazdowa, 2) układ pomiarowy, 3) układ analizujący, 4) silnik elektryczny z korbowym układem wymuszającym

W trakcie badania koło samochodu spoczywa na nieruchomej płycie urządzenia wymuszającego. Dokonywany jest wówczas pomiar statyczny (W_{st}). Dla takiego przypadku wartość współczynnika WE wynosi 100%. Następnie uruchamiany jest układ wymuszający. Płyta osiąga drgania amplitudzie 4-8 mm z częstotliwością około 25 Hz. Po wyłączeniu układu napędowego przy częstotliwości ok. 16 Hz

następuje pomiar minimalnej siły nacisku dynamicznego W_{min} koła jezdnego na płytę. Jeżeli podczas badania przy tej częstotliwości koło oderwie się od płyty, to wartość WE wynosi 0%. Ta filozofia pomiaru jest przejrzysta i logiczna. W wyżej wymienionej metodzie nie jest potrzebna baza danych.

Przykładowy przebieg rejestrowanego sygnału przedstawia rys. 2.



Rys 2. Graficzna interpretacja współczynnika EUSAMA

Wadą opisywanej metody badań jest wrażliwość na sztywność opony i obciążenie statyczne. Sztywność opony zależy między innymi od ciśnienia w ogumieniu. Niskie ciśnienie podwyższa a wysokie obniża wartość współczynnika EUSAMA. Odwrotnie - duże obciążenie statyczne podwyższa, a niskie obniża wartość tego współczynnika.

Przyjęte kryterium oceny jest następujące:

- WE = 0-20% - zły stan techniczny amortyzatorów (niedostateczna wartość tłumienia),
- WE = 21-40% - dopuszczalna wartość tłumienia - amortyzator wymaga sprawdzenia po wybudowaniu na stanowisku indykatorowym,
- WE = 41-60% - dobra wartość tłumienia,
- WE > 60% - bardzo dobra wartość tłumienia.

Różnice wartości EUSAMA pomiędzy stronami pojazdu wynoszące 20-30% świadczą o niesprawności amortyzatora. Wątpliwości budzą testy skuteczności tłumienia drgań zawieszenia osi tylnej lekkich samochodów z przednim napędem. Rozwiązanie tego problemu proponuje firma Hoffman dzieląc samochody na cztery grupy w zależności od masy.

Grupa I - Pojazdy klasy wyższej o masie powyżej 1400 kg otrzymują następujące kryteria oceny

WE	70-100%	stan bardzo dobry,
WE	50-69%	stan dobry,
WE	30-49%	stan dostateczny,
WE	0-29%	stan niedostateczny

Dopuszczalna różnica przyczepności $D_{RP} = 25\%$.

Grupa II - Pojazdy o klasie średniej o masie 900-1399 kg

WE	60-100%	stan bardzo dobry,
WE	40-59%	stan dobry,
WE	20-39%	stan dostateczny,
WE	0-19%	stan niedostateczny.

$D_{RP} = 25\%$.

Grupa III - Pojazdy kompaktowe o masie własnej mniejszej od 899 kg

WE	60-100%	stan bardzo dobry,
WE	30-59%	stan dobry,
WE	20-29%	stan dostateczny,
WE	0-19%	stan niedostateczny.

$D_{RP} = 20\%$.

Uzyskane w wyniku badań metodą EUAMA wskaźniki mają charakter orientacyjny. Wprowadzając banki danych można uzyskać informacje o wartościach odniesienia (np. z badania nowych pojazdów) względnie obserwować trend zmian w trakcie eksploatacji. Przyjęcie w metodzie stałej (16 Hz) częstotliwości rezonansowej ogranicza możliwości metody. Poszczególne koła mogą mieć takie same wartości WE np. 60% ale osiągane przy różnej częstotliwości rezonansowej poszczególnych kół. Takie różnice mają wpływ na bezpieczeństwo jazdy. Było to powodem opracowania zmodyfikowanej metody badań.

2.1.2. METODA BADAŃ AMORTYZATORÓW NA STANOWISKU PŁYTOWYM O ZMIENNEJ CZĘSTOTLIWOŚCI DRGAŃ WYMUSZAJĄCYCH

Metoda badań amortyzatorów w oparciu o zmienną częstotliwość rezonansową występuje pod nazwą EUSAMA plus. W tym systemie badań występują dwa etapy pomiarowe

- faza wstępna
- pomiar współczynnika tłumienia

Faza wstępna, to rozgrzewanie płynu amortyzatorowego celem uzyskania właściwej jego lepkości. Proces ten trwa około 10 sek. przy wymuszeniu o niskiej częstotliwości drgań. W tym czasie dokonywany jest pomiar ciśnienia w ogumieniu. Analizowany jest kontakt płyty z oponą. Jeżeli odchyłka ciśnienia powietrza przekracza 0,05 MPa ciśnienia nominalnego, to następuje przerwanie procesu pomiarowego, sygnalizowanego odpowiednim komunikatem.

Po rozgrzaniu - rozpoczyna się proces pomiarowy ze zmieniającą się co 1 Hz w sposób malejącą częstotliwością drgań płyty od 30 Hz do 8 Hz. Częstotliwość określana jest przez przetworniki optoelektroniczne. Na czas pomiaru przetworniki stabilizują każdą częstotliwość drgań. Szczegółowo analizowany jest przedział 13 do 18 Hz (rezonans masy nieresorowanej) Dla każdego zakresu częstotliwości zmierzone są siły nacisku na płytę podczas ściskania i rozciągania amortyzatorów. Umożliwia to określenie dla każdego typu pojazdu stosunku mas:

$$c = \frac{m_r}{m_n}$$

gdzie: m_r - masa resorowana

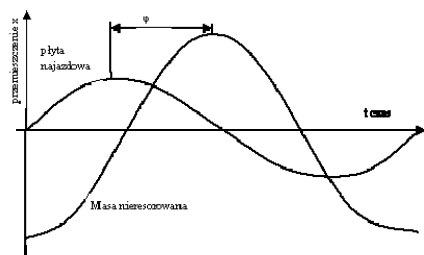
m_n - masa nieresorowana

Precyzyjnie zmieniając częstotliwość drgań płyty stanowiska i uwzględnienie stałej sprężystości ogumienia (ciśnienia powietrza w oponie) świadczy o nowoczesności tej metody oceny stanu

zawieszenia (amortyzatorów). Wprowadzenie fazy rozgrzewania amortyzatorów wpływa korzystnie na dokładność pomiarów.

2.1.3. METODA EUSAMA Z ANALIZĄ FAZOWĄ

Firma Hunter Engineering Company wprowadziła dalsze zmiany w metodzie EUSAMA. Jest nią mierzony kąt przesunięcia fazowego pomiędzy sinusoidalnymi sygnałami przemieszczenia płyty i siły nacisku koła na płytę (rys. 3).



Rys. 3. Przesunięcie fazowe płyty i mas niesorowanych

Przesunięcie masy niesorowanej jest proporcjonalne do chwilowej siły nacisku wywieranego na płytę najazdową. Dlatego osiąga maksimum, gdy kąt fazowy jest równy 0. W zakresie pasma częstotliwości rezonansowej masy resorowanej, gdy brak jest tłumienia w zawieszeniu, wartość kąta fazowego wynosi 180° pomiędzy maksymalną amplitudą przemieszczeń masy niesorowanej i resorowanej.

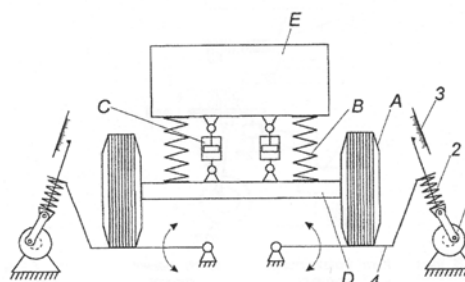
Wartość kąta przemieszczenia fazowego jest wielkością charakteryzującą wielkość tłumienia. Kiedy zawieszenie samochodu ma odpowiednią wartość tłumienia (zamontowany jest odpowiedni amortyzator) to minimalny kąt przesunięcia fazowego pomiędzy częstotliwościami rezonansowymi masy resorowanej i niesorowanej osiągnie wartość powyżej 90° .

2.1.4. METODA BOGE DRGAŃ WYMUSZONYCH W FUNKCJI CZASU

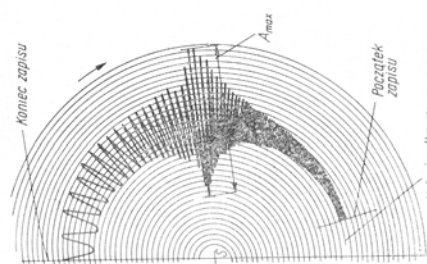
Metoda stosowana przez firmę Boga polega na zastosowaniu do pobudzania drgań stanowiska płytowego, którego ruch wymuszony jest przez układ korbowy. Schemat wyżej wymienionego stanowiska przedstawia rys. 4.

Przebieg drgań jednego koła rejestrowany jest na połowie okrągłej tarczy. Przykładowy wykres przedstawia rys. 5.

Miarą jakości amortyzatora jest podwójna amplituda drgań rezonansowych (A_{max}). Wartość tej amplitudy porównuje się z amplitudą wzorcową amortyzatorów dla danych typów pojazdów. Zastosowanie metody pomiaru dekrementu tłumienia jest słuszne dla układów liniowych.



Rys. 4. Schemat urządzenia firmy Boga:
1) silnik elektryczny, 2) sprężyna, 3) rejestrator,
4) płyta najazdowa, A – badane koło, B – element resorujący, C – amortyzator, D – badana oś, E – nadwozie.



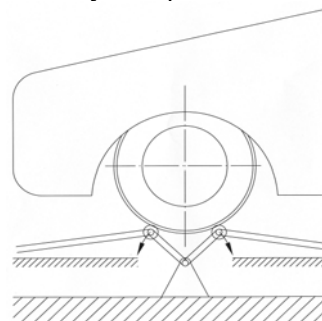
Rys. 5. Przykładowy wynik badań amortyzatora na urządzeniu Boga

2.2. BADANIA AMORTYZATORÓW METODĄ DRGAŃ SWOBODNYCH

Istota metody polega na analizie ruchu podwozia samochodu osobowego po impulsowym pobudzeniu go do drgań. Istnieje wiele metod wymuszenia a mianowicie:

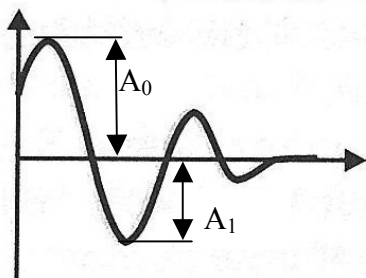
- zrzut,
- spadek z równi pochyłej,
- rozprężenie ściśniętego nadwozia.

Z pośród powyższych metod najszerze zastosowanie znalazła metoda pierwsza. Do tego celu wykorzystuje się najazdowe stanowisko badawcze schematycznie przedstawione na rys. 6.



Rys. 6. Schemat stanowiska

Opadające koło uderza o płytę oporową pobudzając do drgań masę niesorowaną i resorowaną. Zapis amplitudy przemieszczeń drgań w funkcji czasu rejestrowany jest w formie wykresowej. Przykładowy wynik rejestracji przedstawia rys. 7.



Rys. 7. Wynik rejestracji badania metodą „zrzutu” samochodu

Z powyższego wykresu wyznaczona się średnie tłumienie drgań „k” będące stosunkiem drugiej do pierwszej zaobserwowanej amplitudy drgań.

$$k = \frac{A_1}{A_0}$$

Wartości amplitud początkowych powinny być sobie równe. Nie powinny się one różnić o więcej niż 7%.

2.3 BADANIA AMORTYZATORÓW NIE ZABUDOWANYCH W POJEŹDZIE

W badaniach amortyzatorów nie zabudowanych w pojeździe stosuje się stanowiska indykatorowe, na których otrzymuje się wykresy prędkościowe amortyzatorów tzn. zmiany siły tłumienia w funkcji prędkości ruchu tłoka amortyzatora. Przykład takiego stanowiska, które znajduje się w Laboratorium Podstaw Techniki Pomiarowej Wydziału transportu Politechniki Śląskiej przedstawia rys. 8.



Rys. 8. Widok stanowiska indykatorowego

Stanowisko napędzane jest dwubiegowym silnikiem elektrycznym prądu przemiennego. Napęd od silnika przenoszony jest przez przekładnię pasowo-klinową na zębatą skrzynię biegów o dwóch biegach oraz przekładnię nawrotną. Ze skrzynki biegów napęd przenoszony jest na mimośrodowy układ korbowy o zmiennym ramieniu. Przeniesienie wymuszenia ruchu końca amortyzatora z układu korbowego odbywa się za

pomocą mechanizmu suwakowego. Poruszany jest dolny koniec amortyzatora prowadzony osiowo przez układ suwakowy. Górny koniec połączony jest z obudową stanowiska poprzez przetwornik siły. Wykonując badania przy różnych prędkościach obrotowych wału uzyskuje się w efekcie charakterystyki prędkościowe amortyzatora opisane zależnością:

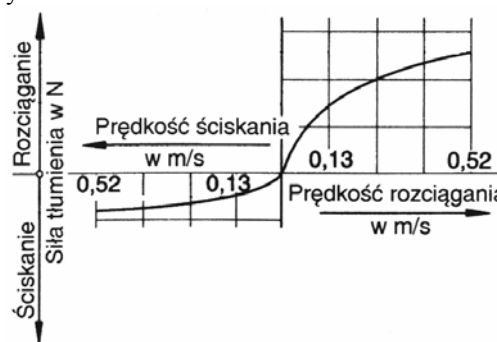
$$F(t) = k_1 \left(\frac{dx}{dt} \right) + k_2 \left(\frac{dx}{dt} \right)^i \quad [N]$$

gdzie:

$k_{1,2}$ – współczynniki określające liniową i nieliniową część tłumienia,

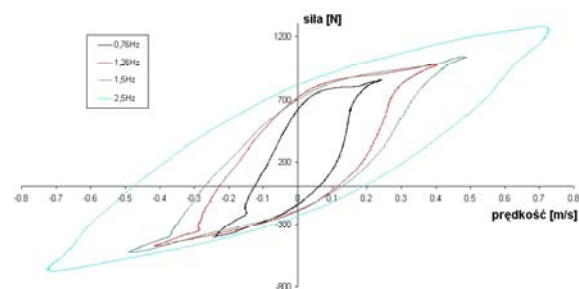
dx/dt – prędkość względna ruchu tłoka amortyzatora względem cylindra [m/s],

i – wykładnik potęgi charakteryzujący nieliniowość siły tłumienia.

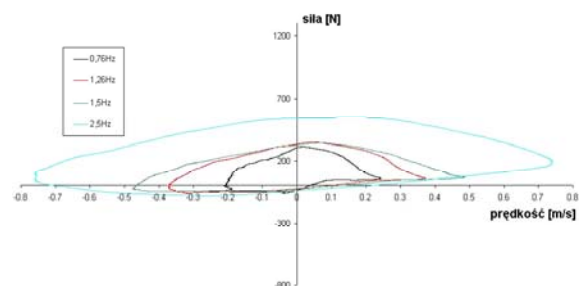


Rys. 9. Przykładowa teoretyczna charakterystyka amortyzatora hydraulicznego

Przykładowe wyniki badań indykatorowych uzyskane dla czterech częstotliwości wymuszeń przedstawiono na rys. 10, 11.



Rys. 10. Wykresy prędkościowe amortyzatora nowego



Rys. 11. Wykresy prędkościowe dla amortyzatora z 75% ubytkiem płynu

Do oceny stanu technicznego amortyzatora wprowadzono nowy parametr diagnostyczny – moc tłumienia. Uzyskuje się go poprzez całkowanie wykresów indykatorów. Powszechnie stosowane są amplitudowe wskaźniki oceny. Zastosowana metoda diagnozowania oparta na mocy tłumienia wykazuje dobre cechy diagnostyczne.

3. METODY BADAŃ AMORTYZATORÓW OPARTE NA WIBROAKUSTYCZNEJ ANALIZIE SYGNAŁÓW DRGANIOWYCH

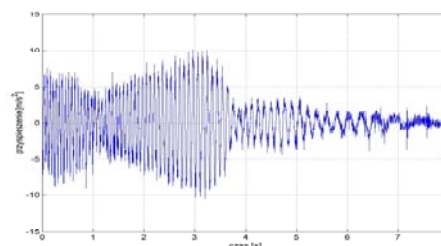
Stanowiskowe metody badań dają wyniki, które można traktować jako przybliżone. Oparta na nich ocena stanu technicznego amortyzatora jest dwustanowa (sprawny, niesprawny). Pozbawiona jest cech identyfikacji jego uszkodzeń. Stało się to podstawą do na opracowywania Laboratorium PTP Wydziału Transportu Politechniki Śląskiej nowych metod diagnozowania amortyzatorów opartych na analizie sygnałów wibroakustycznych. W zaproponowanych metodach założono, że w badaniach zostaną wykorzystane istniejące na stanowiskach kontrolnych amortyzatorów wzbudniki pobudzające do drgań masy kół i zawieszenia oraz nadwozie. W procesie pobudzania wymienionych mas przez układ wymuszający można wyszczególnić trzy następujące fazy:

- rozpędzanie, kiedy następuje przyrost częstotliwości wymuszeń od 0 do częstotliwości górnej (ok. 25 Hz),
- wymuszanie ze stałą górną częstotliwością wymuszeń,
- wybieg urządzenia, podczas którego po wyłączeniu następuje spadek częstotliwości do 0.

Istota zaproponowanej metody badań jest następująca:

- w badaniach analizowany jest proces drganiowy zawieszenia podczas wybiegu, kiedy następuje przejście przez rezonans pobudzonych do drgań mas nieresorowanych i resorowanych. Analizowane drgania są nieustalonym, niestacjonarnym procesem losowym.
- w badaniach amortyzatora wykorzystywane są dwa przetworniki mocowane za pomocą magnesu do dolnego wahacza i gniazda osadczego amortyzatora w nadwoziu. Przetworniki mocowane są osiowo w pobliżu miejsc zabudowy amortyzatorów samochodzie (rys. 13 i 14). Wielkościami mierzonymi są przyspieszenia względne ruchu tłoka amortyzatora. W trakcie pomiarów koło badanej części zawieszenia spoczywa na płycie wzbudnika.
- po dyskretyzacji różnicowy sygnał wibroakustyczny końców amortyzatora poddany jest analizie w środowisku Matlab przy użyciu komputera klasy PC. Wymienione sygnały niestacjonarne poddawane są niżej omówionym analizom dwuwymiarowym, w dziedzinie czasu

i częstotliwości. Przykładowy przebieg drgań nadwozia przedstawiono na rys. 12



Rys. 12. Przyspieszenia drgań nadwozia



Rys. 13. Sposób mocowania przetwornika pomiarowego na nadwoziu

1- pojemnościowy przetwornik przyspieszeń



Rys. 14. Sposób mocowania przetwornika pomiarowego na wahaczu

3.1. KRÓTKOCZASOWA TRANSFORMATA FOURIERA

Skuteczną, stosowaną powszechnie do dziś metodę analizy sygnałów w dziedzinie czasowo-częstotliwościowej zaproponował Gabor (1946 r.). Polega ona na prowadzeniu analizy częstotliwościowej kolejnych fragmentów analizowanego sygnału wyciętych z całości przez tzw. funkcję okna. Podczas analizy okno jest przesuwane w dziedzinie czasu o interwał równy szerokości okna. Analiza częstotliwości prowadzona jest dla kolejnych odcinków niezależnie, co umożliwia powiązanie poszczególnych składowych okresowych widma z czasem i określenie momentów ich występowania. Gabor zaproponował funkcję okna w postaci gausoidy:

$$g(t) = e^{-\frac{1}{2}t^2}$$

oraz sformułował zasadę konstruowania rodziny funkcji analizujących powstałych z funkcji okna przez wprowadzenie przesunięć w dziedzinie czasu b i w dziedzinie częstotliwości ω :

$$w_{\omega,b}(t) = g(t-b)e^{-i\omega t}$$

Wynikiem analizy czasowo- częstotliwościowej sygnału $f(t)$ są współczynniki w postaci:

$$S(\omega,b) = \int_{-\infty}^{+\infty} x(t) \cdot w_{\omega,b}(t) dt = \int_{-\infty}^{+\infty} x(t) \cdot g(t-b)e^{-i\omega t} dt$$

Z powyższych zależności wynika następujący nadający się do praktycznego zastosowania wzór opisujący funkcję analizującą:

$$W_{\omega,b} = e^{-\frac{1}{2} \left(\frac{t-b}{d} \right)^2} \cdot e^{-i\omega t}$$

gdzie:

b – przesunięcie okna

ω – częstości analizujące

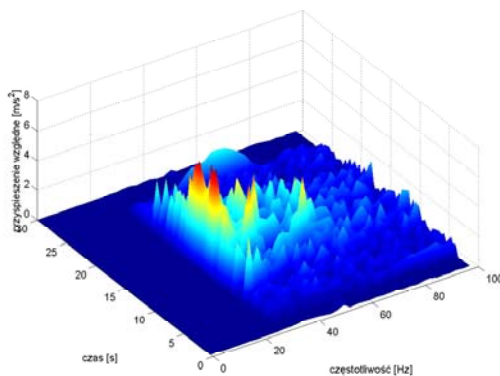
d – szerokość okna

W powyższym wzorze występują trzy współczynniki, z których jeden b odpowiada z przesunięcie okna (parametr czasowy) drugi ω za wartość częstości analizującej wypełniającej okno (parametr częstotliwościowy) trzeci d niezmienny w całej rodzinie funkcji analizujących (parametr wyznaczający szerokość okna).

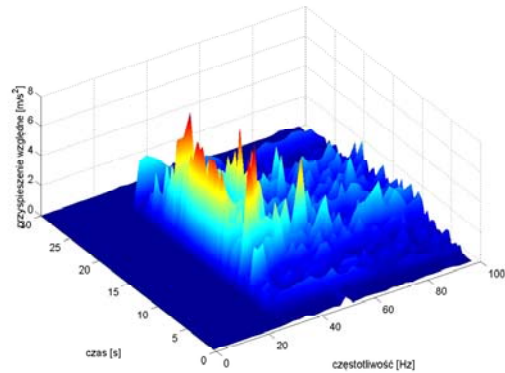
Metoda zaproponowana przez Gabora zwana także krótkoczasową transformatą Fouriera (Short Time Fourier Transform – STFT) lub okienkowym przekształceniem Fouriera opiera się na stałej raz dobranej szerokości okna analizy, które ogranicza jednocześnie precyzję w dziedzinie czasu i zakres częstotliwości analizy.

Optymalnie prowadzona analiza o dokładności ograniczonej jedynie zasadą nieoznaczoności Heisenberga ma miejsce w dolnym zakresie częstotliwości tj. w przypadku funkcji analizujących $W_{\omega,b}$, których okres $1/\omega$ jest równy szerokości d okna gaussoidy. Wszystkie wyższe częstotliwości analizowanego sygnału są wykrywane z tą samą dokładnością lecz znacznie mniejszą od możliwej do uzyskania wynikającej z zasady nieoznaczoności.

Przykładowe wyniki analiz dla amortyzatorów w stanie nowym i 25 % wyciekem płynu przedstawiono na rys. 15 i 16.



Rys. 15. Widmo czasowo-częstotliwościowe przyspieszeń względnych uzyskane dla amortyzatora nowego (100% objętości płynu hydraulicznego)



Rys. 16. Widmo czasowo-częstotliwościowe przyspieszeń względnych uzyskane dla amortyzatora z ubytkiem 25% płynu

Objawy pogłębiającej się niesprawności amortyzatora w mapach spektralnych uwiadcniają się w postaci rosnących wartości amplitud nadrezonansowych częstości harmonicznych. Już niewielka zmiana charakterystyki amortyzatora spowodowana ubytkiem płynu hydraulicznego wywołuje zmiany stosunku wartości amplitud nadrezonansowych. Dalszy rozwój uszkodzenia powoduje gwałtowny wzrost wartości amplitud przyspieszeń w chwilach przechodzenia układu przez częstości drgań własnych zawieszenia widoczny również dla częstości nadrezonansowych. Klasyczne metody diagnozowania stanu technicznego amortyzatora zabudowanego w zawieszeniu wykrywają dopiero duże zmiany jego charakterystyki. Analiza czasowo-częstotliwościowa pozwala uzupełnić wiedzę diagnostyczną informacjami o stosunkowo wczesnych stadiach rozwoju uszkodzenia amortyzatora.

3.2. TRANSFORMATA FALKOWA

Inną metodę analizy czasowo-częstotliwościowej zaproponowali Morlet i Grossman (1984 r.). Podobnie jak w analizie STFT sformułowano funkcję podstawową, reguły konstrukcji na jej podstawie rodziny funkcji analizujących oraz wzór opisujący wynik analizy w dziedzinie t - f (czasowo-częstotliwościowej) w postaci funkcji dwu zmiennych (t i f). Morlet zaproponował funkcję podstawową:

$$\psi(t) = e^{-\frac{t^2}{2}} \cdot \cos(5t)$$

nazwał ją falką-matką (fr. ondelette-mere) ponieważ służy ona do konstrukcji rodziny falek analizujących (fr. ondelette analysamite) $\Psi_{a,b}$;

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \cdot \psi\left(\frac{t-b}{a}\right)$$

gdzie: b jest przesunięciem czasowym;

$a > 0$ jest tzw. współczynnik zwięzania-rozszerzania odpowiada za częstości i czasowy zakres analizy.

Po wprowadzeniu dodatkowego parametru skali a do okna lokalizacyjnego w transformacie Gabora powstała transformata falkowa (Wavelet transform WT).

Funkcja analizująca $\psi\left(\frac{t-b}{a}\right)$ nazywana falką główną posiada współczynnik a który powoduje zmianę czasu trwania falki oraz współczynnik b który zmienia położenie falki na osi czasu. Równanie reprezentuje filtrację pasmowo-przepustową sygnałów za pomocą filtrów o różnych pasmach przepuszczania. Transformata falkowa (WT) jest funkcją dwuwymiarową, w której a – jest parametrem skali (częstotliwości) zaś b – jest parametrem translacji (przesunięcia w czasie).

$$a, b \in R, \quad a \neq 0$$

Współczynniki reprezentujące realizowany sygnał w dziedzinie t-f opisuje zależność:

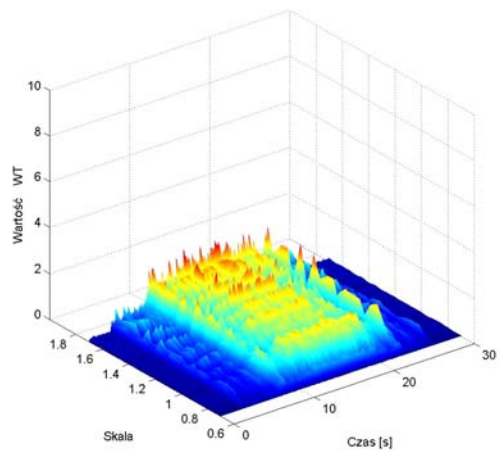
$$WT(a, b) = (x(t) \cdot \psi_{a,b}) = \int_{-\infty}^{+\infty} x(t) \cdot \psi_{a,b}(t) dt = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt$$

Zależność ta podobnie jak podana przez Gabora zależność (3) jest równoważna splotowi sygnału analizowanego $x(t)$ z falką analizującą $\psi_{a,b}(t)$ lub w przypadku analizy Gabora (STFT) z funkcją analizującą $w_{\omega,b}(t)$.

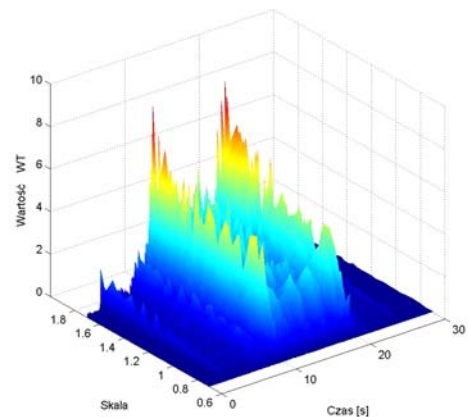
W porównaniu do stosowanych w STFT funkcji okna mających zawsze mniej lub bardziej przybliżoną funkcję gaussoidy, funkcja $\psi(t)$ posiada odmienne cechy. Jest ona parzystą funkcją lokalnie oscylującą a poza spójnym przedziałem oscylacji przyjmuje wartości zerowe. Przebieg funkcji $\psi(t)$ jest podobny do zafalowania, lokalne oscylacje szybko gasną wraz z odległością od centrum – stąd nazwa: falka (ang. Wavelet, fr. ondelette).

W analizie falkowej inna jest w porównaniu z STFT reguła konstruowania rodziny funkcji analizujących. W analizie STFT (Gabora) szerokość okna jest stała co powoduje wzrost oscylacji w oknie wraz ze wzrostem częstotliwości. W analizie falkowej ilość oscylacji falki jest stała a zmianie częstotliwości towarzyszy proporcjonalna zmiana czasowego zasięgu falki.

Analiza niestacjonarnych własności spektralnych sygnału wymaga posługiwania się oknami, które automatycznie zwiększają się przy analizie wysokich częstotliwości i ulegają automatycznemu rozszerzeniu przy analizie niskich częstotliwości.



Rys. 17. Wynik ciągłego przekształcenia falkowego dla przyspieszeń drgań nadwozia dla amortyzatora przedniego samochodu Seicento ze 100% napełnieniem



Rys. 18. Wynik ciągłego przekształcenia falkowego dla przyspieszeń drgań nadwozia dla amortyzatora przedniego samochodu Seicento z 25% ubytkiem oleju

Wartości WT dla amortyzatorów nowych oraz o niewielkich przebiegach eksploatacyjnych są stałe w analizowanym przedziale czas-skala. Zużycie amortyzatora w wyniku dużych przebiegów eksploatacyjnych powoduje wzrost wartości WT w chwilach pracy układu z częstotliwością rezonansową zawieszenia. Takie same efekty uwidaczniają się w rozkładach falkowych uzyskanych dla amortyzatorów z celowo wprowadzonymi usterkami (ubytek płynu oraz uszkodzenie zaworka ruchu sprężania). Starzenie się płynu amortyzatorowego powoduje wzrost wartości amplitud WT w całym analizowanym paśmie.

3.1.4. TRANSFORMATA WIGNERA-VILLE'A (WIGNER-VILLE DISTRIBUTION, WVD)

Widmo wzajemne Wignera-Ville'a dwóch sygnałów $x(t)$ i $y(t)$ definiowane jest następująco:

$$WVD_{xy}(t, f) = \int x\left(t + \frac{\tau}{2}\right) y^* \left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau$$

jeżeli $x(t) = y(t)$ powyższe przekształcenie można zapisać w postaci:

$$WVD_{xx}(\theta, \tau) = \int x^* \left(t - \frac{\tau}{2}\right) x \left(t + \frac{\tau}{2}\right) e^{j2\pi\theta\tau} dt$$

gdzie:

$WVD_{xx}(\theta, \tau)$ – pseudo przekształcenie Wignera-Ville'a

$x^*(t)$ – sygnał urojony sprzężony z $x(t)$

τ – przesunięcie w dziedzinie czasu

θ – przesunięcie w dziedzinie częstotliwości

Transformata ta pozwala przekształcić sygnał zapisany w postaci przebiegu czasowego na widmo czasowo-częstotliwościowe. Program Matlab posiada zaimplementowaną dyskretną wersję tego przekształcenia, która umożliwia obróbkę sygnału zdyskretyzowanego. Ponieważ algorytm realizujący tę transformatę wykorzystuje dwukrotne przekształcenie Fouriera warunkiem uniknięcia zjawiska aliasingu jest próbkowanie sygnału ciągłego z co najmniej dwukrotnie większą częstotliwością od kryterium Nyquista.

W celu redukcji efektów związanych z przeciekiem widma, które utrudniają interpretację uzyskanych wyników, wykonuje się operację filtracji WVD przy pomocy funkcji wagowych. W analizie wykorzystano funkcję wagową Choi-Williams'a w postaci:

$$\phi(\theta, \tau) = \exp(-\theta^2 \tau^2 / \sigma^2)$$

σ – parametr proporcjonalny do amplitudy przecieku widma

Wykorzystane w pracy przekształcenie można zapisać w sposób:

$$WVD(f, \tau) = \int x^* \left(t - \frac{\tau}{2}\right) x \left(t + \frac{\tau}{2}\right) e^{j2\pi f\tau} e^{-\left(\frac{\theta\tau}{\sigma}\right)^2} dt$$

Istotnym problemem związanym z w/w analizami jest forma przedstawienia wyników badań. Istnieje wiele sposobów, wśród których ma szczególną uwagę zasługują:

- wykresy warstwiczne
- wykresy profilowe

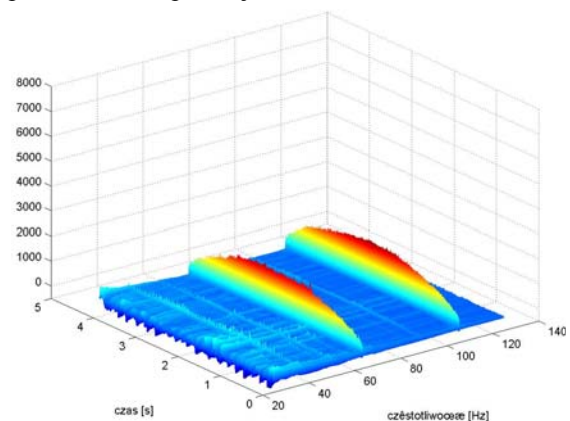
Wykresy warstwiczne są wykresami płaskimi w układzie współrzędnych czas-częstotliwość, na które nanoszone są linie gęstości mocy. Stałe wartości gęstości mocy są reprezentowane przez linie zamknięte (warstwy).

Wykresy profilowe są rysunkami pseudoprzestrzennymi. Z jednej strony ułatwiają one obserwację rozkładu gęstości widmowej mocy w funkcji czasu i częstotliwości. Jednakże „plasterkowy” system wizualizacji może spowodować opuszczenie istotnych wyników, a ponadto na wykresach część wyników zostaje ukryta przez poprzedzające je profile. Rozwiązaniem tego problemu jest wykonanie wykresów zmian mocy w funkcji czasu dla każdej częstotliwości oddzielnie.

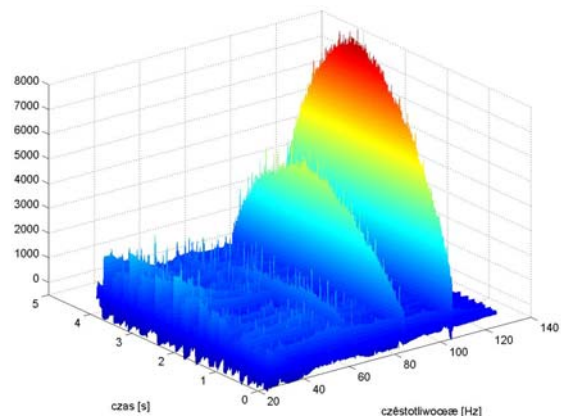
Struktura widmowa sygnałów niestacjonarnych może być określona za pomocą podwójnej transformaty Fouriera niestacjonarnych funkcji korelacji określonych za pomocą wartości oczekiwanych (średnich w zbiorze).

Do analiz wykorzystano pseudoprzekształcenie Wignera-Ville'a z oknem filtracyjnym Choi-Williamsa. Analizy przeprowadzono w środowisku Matlab posiadającym zaimplementowaną dyskretną wersję tego przekształcenia, z postaci przebiegu czasowego na widmo czasowo-częstotliwościowe. Ponieważ algorytm realizujący tę transformatę wykorzystuje dwukrotne przekształcenie Fouriera, w celu uniknięcia zjawiska aliasingu konieczne jest próbkowanie sygnału, z co najmniej dwukrotnie większą częstotliwością od kryterium Nyquista.

Analizom poddano tę część sygnałów drganiowych, w których parametry wymuszenia były stałe w czasie. Układ zawieszenia pracował powyżej podstawowych częstotliwości rezonansowych związanych z masami resorowaną i nieresorowaną. Przykładowy wynik analizy WVD uzyskany dla sprawnego amortyzatora przedstawiono poniżej.



Rys. 19. Widmo WVD przyspieszeń drgań nadwozia – 100% płynu amortyzatorowego



Rys. 20. Widmo WVD przyspieszeń drgań nadwozia – amortyzator z 75% objętości nominalnej płynu amortyzatorowego

4. PROPONOWANA WIBROAKUSTYCZNA METODA DIAGNOZOWANIA AMORTYZATORÓW SAMOCHODÓW OSOBOWYCH ZABUDOWANYCH W POJEŹDZIE

Z omówionych 3 metod analiz niestacjonarnych sygnałów losowych, najkorzystniejsza w badaniach diagnostycznych amortyzatorów okazała się transformata falkowa. Umożliwia ona poprzez dobór odpowiedniego zakresu skali obserwację cech czasowo-częstotliwościowych w wąskim paśmie częstotliwości. Ta właściwość przekształcenia falkowego przy analizie długich szeregów czasowych umożliwia znaczne skrócenie czasu obliczeń

4.1. PROCEDURA ANALIZY

W celu zwiększenia rozdzielczości częstotliwościowej, sygnał poddany jest ponownemu próbkowaniu do częstotliwości 2 [kHz] przy wykorzystaniu funkcji sklepanych. Pasma częstotliwości rezonansowych zawieszenia analizowane jest w zakresie skali 80 do 200 z krokiem co 10 przy wykorzystaniu falki Morleta, której pasmo wynosi: 1,701 [Hz] natomiast częstość środkowa [1/s]. Częstotliwości rezonansowe nadwozia badane są dla zakresu skali 100 do 300 z krokiem 8 falką z rodziny symlet przy częstotliwości próbkowania wynoszącej 0,5 [kHz]. W badaniach stosuje się falkę sym8 (pasmo [Hz] a częstość [1/s]). Samochód pobudzany był w badaniach wzbudnikiem harmonicznym.

4.2. ZASTOSOWANY ESTYMATOR DIAGNOSTYCZNY

Otrzymywany w wyniku analizy falkowej rozkłady czasowo-częstotliwościowy są w powszechnie stosowanej postaci mało przydatne w diagnostyce amortyzatorów. Dlatego poszukiwano takich estymatorów, które umożliwiają jednoznaczną identyfikację rodzaju uszkodzenia. Analiza znanych, powszechnie stosowanych estymatorów nie wykazała oczekiwanych rezultatów. Efektem tych przemyśleń było wyznaczenie nowych estymatorów, które mogą być zastosowane w diagnostyce amortyzatorów. Najkorzystniejsze cechy parametrów diagnostycznych posiadają współczynniki: WT_{zaw} , WT_{nad} , wyznaczone odpowiednio dla częstotliwości rezonansowych mas nieresorowanych i resorowanych. Wartość estymatora WT_{zaw} w przypadku częstotliwości rezonansowej zawieszenia zdefiniowana była następująco:

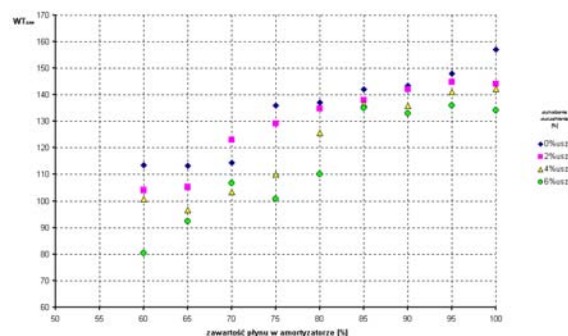
$$WT_{zaw} = \frac{|WT_{max}| + |WT_{min}|}{2},$$

gdzie:

WT_{max} - maksymalna wartość WT w strefie rezonansowej,

WT_{min} - minimalna wartość WT w strefie rezonansowej.

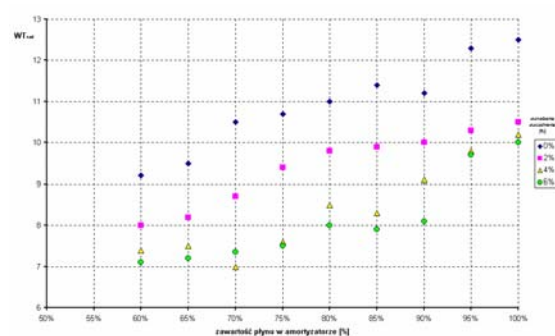
Uzyskane wyniki uśrednione dla amortyzatorów nowych oraz z wyciekami płynu i uszkodzeniami uszczelnienia tłoczka przedstawiono przykładowo na rys. 21



Rys. 21. Przykładowy rozkład wartości współczynnika falkowego WT_{zaw} dla częstotliwości rezonansowej mas nieresorowanych zawieszenia

Maksymalne wartości współczynnika WT_{zaw} występują dla nowego amortyzatora, maleją one wraz z powiększającym się ubytkiem uszczelnienia tłoczka. Dynamika tych zmian jest zachowana dla wszystkich badanych stanów napęnienia amortyzatorów. Ponadto wraz z powiększającym się wyciekami płynu przy określonym ubytku uszczelnienia tłoczka następuje spadek wartości analizowanego współczynnika.

Na rys. 22 przedstawiono wartości współczynników WT_{nad} w strefie rezonansu nadwozia.



Rys. 22. Wartości współczynników falkowych WT_{nad} przy rezonansie nadwozia

Analiza wartości w/w współczynnika w paśmie częstotliwości rezonansowej nadwozia pozwala na stwierdzenie, iż wraz z pogorszeniem się stanu technicznego amortyzatora następuje obserwowalny wzrost wartości WT_{nad} w czasie, gdy działa wymuszenie stanowiska oraz krótko po jego wyłączeniu. Po zmniejszeniu się częstotliwości drgań w chwilach przechodzenia układu przez

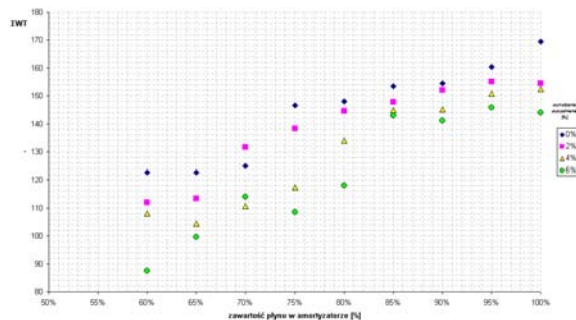
częstotliwość rezonansu nadwozia następuje zmniejszenie wartości WT_{nad} .

Sumę współczynników falkowych ΣWT zdefiniowaną jako:

$$\Sigma WT = WT_{nad} + WT_{zaw}$$

przyjęto jako nowy estymator diagnostyczny, który jako klasyfikator może być zastosowany w identyfikacji rodzaju uszkodzeń eksploatacyjnych amortyzatorów zabudowanych w pojeździe.

Dla przykładowych, omawianych wyników badań postać klasyfikatora przedstawia się następująco (rys. 23).



Rys. 23. Klasyfikator diagnostyczny amortyzatorów oparty na estymatorze ΣWT

4.3. WNIOSKI

Analizowany estymator posiada dużą czułość na zmiany stanu technicznego amortyzatora. Wartości ΣWT maleją wraz z pogarszaniem się jego stanu. Występujące zmiany jego wartości pozwalają na identyfikację stopnia uszkodzenia amortyzatora. Posiada więc te cechy diagnostyczne, których brak powszechnie stosowanym metodom diagnozowania amortyzatorów zabudowanych w samochodzie (np. EUSAMA). Powyższa metoda po opracowaniu automatycznych systemów diagnozowania – co jest przedmiotem dalej prowadzonych prac – może być wykorzystywana w Stacjach Kontroli Pojazdów przy wykorzystaniu istniejących wzbudników drgań.

LITERATURA

1. Batko W., Ziółko M.: *Zastosowanie teorii falek w diagnostyce technicznej*. Problemy inżynierii mechanicznej i robotyki. Kraków, 2002.
2. Bendat J. S., Piersol A. G.: *Metody analizy i pomiaru sygnałów losowych*. PWN, Warszawa 1967.
3. Białasiewicz J. T.: *Falki i aproksymacje*. WNT, Warszawa 2000.
4. Bocheński C., Lozia Z., Mikołajczuk J.: *Próba obiektywnej oceny metody badań amortyzatorów zamontowanych w pojeździe zalecanej przez zrzeszenie EUSAMA*. X Konferencja Diagnostyka Maszyn Roboczych i Pojazdów, Bydgoszcz-Borówno 1999, str. 59-70.
5. Gardulski J., Łazarz B.: *The Influence of the Kinds of Kinematics Forcing, Coming from Uneven Road Surface, on the Acceleration of Motor-Car Vertical Vibrations*. INTER-NOISE, Floryda, USA, V. 1, 1999, s. 751-756.
6. Gardulski J., Warczek J.: *Analiza dwuwymiarowa przyspieszeń drgań pojazdu samochodowego w aspekcie diagnostyki układu zawieszenia*. XXX Jubileuszowe Ogólnopolskie Sympozjum Diagnostyka Maszyn, Węgierska Góra, 2003, str. 38.
7. Gardulski J., Warczek J.: *Diagnostics of Shock Absorbers Built in a Vehicle Based on Vibration Energy Analysis*. Machine Dynamics Problem, 2003, Vol. 27, No 3.
8. Gardulski J., Warczek J.: *Diagnostyka amortyzatorów zabudowanych w pojeździe oparta na analizie efektów nieliniowych*. Zeszyty Naukowe Politechniki Śląskiej, z. 48, ser. Transport, Gliwice 2003, str. 29-36.
9. Gardulski J., Warczek J.: *Identyfikacja stanu technicznego hydraulicznego amortyzatora teleskopowego na podstawie analizy jego charakterystyki*. XXX Jubileuszowe Ogólnopolskie Sympozjum Diagnostyka Maszyn, Węgierska Góra, 2003, str. 39.
10. Gardulski J., Warczek J.: *Moc tłumienia jako parametr diagnostyczny amortyzatorów samochodowych*. DIAGNOSTYKA, vol. 29, 2003r. Warszawa, str. 69-72.
11. Gardulski J., Warczek J.: *Zastosowanie transformaty falkowej do wyznaczania parametrów diagnostycznych elementów zawieszeń samochodowych*. III Międzynarodowy Kongres Diagnostyki Technicznej, Poznań 2004.
12. Gardulski J., Warczek J.: *Wykorzystanie analizy czasowo-częstotliwościowej w diagnostyce zawieszeń pojazdów samochodowych*. Przegląd Mechaniczny nr 7-8 2003, str. 46-48.
13. Gardulski J., Warczek J.: *Względne przyspieszenia drgań jako podstawa diagnostyki elementów zawieszeń pojazdów samochodowych*. V Krajowa Konferencja Diagnostyka Techniczna Urządzeń i Systemów. 13-17.X.2003r. Ustroń, str. 93-99.
14. Gardulski J., Warczek J.: *Zastosowanie krótkoczasowej transformaty Fouriera w diagnostyce zawieszeń pojazdów samochodowych*. Zeszyty Naukowe Politechniki Śląskiej, nr 1562, 2002r, str. 23-29.
15. Gardulski J.: *Bezstanowiskowa metoda oceny stanu technicznego zawieszeń samochodów osobowych*. Wydawnictwo Instytutu Technologii Eksploatacji, Radom 2003.
16. Gardulski J.: *Simulation studies of mechanical system with non-linear parameters of the structure for operating construction needs*.

- Machine Dynamics Problems. Vol. 23 No 3, 1999 (5).
17. Gardulski J.: *Symptom wibroakustyczny jako miara diagnostyczna stanu technicznego amortyzatorów samochodów*. Mat. II Międzynarodowego Kongresu Diagnostyki Technicznej DIAGNOSTYKA 2000, Warszawa 19-22.09.2000, s. 237-242.
 18. Gardulski J.: *The Simulating Researches the Influences of Operational Wear of a Suspension Spring with Nonlinear Characteristics on Motor-car Body Vibrations*, INTER-NOISE, Florida USA, V. 1, 1999, s. 867-870.
 19. Gardulski J.: *Use of Non-Linear Effects in Vibroacoustic Diagnosis*. INTRONAC-7. V.B, pp. 230-237, 1999.
 20. Sikorski J.: *Amortyzatory pojazdów samochodowych – budowa, badania, naprawa*. WKiŁ, Warszawa 1984.
 21. Sitek K.: *Diagnozowanie samochodów w zakresie bezpieczeństwa jazdy*. Spółka wydawnicza „AS”, Piła 1992.
 22. Stańczyk T. L., Strachowski P.: *O możliwości ulepszenia metody badania amortyzatorów wg zaleceń EUSAMA*. II Konferencja Problemy Bezpieczeństwa w Pojazdach Samochodowych, Kielce 2000, str. 235-240.
 23. Warczek J.: *„Sygnały drganiowe jako miary diagnostyczne amortyzatorów samochodów osobowych”* Rozprawa doktorska Politechnika Śląska Wydział Transportu, Katowice 2004r.
 24. Warczek J.: *Variation of energy vibrations in aspect of diagnostics parts of car suspensions*. Pernest Contact 2004. Pardubice, Slovakia Maszyn i Komitetu Transportu PAN.



Prof. nz. dr hab. inż.

Janusz GARDULSKI

jest pracownikiem naukowym Katedry Budowy Pojazdów Samochodowych Wydziału Transportu Politechniki Śląskiej. Zainteresowania badawcze: diagnostyka wibroakustyczna, dynamika zawiesz

pojazdów samochodowych, modelowanie nieliniowych obiektów mechanicznych, minimalizacja hałasu i drgań w obiektach technicznych. Jest autorem i współautorem 3 monografii, książek i skryptów, ok. licznych artykułów opublikowanych w czasopiśmie i materiałach konferencyjnych. Członek PTPE, PTDT, oraz różnych sekcji Komitetu Budowy Maszyn i Komitetu Transportu PAN.

PLD WE WSPOMAGANIU DEMODULACJI DRGAŃ MASZYN

Piotr KRZYWORZEKA, Witold CIOCH

Katedra Mechaniki i Wibroakustyki, AGH
Al. Mickiewicza 30, 30-059 Kraków
email: krzyworz@agh.edu.pl

Streszczenie

W diagnozowaniu maszyn, dla których realizacja celu podstawowego obejmuje powtarzanie sekwencji zdarzeń podobnych, można wykorzystać PLD dla wspomagania demodulacji sygnału drganiowego w niestacjonarnych stanach pracy maszyny. Zastosowana transformacja skali czasu powoduje wyrównanie długości dyskretnych reprezentacji kolejnych cykli i zbliża rezultat do sygnału stacjonarnego zachowując modulacje AM i PM. Nawet zastosowanie demodulatorów asynchronicznych prowadzi wówczas do wyraźnego zmniejszenia uchybu obciążenia rezultatu. zmniejszenia obciążeń wyników. Przedstawiono rezultaty badań dla sygnałów modelowych i rzeczywistych.

Słowa kluczowe: diagnostyka, modulacja, synchronizm.

PLD AIDED DEMODULATION OF MACHINE VIBRATIONS

Summary

This paper is concerned with the PLD (Procedure of Linear Decimation) aided output biasing elimination which takes into account some particularities of rotating machinery signal modulation. Rotating speed fluctuations or trend affect demodulation procedure due to nonlinear trend of carrier frequency. The proposed time scale transformation, denoted as PLD, aims to reduce demodulation error by employing linear approximation of carrier cycle short-term changes. Some examples of simulated and real signal asynchronous demodulations are also discussed hereby.

Keywords: diagnostics, modulation, synchronism.

1. DLACZEGO DEMODULACJA?

Rozważania modelowe, uogólniające wieloletnie doświadczenia diagnostyki wibroakustycznej implikują demodulację jako sposób odtworzenia zmienności właściwej wczesnym stadium ewolucji wielu kategorii uszkodzeń w prostszej formie i przy mniejszych zakłóceniach, niż ujawnia to bezpośrednio drganiowy sygnał mierzony y .

Mimo pewnych analogii MSM modulacji zamierzonych amplitudy AM, kąta PM, FM, czy położenia impulsów PPM i użytecznych kategorii modulacji sygnału maszyny MSM, utrudnienia powodują odmienności tych ostatnich. Wymienić tu warto:

- brak wyraźnego rozgraniczenia między sygnałem nośnym y_c (nieznanym); a zmodulowanym (mierzonym) y ;
- bezpośrednią niedostępność sygnału nośnego;
- nieznaną modulujący sygnał wartościowy x ;
- niepełną (lub brak) informację o modulatorze;

- niestacjonarność procesu MSM; zwłaszcza krótkoterminowe trendy i nieinformacyjne fluktuacje częstotliwości nośnej.

Wybór właściwej procedury demodulacji PD utrudnia ich brak. Stosowane zostały opracowane bez uwzględnienia specyfiki MSM.

Stąd przydatność konkretnych PD w diagnozowaniu nie bywa bezwarunkowa, a ich skuteczność pozostawia wiele do życzenia.

Nie wnikając w szczegóły, w każdym przypadku rezultaty demodulacji mogą stanowić wiarygodną bazę oceny stanu technicznego jedynie wówczas, jeśli PD będzie estymatorem efektywnym, to jest nieobciążonym i o minimalnej wariancji. Wariancja oceny rezultatu i jej liniowa reprezentacja – uchyb statystyczny, nie stanowią tu tak istotnego problemu jak w innych zastosowaniach pomiarowych, natomiast obciążenie oceny porównywalne z interwałami skali może prowadzić do mylnej diagnozy.

Jak się jednak okaże, dotychczas stosowane sposoby demodulacji drgań w czasie rzeczywistym nie kontrolują obciążenia rezultatu, bądź nie uwzględniają jego przyczyn, ponieważ są

optymalizowane dla określonych i stacjonarnych sygnałów nośnych.

2. PRZESŁANKI WSPOMAGANIA PLD

2.1 Procedura liniowej decymacji

Wspomaganie PLD dotyczy MSM maszyn dla których realizacja celu podstawowego wiąże się z cyklicznym powtarzaniem sekwencji zdarzeń podobnych, lecz nie identycznych.

Obejmuje wybrane kategorie demodulacji czasu rzeczywistego (tu PM i AM) w niestacjonarnych stanach eksploatacyjnych

Pomijając szczegóły (wielokrotnie opisywane), PLD prowadzi do opisu zmienności sygnału w innej skali czasu, takie że:

- opis zmienności staje się prostszy,
- zachowane są istotne cechy zmienności informacyjnej,
- zmienność nieinformacyjna ulega redukcji.

Taka transformacja pozwala w pewnych przedziałach czasowych i dla pewnych kategorii niestacjonarności zredukować niestacjonarność sygnału do stopnia umożliwiającego stosowanie prostszych metod analizy (np. widmowej) oraz procedur demodulacji właściwych sygnałom stacjonarnym. Polega ona na synchronizacji nowego zegara właściwie dobranym ciągiem zdarzeń referencyjnych. Nie są one tożsame z cyklem chwilowym (to byłby przypadek analizy rzędów), lecz pewną jego charakterystyką synchronizującą $\psi(\Theta)$.

Po wielu próbach dokonano zadawalającego wyboru tej charakterystyki aproksymując w oknie analizy T_k monotoniczny trend referencyjnego cyklu chwilowego liniowym przyrostem $\psi(\Theta)$ co pozwoliło na opracowanie przybliżonej metody synchronizacji dyskretnych reprezentacji sygnałów drgań ciągiem próbek $\psi(\Theta)$, uzyskanej bądź na drodze pomiaru kąta obrotu, bądź drogą autosynchronizacji.

Poprzedzając demodulację PLD:

- redukuje nieinformacyjny trend cyklu (monotoniczną zmianę Θ),
- zachowuje w stopniu na ogół wystarczającym informacyjne fluktuacje cyklu chwilowego Θ reprezentujące modulacje położenia impulsu PPM oraz kąta PM,
- zachowuje także (co oczywiste), modulację amplitudy AM.

2.2. Kategorie modulacji a celowość wspomagania

Potrzeba, rodzaj i skuteczność wspomagania demodulacji niekoniecznie bywają *a priori* oczywiste. Szereg pytań wymaga najpierw odpowiedzi – a przede wszystkim to:

- a) jaki rodzaj modulacji dominuje w badanym sygnale;
- b) co chcemy uzyskać – jaka charakterystyka ξ rezultatu PD jest użyteczna w diagnozowaniu, także w sensie możliwości stworzenia skali pomiarowej wybranej własności stanu;
- c) jakich kategorii niestacjonarności wspomaganie może i ma dotyczyć;
- d) jak wspomaganie wpływa na niepewność oceny ξ ;
- e) obecność i rodzaj związków sygnałów wejściowych modulatora (x, y_c).

Obecność (b) jest cechą charakterystyczną wyróżniającą MN, bowiem:

W modulacjach zamierzonych (np. telekomunikacyjnych), x i y_c są nieskorelowane. Sygnał nośny nie ma wpływu na informację wartościową, a jedynie na jej transfer – inaczej mówiąc sygnały modulujący i nośny nie mają wspólnej przyczyny

W modulacjach niezamierzonych natomiast, jeśli ten sam ruch jest przyczyną zmian modulujących, jak też tych, które je przenoszą (sygnału nośnego), często występuje synchronizacja kodowania IW cyklem kinematycznym.

Wówczas procesy drganiowe $\{x\}$ i $\{y_c\}$ związane z zamierzonymi własnościami dynamicznymi i eksploatacyjnymi, taktowane są tym samym zegarem, którego zbiór zdarzeń referencyjnych wyznacza roboczy cykl kinematyczny (por. rys. 1).

Reasumując, Im bliżej badanej MSM do zamierzonego analogu, tym mniej użyteczna może okazać się PLD (co nie wyklucza innych metod wspomagania – np. PLL [6]).

2.3. Ocena jakości wspomagania

Postulat użyteczność rezultatu w diagnozowaniu stawia procedurom demodulacji wymagania nieco inne niż system telekomunikacyjny.

Dokładne odtwarzanie kształtu x nie jest generalnie konieczne, nie bywa on zresztą jednoznacznie zdeterminowany, gdzie losowa natura manifestuje się zwłaszcza w przedziale obserwacji jednostkowej T .

Dla tworzenie skali diagnostycznej kluczowym jest znalezienie wiarygodnych i odtwarzalnych wzorców opartych o uśrednione charakterystyki ξ realizacji sygnału x . ξ może reprezentować np.:

- średnią energię w przedziale pomiaru (RMS);
- ocenę kształtu sygnału modulującego (estymaty bezwymiarowe, *skewness*, *kurtosis*);
- informacyjny i poprawnie odtworzony fragment widma sygnału x .

Stworzenie syntetycznego wskaźnika oceniającego ilościowo różne cechy odtworzonego sygnału modulującego utrudnia przede wszystkim ich niepełna znajomość w praktyce.

Przyjmując za podstawowe postulat poprawnego odtwarzania wartości skutecznej modulacji określonej kategorii (AM, lub PM) można zaproponować wskaźnik porównawczy interpretowany dalej jako uchyb resztkowej niestacjonarności δ_{RNS} – względną różnicę wartości skutecznej ocenianej i referencyjnej.

Formalizacja wiąże się z rodzajem modulacji i wyborem ξ .

2.4 Wpływ wyboru $\psi(\Theta)$

Oprócz oczywistego znaczenia skali czasu dynamicznego, wartości zainteresowania okazują się dwa inne przypadki – ich możliwości zestawia rys. 1.

PLD wprowadza czas η synchronizowany pewną charakterystyką cyklu $\psi(\Theta)$. Jej wybór okazuje się krytyczny zwłaszcza w przypadku modulacji PM.

Gdy $\psi(\Theta)$ odpowiada cyklowi chwilowemu Θ , opis sygnału w dziedzinie częstotliwości przedstawia widmo rzędów wykorzystywane również w diagnozowaniu [8] (rys. 1 b). Widma rzędów cechuje dobra rozdzielczość i dynamika. Ich estymacja natomiast wymaga złożonych obliczeń i precyzyjnego pomiaru fazy (położenia) [ibidem]. Ponadto w tak zdefiniowanej skali 'η' tracona jest informacja o fluktuacjach cyklu, które mogą być wynikiem modulacji PPM drgań przez dynamiczne procesy resztkowe, cenne źródło wczesnej informacji diagnostycznej [7].

Natomiast, w stanach niestacjonarnych demodulacja składowych PPM mogłaby być wspomagana przez transformację skali czasu zachowującą wartościową składową fluktuacji częstotliwości chwilowej, a redukującą monotoniczny trend $\Theta_C(\eta)$.

Niedogodności powyższej pozwala w wielu przypadkach uniknąć liniowa aproksymacja monotonicznych zmian cyklu nośnego. Wówczas w przedziale T:

$$\psi(\Theta) = \Theta_{ref} \left(1 + \frac{\Delta_{\Theta}}{\Theta_{ref}} t \right) \quad (1)$$

3. WSPOMAGANIE DEMODULACJI AMPLITUDY I KĄTA

3.1 Sygnał analityczny w opisie modulacji

Dogodne narzędzie formalne opisu stanów ustalonych procedur demodulacji stanowi analityczna reprezentacja sygnału oparta o transformatę Hilberta [4].

Jej postać wykładnicza:

$$Y(t) = A(t)e^{j\Phi(t)} \quad (2)$$

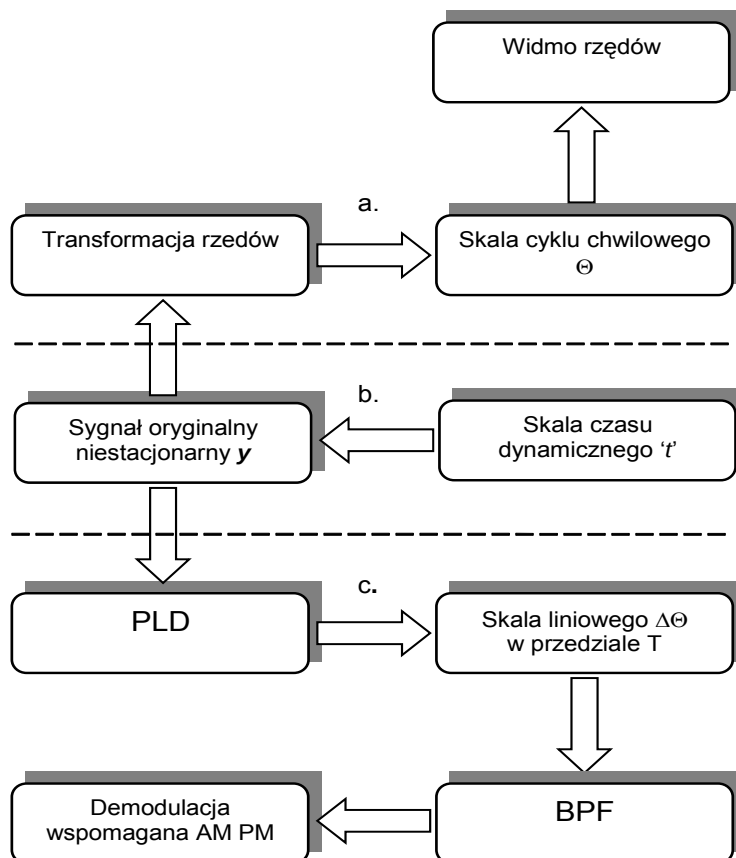
pozwała na rozróżnienie amplitudowej $A(t)$ i fazowej $\Phi(t)$ składowej zmienności y w dziedzinie czasu, lecz jedynie dla sygnałów wąskopasmowych $\Delta f/f_c \ll 1$ [4].

Formuła (2) często stosowana bywa bezpośrednio dla demodulacji amplitudy oraz (po dodatkowych zabiegach, także domniemanych modulacji niezamierzonych kąta [4-6]. Metody te, określane mianem niekoherentnych, lub asynchronicznych AD [35] nie wymagają dostępności, ani odtwarzania określonej formy sygnału nośnego y_C . Jedynie pewne jego parametry, jak na przykład f_C , definiują parametry niezbędnej wstępnej filtracji pasmowej. Czy stanowią zadawalające rozwiązanie przy niestacjonarnym sygnale nośnym – o tym dalej?

3.2. Wspomaganie PLD w demodulacji amplitudy

Jeśli przyjąć, że AM sygnału maszynowego jest modulacją wąskopasmową o paśmie $2F_X$, (F_X pasmo sygnału modulującego), to metoda obwiedni – $Abs[Y]$ (2) wymaga najmniej ograniczeń, i daje dobre rezultaty dla sygnałów o symetrycznej obwiedni (a po pewnej modyfikacji, także niesymetrycznej). Nie wymaga odtworzenia częstotliwości nosnej a jedynie lokalizacji pasma modulacji wokół domniemanej f_C . Prosty filtr pasmowy okazuje się wystarczający w stanach stacjonarnych. Zmiany częstotliwości nośnej wymagają dostosowania częstotliwości centralnej filtru, bowiem metoda $Abs[Y]$ nie jest selektywna, a ponadto wykazuje znaczną wrażliwość na interferencje innych modulacji bliskich [6].

Wspomagająca rola PLD ogranicza się do transformacji sygnału zmodulowanego y .



Rys. 1. Możliwości reprezentacji sygnału niestacjonarnego w różnych skalach czasu

- a) czas dynamiczny, charakterystyki dwuparametrowe;
- b) czas cyklu chwilowego, widma rzędów;
- c) czas aproksymacji liniowej $\psi_T(\Theta)$, wspomaganie demodulacji

3.3. Wspomaganie PLD w demodulacji kąta

W przypadkach niestacjonarności eksploatacyjnej maszyny, kiedy prędkość obrotowa, a zatem i częstotliwość nośna nie są stałe, źródeł błędów obciążenia rezultatu DP należy szukać w odtwarzaniu niewłaściwych składowych fazy chwilowej. Można wyróżnić cztery składowe o różnej jakościowo zmienności i odmiennej, choć nie zawsze jednoznacznej interpretacji[6]:

$$\text{Arg}[Y(jt)] = \Phi(t) = \Phi_C(t) + \Phi_x(t) + \Phi_F(t) + \Phi_T(t) \quad (3)$$

gdzie:

1. $\Phi_C(t)$ – faza sygnału nośnego przy ustalonym cyklu Θ_C ;
2. $\Phi_T(t)$ – rezultat monotonicznej zmiany cyklu nośnego;
3. $\Phi_x(t)$ – składowa informacyjna, tu otworzenie zmienności DPR;
4. $\Phi_F(t)$ – nieinformacyjne fluktuacje związane ze stanem eksploatacyjnym, lub niedoskonałością demodulacji.

Pasma częstotliwości odpowiadające pochodnym składowych 1,...,4 powinny być rozłączne, ale nie zawsze są

W przyporządkowanie różnych kategorii zmienności odmiennym własnościom realnego obiektu diagnostyki [6] pozwala w ogólnym przypadku przyjąć, że:

- a) fluktuacje cyklu chwilowego reprezentują zmienność informacyjną;
- b) krótkoterminowy monotoniczny trend cyklu nośnego reprezentuje zakłócającą zmianę kategorii stanu eksploatacyjnego.

Demodulacja jest w tej sytuacji formą selekcji zmienności (a) przy zakłócającym, lecz potencjalnie wspomagającym działaniu (b).

Właściwie dobrana charakterystyka $\psi(\Theta)$ redukuje Φ_T (PLD).

Operacja *detrend* eliminuje Φ_C :

$$x \Rightarrow \Phi_x = \text{detrend}\{\text{Arg}[Y(t)]\} \quad \text{dla skali } 't' \quad (4)$$

$$x_u \Rightarrow \Phi_{xu} = \text{detrend}\{\text{Arg}[U(\eta)]\} \quad \text{dla skali } '\eta'.$$

4. EKSPERYMENTY SYMULACYJNE

4.1. Założenia

Weryfikacja asynchronicznych metod demodulacji AM i PM, wspomaganych, lub nie, była przedmiotem eksperymentu symulacyjnego, którego etapy zestawia rys. 6.6.

Porównanie wyników dotyczyło RMS odtworzonych sygnałów x oraz ich widm RMS w zakresie domniemanej użyteczności dla diagnozowania.

Dla przydatności wspomagania metody obwiedni przez PLD badano także obecność i rodzaj związków sygnałów wejściowych modulatora (x , y_c), a konkretnie, czy jest synchronizowany cyklem nośnym.

W przypadku PM symulacje tu prezentowane miały przede wszystkim na celu ocenę skuteczności PLD w przybliżaniu cyklu nośnego i w konsekwencji redukcji fluktuacji resztkowych.

4.2. Sygnały modelowe, wzorcowanie DP

Z powodu niedostępności sygnału modulującego MSM ocena skuteczności demodulacji może być zrealizowana jedynie poprzez wzorcowanie PD dwoma sygnałami modelowymi, podobnie zmodulowanymi, lecz o różnych składowych nośnych, Są to:

- sygnał referencyjny o $f_c = \text{const}$;
- sygnał o niestacjonarnej składowej nośnej – $f_c = \text{var}$.

Obie kategorie obejmowały:

- jednoczesne modulacje AM i PM tego samego sinusoidalnego sygnału nośnego;
- modulacje sygnałem jedno lub dwutonowym, albo falą prostokątną;
- sinusoidalne składowe niskoczęstotliwościowe;
- szum addytywny.

Testy dotyczyły liniowego trendu nośnej f_c , aproksymowanego przez PLD liniową zmianą cyklu.

Jednocześnie demodulowano sygnał referencyjny o $f_c = \text{const}$.

Ponadto zbadano skuteczność PLD dla: sygnału modulującego AM:

- o stałej częstotliwości;
- o trendzie częstotliwości podobnym trendowi f_c ,

oraz parametry charakterystyczne PLD, decydujące o jej skuteczności.

Jak wspomniano przy omawianiu założeń PLD, po transformacji $t \Rightarrow \eta$ $f \Rightarrow f'$ otrzymujemy zapis sygnału próbkowany jednolicie,

$h = \text{const}$ w nowej skali czasu. Wartość h wybierano arbitralnie, tak jednak by zachować sens fizyczny charakterystyk ξ w odniesieniu do badanego obiektu w ruchu – tj. przede wszystkim zidentyfikować składowe kinematyczne widma w rzeczywistym zakresie ich częstotliwości jakie występują w trakcie eksploatacji.

W przeprowadzonych symulacjach, dla prezentacji widm $U(f)$ przyjęto jako referencyjną:

$f_a = (f_0 + f_1)/2$ – tj. średnią częstotliwość nośną w przedziale T ;

Ponadto rejestrowano istotne dla jakości, a nawet wykonalności PLD wskaźniki:

- $\delta f_c = 100(f_0 - f_1)/f_a$ – procentowy przyrost częstotliwości nośnej w przedziale T ;
- $\delta \Theta$ – procentowy przyrost cyklu aproksymującego (na cykl).

4.3. Ocena rezultatu DP

Jako reprezentacja błędu odtworzenia x posłużył w pierwszym przybliżeniu uchyb średniokwadratowy, ściślej jego liniowa reprezentacja – wartość skuteczna RMS.

Nie jest to ocena selektywna. Jednak stosowane w Europie skale diagnostyczne często klasyfikują wg. wartości poziomu RMS drgań, a kolejne wartości progowe są dość odległe (> 6 dB), łączna ocena niepewności i obciążenia okazuje się wystarczająca. Bardziej wyrafinowane metody oceny jakości demodulacji znajdzie dociekliwy czytelnik w [6].

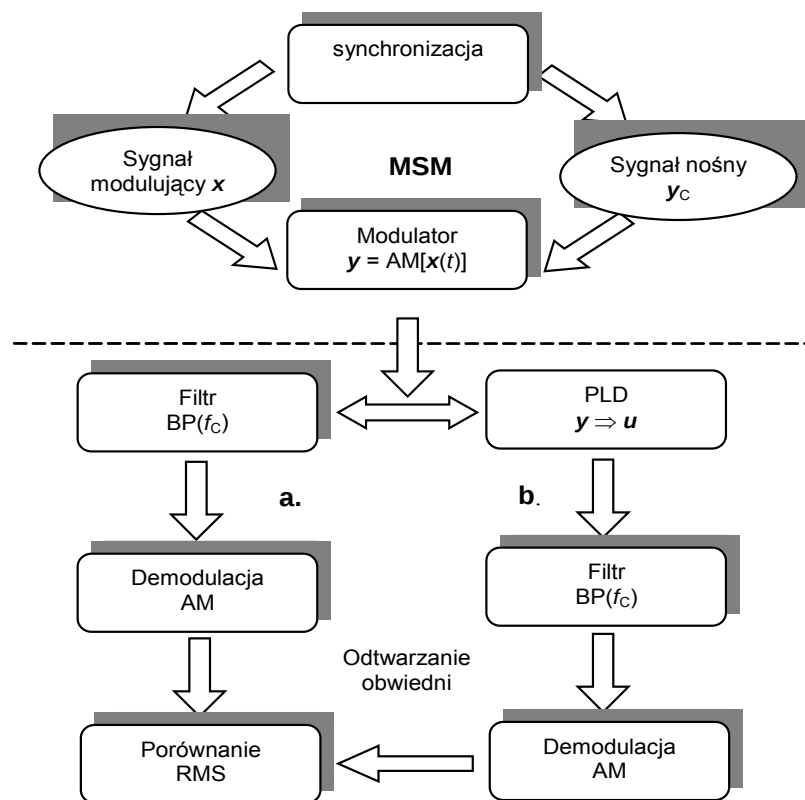
Zamieszczone w tabeli 1 uchyby względne pozwalają ocenić

- czy stopień wspomagania PLD jest akceptowalny?
- czy filtracja fluktuacji resztkowych jest skuteczna?

4.4. Wybrane rezultaty i wnioski

Przykłady ilustrują przypadki charakterystyczne dla możliwych rzeczywistych AM i PM sygnałów, a także możliwości PLD redukcji skutków trendu f_c .

Autorzy pragną usprawiedliwić brak rzeczywistej skali rzędnych wyników pomiaru drgań Nie stanowiąc tu informacji wartościowej byłyby one czystą redundancją w toku wywodów generalnie dotyczących wykrywania zmian innowacyjnych, lub zakłócających, ale zawsze względem referencji w podobnej skali.



Rys. 2. Alternatywne tory demodulacji AM:

- a) bez wspomagania;
b) wspomaganie PLD

- $\varepsilon_1 = (-X_0 + X_R) / X_0 * 100\%$
- $\varepsilon_2 = (-X_0 + X_{UF}) / X_0 * 100\%$
- $\varepsilon_3 = (-X_0 + X_U) / X_0 * 100\%$
- $\varepsilon_4 = (-X_0 + X_D) / X_0 * 100\%$
- $\varepsilon_5 = (-X_R + X_{UF}) / X_0 * 100\%$
- $\varepsilon_6 = (-X_R + X_D) / X_0 * 100\%$

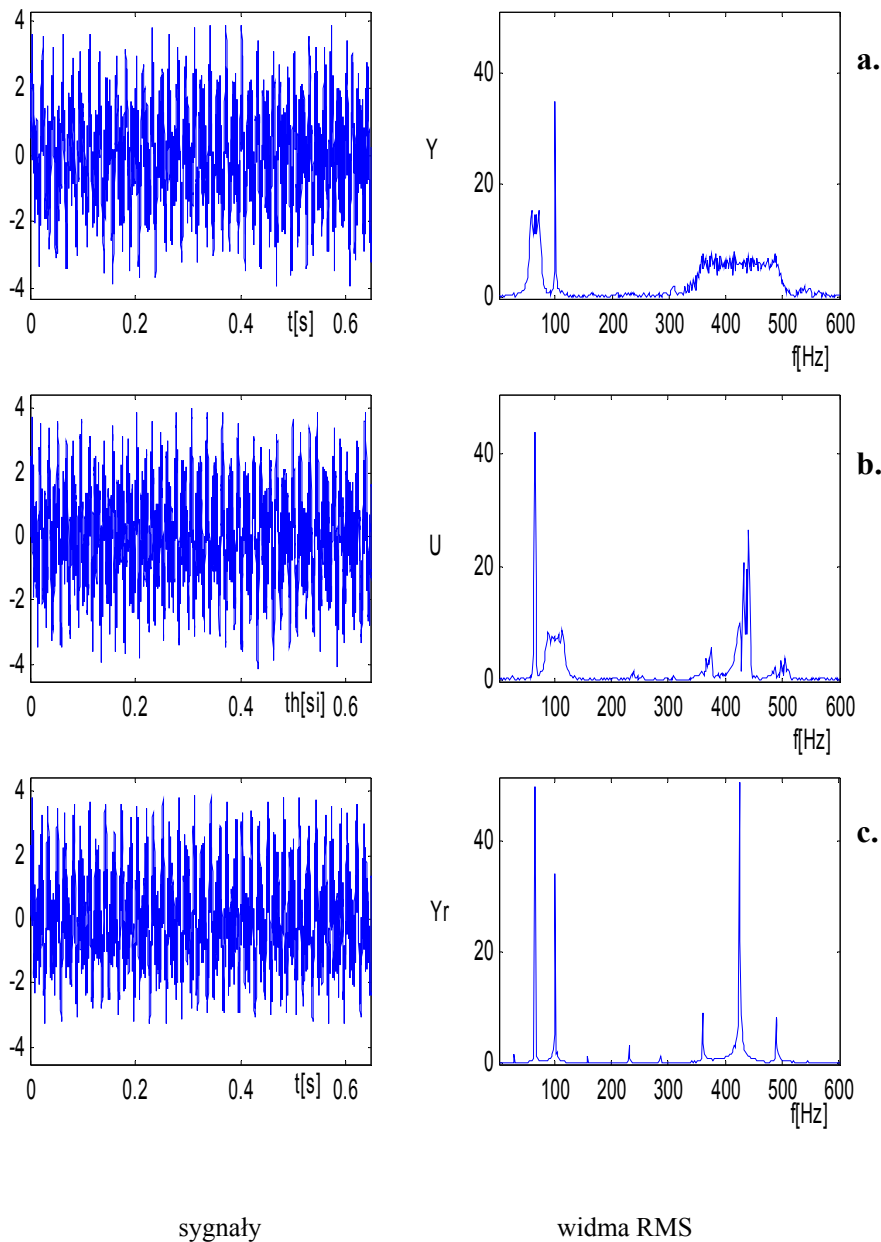
- X_0 – sygnału modulującego x_0 ;
- X_D – rezultat PD bez PLD
- X_U – rezultat PD, PLD;
- X_{UF} – rezultat PD, PLD, HP
- X_R – rezultat PD sygnału referencyjnego y_R ,
- $f_c = \text{const}$

Wartości $\delta f_c\%$, $\delta \Theta\%$ (Tab. 1) określają w przybliżeniu górną granicę przedziału skuteczności PLD w przeprowadzonych eksperymentach symulacyjnych.

Dla MSM wobec niedostępności realnych sygnałów modulujących przyjęto jako referencyjne rezultaty PD wspomaganej – X_U dla AM, X_{UF} dla PM

Tabela 1.

no	Typ Mod.	uchyby RMS %						dane PLD		uwagi
		ε_1	ε_2	ε_3	ε_4	ε_5	ε_6	$\delta f_c\%$	$\delta \Theta\%$	
1	AM	-25,4	-15,1	-14,8	-20,1	10,6	5,4	35,8	-3,02	$x - \text{sqr AM}$ $f_x = \text{const}$
2	PM	-12,3	-12,9	786	8537	-0,6	8550			
3	AM	-18,8	-15,1	-14,6	-19,3	4,2	-0,56			$x - \text{sqr AM}$ $f_x = \text{var}$
4	PM	-19,2	-13,8	784	8546	5,5	8565			
5	AM	-20,7	-9,52	-9,2	-17,4	11,6	3,35			$x - \text{sin AM}$ $f_x = \text{const}$
6	PM	-11	-3	1148	9980	7	9990			



Rys. 3. Symulacja DAM sygnału modulowanego AM i PM falą prostokątną, $f_M = \text{const}$

- a) sygnał oryginalny;
- b) sygnał po transformacji PLD;
- c) sygnału referencyjny, $f_C = \text{const}$

4.5. Uwagi

Jeśli niestacjonarność sygnału nośnego potraktować jako zakłócenie obciążające rezultat DMS, to w określonych przypadkach wspomaganie może znacznie zredukować obciążenie estymat RMS odtwarzanego x : Wskazują to wartości ε_2 i ε_5 (Tab. 1).

Wspomaganie jest tym bardziej skuteczne im lepiej PLD odtwarza, (a właściwie przywraca) skalę czasu dynamicznego. Odtworzenie dokładne,

w praktyce na ogół niepotrzebne, sygnału nośnego o częstotliwości oryginału, pozwala odzyskać oryginalny kształt sygnału modulującego w dziedzinie czasu η (por rys 6.13c)

Aproksymacja liniowo-odcinkowa cyklu nośnego w przedziałach o monotonicznej zmianie częstotliwości nośnej rzędu nawet 30–40 % pozwala na odtworzenie wartości skutecznej modulacji (RMS x) z dokładnością rzędu 10% wystarczającą w zastosowaniach diagnostycznych.

Ponieważ PLD przybliża synchronizm sygnału u z cyklem nośnym, asynchroniczne z nim śladowe sygnały y ulegają osłabieniu a odpowiadające im prążki widma U , rozmyciu, jeśli zatem asynchronizm dotyczy sygnałów modulujących amplitudę wspomaganie okazuje się niecelowe (por. rys. 4.b).

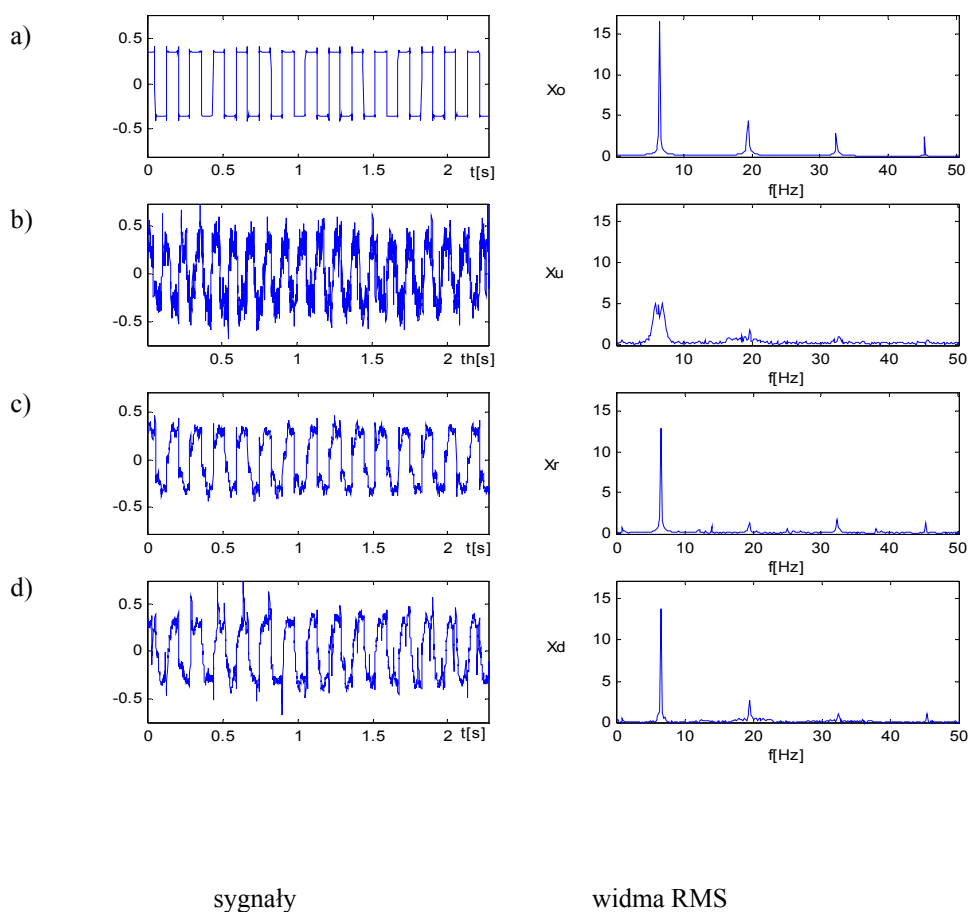
Błąd odtwarzania Θ_C powoduje fluktuacje resztkowe RF oceny x (por. np. rys. 3.c) i w rezultacie nieakceptowane obciążenie oceny wartości skutecznej rezultatu demodulacji PM.

Właściwy dobór parametrów PLD pozwala obniżyć częstotliwość FR poniżej F_X , co ułatwia ich zadowalające odfiltrowanie (por. rys. 4.b i 6. c).

5. DEMODULACJA DRGAŃ MASZINY

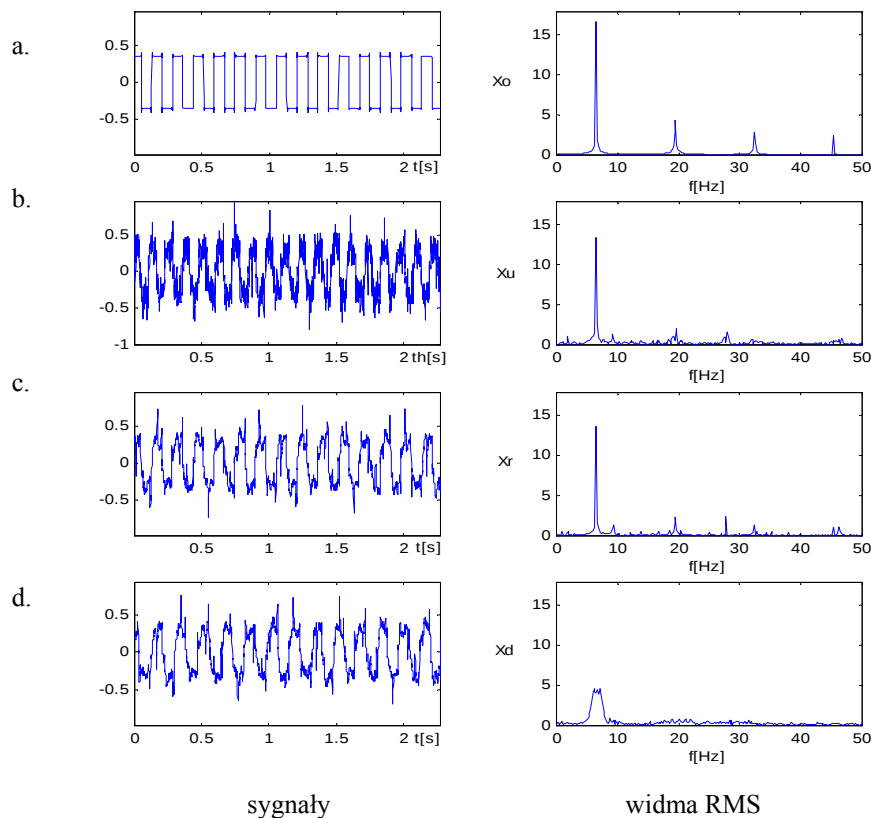
Tu sytuacja nie jest tak jednoznaczna, jak w przypadku symulacji. By PD miała sens trzeba najpierw odpowiedzieć na pytania:

- czy występuje modulacja i jaka?
- jaka cecha zmienności stanowi jej reprezentację w sygnałach, lub charakterystykach wyjściowych?
- czy częstotliwość nośna, ew. pasmo modulacji zlokalizowano trafnie?
- czy wstępna filtracja pasma domniemanej PM jest konieczna, użyteczna, zbędna?
- co przyjąć jako wynik referencyjny?



Rys. 4. Symulacja DAM sygnału modulowanego AM i PM falą prostokątną, $f_X = \text{const}$

- a. sygnał modulujący
- b. DAM, wspomaganie PLD;
- c. DAM sygnału referencyjnego, $f_C = \text{const}$;
- d. DAM, bez wspomagania



Rys. 5. Symulacja DAM sygnału modulowanego AM i PM falą prostokątną $\delta_x/\delta_c = 1$:

- a) sygnał modulujący;
- b) DAM, wspomaganie PLD;
- c) DAM sygnału referencyjnego, $f_c = \text{const}$;
- d) DAM, bez wspomagania

Ponieważ prezentacja rezultatów (rys. 7, 8 i 9) służy jedynie porównaniu efektów wspomaganie PLD z jego brakiem, pominięto skalowanie osi rzędnych w jednostkach bezwzględnych zachowując jedynie podobieństwo skali. Także wybór obiektu podyktowała dostępność i łatwość eksperymentowania. Selekcja pasma MSM nie sprawiała trudności (por. rys. 7). Wyniki dotyczą małego fragmentu rozbiegu (12 obrotów wałka referencyjnego) co odpowiada względnej zmianie cyklu referencyjnego ok. $-4,8\%$ (na cykl $\delta\Theta = 0,4\%$).

Mimo tak małej wartości $\delta\Theta$ uchyb obciążenia demodulacji PM okazuje się nieakceptowalny (rys.9). $\epsilon_2 = -426\%$. Natomiast dodatkowa filtracja HP fluktuacji resztkowych zmniejsza obciążenie o $\epsilon_1 = -16,4\%$.

W przypadku AM PLD poprawia nieco czytelność widma, natomiast zmiana RMS rezultatu jest pomijalna – $\epsilon_a = 5,18\%$ (rys. 8 a,b).

6. WNIOSKI

Pomimo prostoty samej procedury okazuje się PLD dość skutecznym narzędziem wspomagającym demodulację kąta, zwłaszcza, jeśli ten sam ruch jest cykliczną przyczyną zmian modulujących, jak też tych, które je przenoszą.

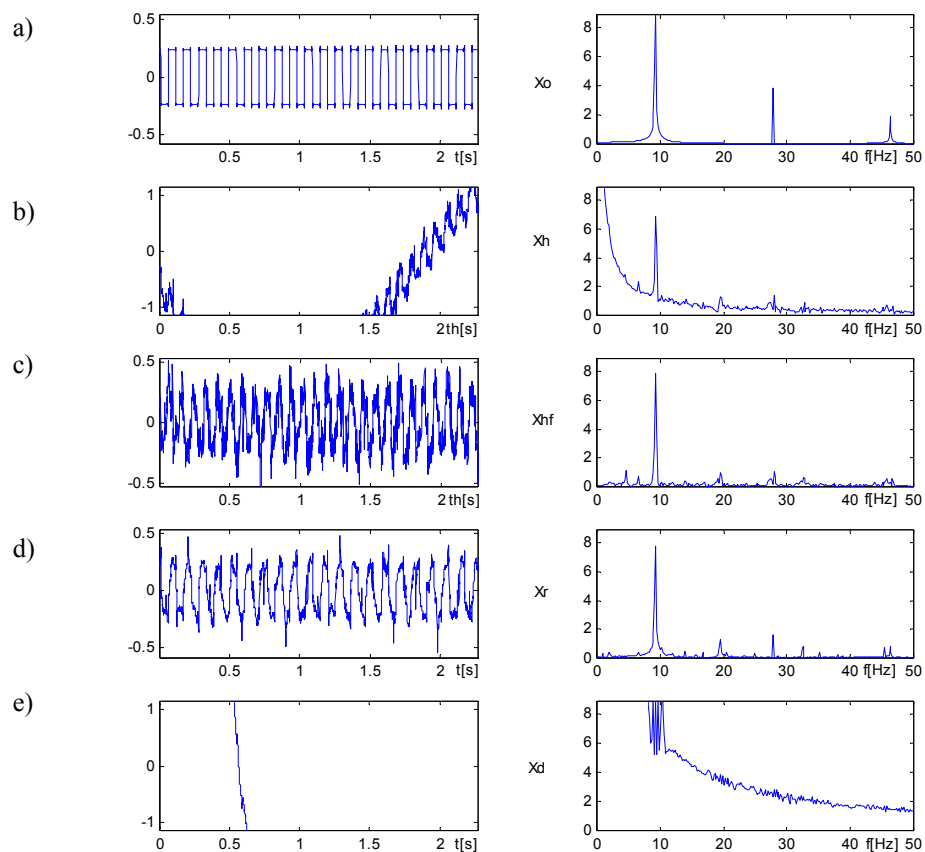
Jednakże, w przypadku PM, wspomaganie demodulacji przez PLD nie jest rozwiązaniem optymalnym, lecz kompromisowym.

asynchronizm kodowania sygnałów modulujących i nośnych AM powoduje nieskuteczność PLD. W pozostałych badanych przypadkach wspomaganie PLD pozwala uzyskać odtwarzanie RMS modulacji z błędem poniżej 10%.

DPM wymaga jeszcze na ogół filtracji fluktuacji resztkowych.

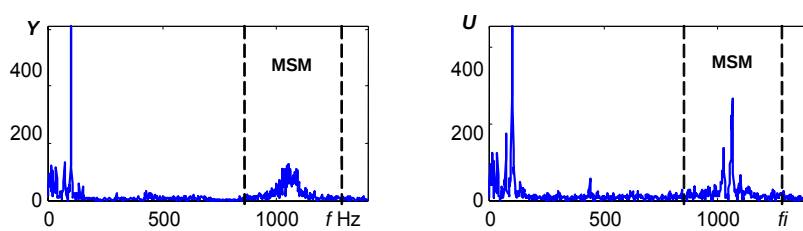
PLD ułatwia selekcję pasma modulacji.

Demodulatory wspomagane oparte o sygnał analityczny wymagają dużego nadpróbkowania (<100) przy rejestracji sygnału MS.



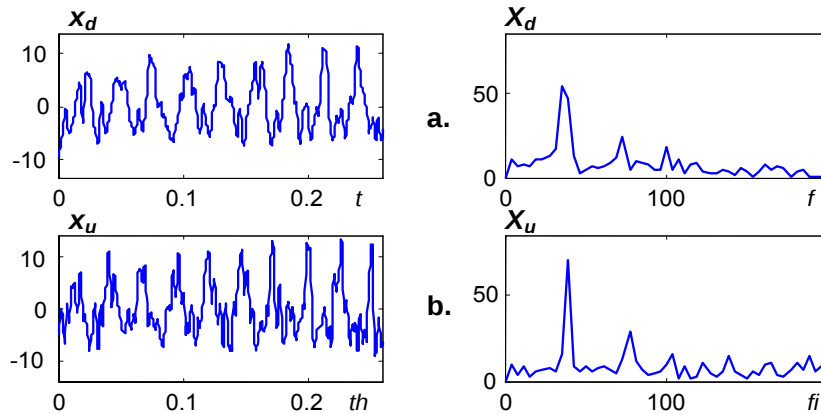
Rys. 6. Symulacja DPM sygnału modulowanego AM i PM falą prostokątną $\delta_M/\delta_C = 1$:

- a) sygnał modulujący;
- b) DPM, wspomaganie PLD;
- c) DPM, wspomaganie PLD; filtracja pasmowa;
- d) DPM sygnału referencyjnego, $f_C = \text{const}$;
- e) DPM, bez wspomagania



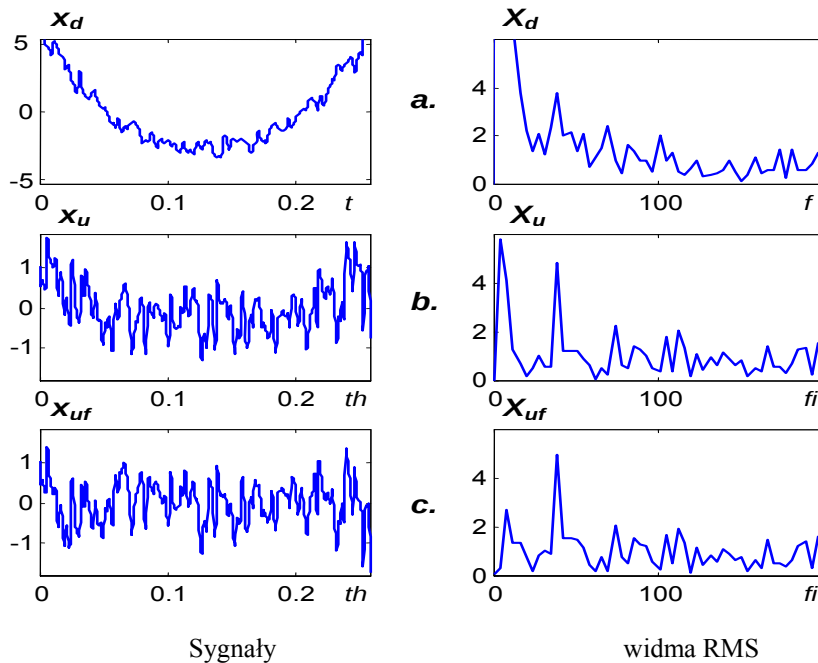
Rys. 7. Strefy domniemanej modulacji w widmie drgań przekładni:

- a) sygnał oryginalny;
- b) po transformacji PLD



Rys. 8. Wpływ wspomagania PLD na demodulację AM prędkości drgań przekładni $\delta_M/\delta_C = 1$:

- a) bez wspomagania;
 b) wspomaganie PLD



Sygnaly

widma RMS

Rys. 9. Wpływ wspomagania PLD na demodulację PM prędkości drgań przekładni:

- a) bez wspomagania;
 b) wspomaganie PLD;
 c) wspomaganie PLD; filtracja HP

BIBLIOGRAFIA

- [1]. Bigret R., Feron J. L.: *Diagnostic – maintenance disponibilite des machines tournantes*. Paris, Masson, 1995.
- [2]. Cioch W., Krzyworzecka P.: *Modulacje niezamierzone w diagnozowaniu głośnika*. Diagnostyka, vol. 28, 2003, 29 – 36.
- [3]. KRZYWORZEKA P., CIOCH W. *Machine diagnostics in cycle-time scale using linear decimation procedure*: 1st Int. Conf. on Experiments/Process/System.Modelling/Simulation/Optimization : 6–9 July, 2005 : book of abstracts / & CD-ROM Proc. — [1–6]. Univ. of Patras. LFME Athens, Greece.
- [4]. Haykin S: *Systemy telekomunikacyjne*. W-wa, WKiŁ, 1998.
- [5]. Krzyworzecka P.: *Demodulacja drgań maszyn przy zmianach prędkości obrotowej*. Diagnostyka, vol. 31, 2004, 21 – 28.
- [6]. Krzyworzecka P.: *Wspomaganie synchroniczne w diagnozowaniu maszyn*. Kraków-Radom. BPE, Wyd. ITE. 2004.

- [7]. Radkowski St.: *Wibroakustyczna diagnostyka uszkodzeń niskoenergetycznych*. W-wa, Radom, WITE, 2002.
- [8]. Piotrowski A., Stankiewicz A., Balunowski J., Solbut A.: *Diagnostics symptoms forming in the rotating machine monitoring based on the order tracking analysis*. II International Congress of Technical Diagnostics, Diagnostika'2000, Poland, Warsaw 19-22 September 2000.

Publikacja powstała w ramach projektu KBN
nr 4 T07B 05426

Ważniejsze symbole i skróty

AM	–	modulacja amplitudy
IW	–	informacja wartościowa
MSM	–	modulacja sygnału maszyny
SW	–	sygnał wartościowy
SN	–	sygnał nośny
MS	–	sygnał modulowany
VSD	–	demodulacja sygnału drgań
PM	–	modulacja fazy
PDI	–	procedura identyfikacji diagnostycznej
PLD	–	procedura liniowej decymacji
PLL	–	pętla synchronizacji fazy (<i>Phase-locked loop</i>)
PPM	–	modulacja położenia impulsu
AM	–	modulacja amplitudy
MSM	–	modulacja sygnału maszyny
SW	–	sygnał wartościowy
SN	–	sygnał nośny
MS	–	sygnał modulowany
VSD	–	demodulacja sygnału drgań
$\psi(\Theta)$	–	synchronizująca charakterystyka cyklu
ε	–	błąd względny
Φ_C	–	faza nośna
Φ_X	–	faza wartościowa
Φ_F	–	faza fluktuacji
Φ_T	–	faza trendu
Ω_C	–	częstość nośna
Θ_C	–	cykl nośny
x	–	sygnał modulujący
y	–	zmodulowany sygnał drganiowy
u	–	zmodulowany sygnał drganiowy po transformacji PLD
y_C	–	sygnał nośny
s_F	–	sygnał charakterystyczny
t	–	czas dynamiczny
T	–	przedział obserwacji sygnału
η	–	czas cyklu
ϑ	–	czas eksploatacyjny
δ_Θ	–	względna zmiana cyklu

Dr hab. inż. Piotr KRZYWORZEKA studiował na Wydziale Elektrotechniki AGH. Pracuje na tejże uczelni do dziś, obecnie jako prof. W Katedrze Mechaniki i Wibroakustyki. Wykładał kilka lat w Algierii. Jest autorem ok. 100 publikacji, głównie o tematyce diagnostycznej, rzeczoznawcą SEP w zakresie elektroakustyki oraz członkiem PTDT od momentu jego powstania. Był wieloletnim sekretarzem Polskiej Sekcji *Audio Engineering Society*. Interesuje się także psychologią i filozofią. Jako środek transportu preferuje rower.

Dr inż. Witold CIOCH jest absolwentem Wydziału Budowy Maszyn i Lotnictwa PRz., specjalność napędy lotnicze oraz Wydziału Inżynierii Mechanicznej i Robotyki AGH, specjalność wibroakustyka. Obecnie pracuje jako adiunkt w Katedrze Mechaniki i Wibroakustyki AGH. Autor 34 publikacji z tematyki diagnostyki technicznej i wibroakustyki. Zajmuje się zagadnieniami cyfrowego przetwarzania sygnałów, sieciami neuronowymi i analizą ryzyka eksploatacyjnego.

ENERGETYCZNE CHARAKTERYSTYKI DEGRADACJI SYSTEMU

Henryk KAŻMIERCZAK

Przemysłowy Instytut Maszyn Rolniczych
61-963 POZNAŃ, ul. Starołęcka 31, tel. 061 8712279, e-mail: kazmhenr@pimr.poznan.pl

Streszczenie

System mechaniczny modelowany jest jako procesor energii, który przetwarza moc wejściową na moc użyteczną i moc degradacji jej elementów. Elementy obiektu mechanicznego ulegają zużyciu z różną intensywnością. Jest to odzwierciedlane w widmowych charakterystykach dynamicznych obiektu i charakterystykach mocy obciążeń dynamicznych powodujących degradację maszyny. Przedstawiono ideę energetycznej metody identyfikacji stanu degradacji strukturalnej maszyny. Metoda stosowana jest w badaniach rozplywu energii i zmian strukturalnych w obiektach mechanicznych.

Słowa kluczowe: obiekty mechaniczne, degradacja maszyny, metody identyfikacja procesu degradacji.

THE ENERGETICS CHARACTERISTICS OF MACHINE'S DEGRADATION

Summary

The mechanical system is modelled as an energy processor, which transforms the input power into effective power and destruction power of its elements. The elements of the mechanical object are subject to wear with different intensity. This is reflected in the dynamic characteristics of the object and the vibration load power characteristics causing the destruction of the machine. The paper presents the concept of an energy-based vibration identification method of the degradation of objects. The method is used in testing the energy propagation and the structural changes in mechanical objects.

Keywords: mechanical objects, machines degradation, identification method of the degradation.

1. CHARAKTERYSTYKA PROCESU DEGRADACJI OBIEKTU MECHANICZNEGO

Podzespoły obiektu mechanicznego ulegają zużyciu z różną intensywnością. O stanie obiektu decyduje najsłabszy element. Niepożądane oddziaływania dynamiczne (drżania) dezorganizują obiekt funkcyjnie i strukturalnie, stanowiąc o ewolucji jego stanu technicznego. Kumulacyjne skutki występowania obciążeń (przeciążeń), powodujących degradację węzłowych elementów konstrukcji stanowią o potrzebie identyfikacji rozkładu przestrzennego obciążeń i ich wzrostu w funkcji czasu ewolucji stanu systemu.

Procesy stanowiące o charakterystyce trwałościowej obiektu mechanicznego mają wymiar energetyczny. W badaniach energetycznych charakterystyk trwałości zmęczeniowej maszyn stosowana jest metoda podatności dynamicznej do budowy analitycznego modelu systemu i metoda eksperymentalnej analizy modalnej do identyfikacji charakterystyk dynamicznych maszyn oraz metoda rozkładu mocy obciążeń w wielowęzłowych – wielowęzłowych systemach mechanicznych [6].

Własności dynamiczne maszyny są pochodną przestrzennego rozkładu parametrów fizycznych maszyny oraz warunków brzegowych (sposobu

posadowienia maszyny na podłożu), a także sposobu agregowania jej ze źródłem napędu [11].

Od ogólnego modelu do szczegółów to podstawa idei modelowania systemów mechanicznych. Stworzenie dobrego modelu części, podzespołu może stanowić o możliwych jego dalszych zastosowaniach. Informacje z testów eksperymentalnych są konfrontowane z modelem wirtualnym. Istotną jest integracja między badaniami testowymi i wirtualnymi (warunki brzegowe). W modelowaniu analitycznym obciążeń systemu praktykowane jest podejście hybrydowe z użyciem danych wejściowych z eksperymentu. Analiza dynamiki systemu oparta jest na obciążeniach wyznaczonych z eksperymentu. Wymagana jest pełna integracja testu wirtualnego z testem eksperymentalnym [8].

Przy dużej liczbie różnorodnych czynników konstrukcyjnych, technologicznych i eksploatacyjnych (zmienność obciążeń) wpływających na stan obiektu, charakterystyki obciążeń uzyskane podczas badań eksploatacyjnych całej maszyny pozwalają wskazać elementy poddane intensywnym obciążeniom [4].

W silnikach, w układach przenoszenia napędu, oraz w maszynach roboczych do najbardziej obciążonych elementów należą łożyska, których uszkodzenie może prowadzić do katastrofy

obiektów, w których są one instalowane. Uszkodzenie może wystąpić w wyniku ewolucyjnego procesu destrukcji lub w wyniku chwilowego przeciążenia [9, 11]. Dotyczyć to może np. łożysk w obiektach latających lub w generatorach w elektrowniach, a także w maszynach roboczych, itp. Uszkodzenie układu mechanicznego następuje wówczas, gdy porcja energii (mocy) spowoduje zmianę jego własności strukturalnych, gdy przy ciągłej kumulacji skutków pracy sił destrukcji nastąpi przekroczenie wartości granicznych charakterystyk wytrzymałościowych, powodując niszczenie węzła konstrukcyjnego. Procesy te mają wymiar energetyczny [2, 4].

Efektywny system oceny stanu maszyny, dostarczający szczegółowych informacji o własnościach części maszyny oraz dążność do budowy urządzeń odpornych na szkodliwe oddziaływanie zewnętrzne, mogą w istotny sposób zmniejszyć przypadkowość w funkcjonowaniu maszyny i umożliwić stworzenie teorii niezawodności opartej na innej podstawie niż na rachunku prawdopodobieństwa. Przyjmuje się twierdzenie: jeśli uległy zmianie cechy procesów wyjściowych maszyny oznacza, że zostało to spowodowane zmianą warunków zewnętrznych (wymuszeń) lub zmianą stanu technicznego maszyny.

2. KRYTERIA WYZNACZANIA TRWAŁOŚCI OBIEKTU MECHANICZNEGO

Kryteria obliczania trwałości zmęczeniowej oparte na koncepcji płaszczyzny kryterialnej, umożliwiają szacowanie trwałości zmęczeniowej dla różnych stanów wieloosiowego zmęczenia oraz określania kierunku pęknięcia zmęczeniowego. Według tego kryterium pęknięcie nastąpi na płaszczyźnie najbardziej wyężonej. Elementy maszyn poddawane są często cyklicznym obciążeniom zginającym, skręcającym oraz często obu tym obciążeniom, a także obciążeniom statycznym (ciężar i obciążenie wstępne), które związane są z funkcjami spełnianymi przez maszynę. Obciążenia te powodują procesy zmęczeniowe materiałów konstrukcji w punktach kumulacji naprężeń.

Trwałość jest podstawowym parametrem jakości i bezpieczeństwa konstrukcji. Przyczyną ograniczonej trwałości maszyny jest występowanie nadmiernych naprężeń dynamicznych. Naprężenie dynamiczne w drgającym dowolnie, stacjonarnie elemencie jest funkcją prędkości drgań. Wynika stąd związek między trwałością materiałów a amplitudą prędkości drgań maszyn.

W wyniku procesu degradacji (zmiana miar parametrów mechanicznych) następuje: obniżenie sztywności, spadek trwałości, złuszczenie bieżni, pękanie zmęczeniowe.

Rozróżnia się dwa mechanizmy rozpraszania energii mechanicznej:

1. Tarcie w warstwach poślizgu w węzłach kinematycznych

– węzły zaprogramowane (konstrukcyjne)

– węzły awaryjne: pęknięcia oraz powierzchnie poślizgu przy odkształceniu plastycznym

2. Drgania

Naprężenia zmienne wywołane drganiami konstrukcji koncentrują się na krawędziach szczelin i wad materiałowych. Pękanie wynika na ogół z koncentracji energii odkształcenia sprężystego lub plastycznego na krawędziach wad materiałowych. Proces pękania to proces propagacji wad.

Należy podkreślić synergizm czynników obniżających trwałość konstrukcji [10]:

– korozyjne, zmęczeniowe, oddziaływanie środowiska (wilgotność), temperatura,

– niejednorodność struktury,

oraz synergizm technologii inżynierii powierzchni w zwiększaniu trwałości: umocnienia warstwy wierzchniej, obróbka cieplna i laserowa, unikanie karbu.

Poziom naprężen w materiale mieści się w zakresie liniowo-sprężystym, gdy naprężenie w materiale znajduje się w liniowej części jego charakterystyki naprężenie – odkształcenie i jest niższe od jego granicy sprężystości. Naprężenie, przy którym następuje początek płynięcia nosi nazwę granicy plastyczności. Plastycznością nazywana jest cecha materiałowa, gdy poza zakresem liniowym zachodzą odkształcenia niesprężyste. Przy obciążeniu materiału do zakresu plastycznego, zmiana ulega jego struktura wewnętrzna (zmiana własności). Charakterystyki różnych materiałów różnią się między sobą dla różnych prędkości odkształceń i różnych temperatur.

Materiały kruche charakteryzują się tym, że pęknięcie następuje przed wystąpieniem jakichkolwiek zauważalnej zmiany wydłużenia.

Naprężenia zmienne wywołane drganiami konstrukcji koncentrują się na krawędziach szczelin i wad materiałowych. To powoduje, że rozpraszanie energii w materiale jest również skoncentrowane w tych miejscach. Propagacja wad materiałowych na skutek zmiennych naprężeń daje w wyniku złom zmęczeniowy.

Badania trwałości zmęczeniowej maszyn obejmują znajomość:

1) obciążeń eksploatacyjnych (różnorodności charakterystyk obciążeń),

2) charakterystyk zmęczeniowych elementów,

3) hipotez sumowania uszkodzeń (np. liniowe).

Środowisko korozyjne obniża trwałość (wytrzymałość zmęczeniową). Rozróżnia się: zmęczenie objętościowe i pitting warstwy wierzchniej.

Resztkowa trwałość zmęczeniowa jest funkcją historii widma obciążeń i trwałości początkowej (projektowej). Degradacja obiektu jest wynikiem kumulowania się skutków procesów zużycia.

Graniczna trwałość zmęczeniowa jest funkcją rodzaju materiału i poziomu obciążeń.

Podstawowym problemem jest wyznaczenie poziomu trwałości resztkowej całej maszyny.

W warunkach wieloosiowych obciążeń losowych wymagane jest sformułowanie energetycznej charakterystyki stanu wytężenia zmęczeniowego konstrukcji. Parametrem opisującym uszkodzenie układu jest liniowa kombinacja składowych wytężenia w ustalonej płaszczyźnie węzłów konstrukcji.

3. WYZNACZANIE TRWAŁOŚCI ZMĘCZENIOWEJ MASZyny Z WYKORZYSTANIEM METODY WIDMOWEJ ANALIZY ROZKŁADU MOCY OBCIĄŻEŃ DYNAMICZNYCH

Na trwałość maszyny jako układu złożonego mają wpływ własności dynamiczne jej elementów, a także mechaniczne właściwości materiałów z których są wykonane. Niszczenie układu mechanicznego jest związane ze zmianą jego własności strukturalnych, a te zachodzą przy udziale energii, która jest różnicą między pracą włożoną w układ mechaniczny a pracą użyteczną przez niego wykonaną. Rozpraszanie energii jest powodowane procesami tarcia zewnętrznego i wewnętrznego oraz odkształceniami trwałymi zachodzącymi w materiale. Uwzględniając, że część energii dyssypowanej jest rozpraszana w postaci ciepła, można z dużą dokładnością określić energię powodującą odkształcenia trwałe poszczególnych elementów maszyny. Energia ta jest przyczyną zmęczeniowego zniszczenia materiału.

Kryteria uwzględniające różnice faz przebiegu składowych stanu odkształcenia i naprężenia klasyfikuje się na kilka różnych sposobów. Jednym z nich jest podział według rodzajów przyjętych parametrów opisujących proces zmęczenia. Wyróżnia się trzy grupy kryteriów: naprężeniowe, odkształceniowe, energetyczne.

Hipoteza wytrzymałościowa Hubera, sformułowana w 1904 roku zakłada, że o zniszczeniu materiału decyduje krytyczna wartość energii właściwej odkształcenia postaciowego [18]. Dla małych odkształceń postaciowych materiału energia odkształcenia postaciowego jest zbliżona do energii materiału liniowo sprężystego, a zatem właściwości fizyczne materiału będą również zbliżone w początkowej fazie odkształcania do właściwości fizycznych materiałów liniowo sprężystych, czym charakteryzuje się większość znanych materiałów o dużym znaczeniu praktycznym, wykorzystywanych jako tworzywa konstrukcyjne.

3.1. Analiza rozpręgu energii w maszynie

Przepływ energii drganiowej jest modelowany podobnie jak przepływ ciepła w strukturze. Równanie różniczkowe przewodnictwa cieplnego opisuje stan drganiowy. Metoda analizy przepływu energii (SEA) bierze pod uwagę przestrzenną zmianę

energii w indywidualnych podsystemach oraz przepływ energii między podsystemami [9]. Zamiast charakteryzowania każdego podsystemu poprzez pojedynczą wartość energii, analiza przepływu energii opisuje straty i przewodzenie energii drganiowej w każdym podsystemie. Przez rozwiązanie odpowiednich równań energii na rozkład gęstości energii w każdym podsystemie, uzyskuje się przestrzenną zmianę odpowiedzi drganiowej (w dziedzinie przyspieszenia, naprężenia, ciśnienia akustycznego, itp.).

Główną ideą analizy przepływu energii jest fakt, że stan dynamiczny systemu mechanicznego może być reprezentowany przez gromadzoną, dyssypowaną i przenoszoną energię. Analiza przepływu energii pozwala na predykcję uśrednionego (przestrzennie i w dziedzinie częstotliwości) drganiowego zachowania się struktury analogicznie jak metoda SEA. Metoda analizy przepływu energii wykorzystuje objętościowe podejście różniczkowe do wyprowadzenia równań na gęstość energii podsystemu. Jedną z zalet wykorzystania równania różniczkowego do opisu stanu energii drganiowej w polu dalekim jest fakt, że możliwe jest modelowanie efektów lokalnych takich jak lokalne tłumienie i siły punktowe. Inną ważną zaletą jest fakt, że można zdefiniować drgania przestrzenne gęstości energii pola dalekiego i przepływu energii pola dalekiego.

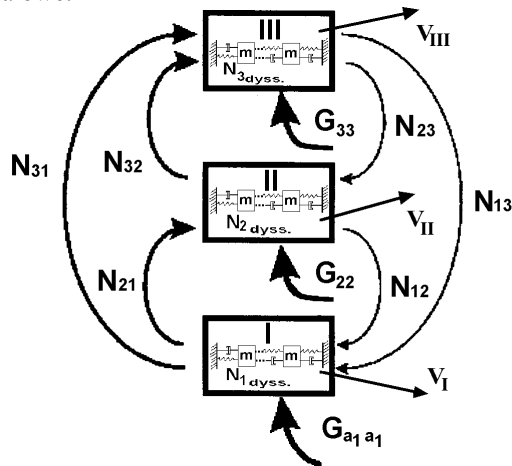
3.2. Metoda analizy rozkładu mocy obciążeń dynamicznych dla wielu podsystemów

Złożony system mechaniczny, analizowany jako układ dynamiczny modelowany energetycznie, może zawierać wiele podsystemów multimodalnych, a każdy podsystem może być sprzężony z więcej niż jednym z pozostałych podsystemów. Ponadto, każdy podsystem może być modelowany jako zbiór modów energetycznych [6]. Na przykład w przypadku obiektu mechanicznego typu belka, mody drgań poprzecznych w odpowiednim paśmie częstotliwości mogą być traktowane jako podsystem, natomiast mody obciążeń dynamicznych wzdłużnych jako inny podsystem. Wzdłużne i poprzeczne własności struktury są analizowane z energetycznego punktu widzenia jako oddzielne, separowane podsystemy ze względu na różnice w ich impedancjach i gęstościach modalnych, definiowanych jako iloraz ilości n modów energetycznych do szerokości pasma częstotliwości $n/\Delta f$ [6].

Przykładowy schemat idei dystrybucji mocy w systemie mechanicznym pokazano na rys. 1 [6]. System jest złożeniem podsystemów I, II, III, z których każdy ma charakterystyczne dlań własności dynamiczne. Podsystemami tymi mogą być np.:

I-konstrukcja wsporcza, II-układ izolatorów, III-maszyna;

I-koła jezdne ciągnika, II-podsystem resorowania, III-korpus z agregowaną maszyną i siedzisko z kierowcą;
przęsło struno-betonowe: I-belka z betonu, II-pręty stalowe.



Rys. 1. Schemat rozdziału mocy obciążeń w systemie mechanicznym (N_{ik} – elementy macierzy przenoszenia mocy, G_{ak} – gęstość widmowa mocy wymuszeń zewnętrznych) [6]

Wówczas przykładowo:

G_{a1a1} – jest mocą wejściową, wynikającą np. z występowania wymuszeń kinematycznych (nierówność jezdni),
 G_{a2a2} – mocą wejściową wymuszeń układu napędowo-roboczego,
 N_{ik} – straty energii systemu "i" w wyniku sprzężenia z podsystemem "k".

Zatem ogólnie G_{a1a1} , G_{a2a2} , G_{a3a3} są gęstościami widmowymi mocy wymuszeń zewnętrznych, N_{ik} są to moce przenoszone z podukładu k na podukład i, $N_{I_{dyss}}$, $N_{II_{dyss}}$, $N_{III_{dyss}}$ są mocami dyssypowanymi wewnątrz w poszczególnych podsystemach, V_I , V_{II} , V_{III} – moce dyssypacji zewnętrznej.

Jeśli system złożony jest z dwóch podsystemów, z których pierwszy jest wymuszany przez siły zewnętrzne a drugi wyłącznie przez sprzężenie, równanie równowagi energii ma postać.

$$\begin{cases} N_1 = \omega \eta_1 E_1 - \omega \eta_{21} E_2 \\ 0 = \omega \eta_2 E_2 + \omega \eta_{12} E_1 \end{cases} \quad (1)$$

stąd:

$$\frac{E_1}{E_2} = \frac{\eta_{12}}{\eta_2 + \eta_{21}} \quad (2)$$

Wyrażenie $\omega \eta_{12} E_1$ przedstawia energię utraconą przez podsystem 1 w wyniku sprzężenia

z podsystemem 2; $\omega \eta_1 E_1$ przedstawia energię utraconą przez podsystem 1 w wyniku tłumienia. Równanie powyższe pokazuje, że:

- Jeśli $\eta_2 \ll \eta_{21}$, to niezależnie od aktualnej wielkości sprzężenia, stosunek energii przechodzi w η_{12}/η_{21} lub η_2/η_1 (w wyniku zasady wzajemności), tzn. w stosunek między gęstościami modalnymi podsystemów. Jednym z praktycznych wyników jest fakt, że zastosowanie dodatkowego tłumienia nie będzie efektywne dopóki η_2 jest tego samego rzędu co η_{21} . Warunek spełnienia równania $E_1 \eta_2 = E_2 \eta_1$ jest podawany jako warunek równomiernego rozdziału energii i jest wzmacniany przez bliskie częstotliwości własne, małe tłumienie i duże sprzężenie. W przypadku równomiernego rozdziału energii zanika przepływ energii między podsystemami.
- Energia modalna podsystemu 2 nigdy nie przewyższa energii modalnej podsystemu wymuszanego dopóki współczynniki strat sprzężenia są dodatnie.
- Dla $\eta_{21} \ll \eta_2$ poziom energii drugiego podsystemu jest odwrotnie proporcjonalny do jego tłumienia i wprost proporcjonalny do współczynnika sprzężenia η_{12} . Jednakże w niektórych przypadkach, dla których powinna być osiągnięta mała wartość E_2 , wzrost tłumienia drugiego podsystemu nie może być bardzo efektywny i wtedy preferowane jest obniżenie współczynnika sprzężenia η_{12} .

Podstawowa zależność między sprzężonymi współczynnikami strat SEA [10] i gęstościami modalnymi jest nazywana zależnością wzajemności (równanie 1). Wzajemność wynika z faktu, że przepływ energii między dwoma podsystemami jest proporcjonalny do różnicy energii modalnych (równanie 1).

W dziedzinie częstotliwości model stanu obciążeń maszyny opisywany jest przez macierz gęstości widmowych mocy obciążeń dynamicznych w systemie mechanicznym [6, 10 - 12]:

$$\{G_{N_{ik}}(j\omega, \Theta)\} = H_{V_{ik}}(j\omega, \Theta) \cdot G_{F_{kF_k}}(j\omega, \Theta) \quad (3)$$

gdzie: $H_{V_{ik}}(j\omega, \Theta)$ – macierz mobilności dynamicznej systemu,

$G_{F_{kF_k}}(j\omega, \Theta)$ – macierz gęstości widmowych wymuszeń zewnętrznych.

Elementy macierzy charakterystyk dynamicznych [9]:

$$H_{ik}(j\omega, D(r, \Theta)) = \begin{Bmatrix} H_{i1}[j\omega, D_{11}(\Theta)] & \dots & H_{in}[j\omega, D_{1n}(\Theta)] \\ \vdots & \ddots & \vdots \\ H_{n1}[j\omega, D_{n1}(\Theta)] & \dots & H_{nn}[j\omega, D_{nn}(\Theta)] \end{Bmatrix} \quad (4)$$

są funkcjami przestrzennej miary degradacji systemu mechanicznego,

przy czym miara degradacji [2, 4]:

$$D(r, \Theta) = \frac{E_d(r, \Theta)}{E_b(r)} = \frac{\int_0^{\Theta} P[(r, \Theta), N_i] d\Theta}{E_b(r)} \quad (5)$$

W modelu dyskretnym każdy z elementów macierzy miary degradacji jest [3, 10]: $D_{ik}(\Theta) = \frac{E_{ik_d}(\Theta)}{E_{ik_{db}}}$,

przy czym: $i, k=1, 2, \dots, n$ są punktami węzłowymi obiektu,

$E_{ik_d}(\Theta)$ - zbilansowana energia degradacji elementu

„i-k” systemu mechanicznego,

$E_{ik_{db}}$ - pojemność dyssypacyjna elementu „i-k”

systemu mechanicznego.

Zastosowanie modelu rozkładu mocy wymuszeń zewnętrznych oraz modelu rozptywu mocy obciążeń w obiekcie mechanicznym daje nowoczesną, energetyczną metodę analizy trwałościowej obiektu. Umożliwia ocenę wpływu poszczególnych wymuszeń na charakterystykę trwałościową obiektu. Może stanowić również metodę weryfikacji eksperymentalnej numerycznych modeli wytrzymałościowych otrzymywanych metodami elementów skończonych. Przestrzenna charakterystyka wyczerpywania zapasu trwałości elementów systemu wyznaczana jest przez macierz mocy dyssypowanej w systemie oraz mocy sił bezwładności i sztywności dynamicznej. Elementy macierzy mocy obciążeń (część rzeczywista i część urojona) obliczane są w oparciu o wyznaczone eksperymentalnie elementy wektora odpowiedzi i elementy macierzy charakterystyk strukturalnych obiektu (metody analizy modalnej), uzyskanej z prostych pomiarów drganiowych. Intensywność procesów degradacji odzwierciedlana jest w wartościach mocy obciążeń przenoszonych i obciążeń destrukcyjnych, których skutki działania akumulowane są w podzespołach.

Przykład

Budując holistyczny model obciążeń maszyny należy uwzględnić moc wejściową N_i , moc użyteczną N_u i moc dyssypowaną N_d , traconą na procesy degradacji obiektu.

W agregacie: źródło napędu – obiekt, traktowanym jako system mechaniczny, tworząc bilans mocy można zapisać:

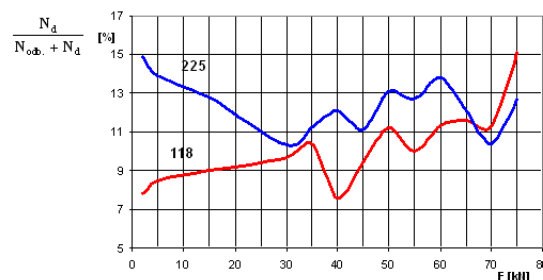
$$N_{zr.} = N_{odb.} + N_u + N_d \quad (6)$$

gdzie: $N_{zr.}$ - moc źródła

$N_{odb.}$ - moc odbitej

N_u - moc użyteczna

N_d - moc procesów degradacji maszyny



Rys. 2. Wykres funkcji $\frac{N_d}{N_{odb.} + N_u}$,
opisującej podatność degradacyjną
obiektu mechanicznego

Na rys. 2 przedstawiono wykres funkcji $\frac{N_d}{N_{odb.} + N_u}$ uzyskanej z wyników badań

atestacyjnych procesu degradacji (aż do pęknięcia) belek struno-betonowych. Na stanowisku badań przyspieszonych realizowano badania diagnostyczne stanu technicznego dwóch belek struno-betonowych sprężonych, różniących się wielkością napęcia wstępnego prętów stalowych (belka 125 – wyższe napiecie wstępne), poddawanych wymuszeniom testującym zadawanym kolejno w narastających stanach degradacji (wzrost głębokości szczeliny). W wyniku zadawania kolejno następujących obciążeń statycznych F , przykładanych w środku belki: 5, 10, ..., 75 kN następowała degradacja belki.

Funkcja $\frac{N_d}{N_{odb.} + N_u}$ opisuje zmienną podatność

degradacyjną obiektu w wyniku jego zmian strukturalnych, spowodowanych wzrastającym obciążeniem statycznym. Z wykresu można odczytać duże wartości energii odbitej oraz zmiany wartości mocy sił degradacji belki w funkcji wzrostu obciążenia, a także wartości obciążenia powodującego inicjację procesu pęknięcia belki.

4. WNIOSKI

1. Procesy stanowiące o charakterystyce trwałościowej maszyny mają wymiar energetyczny.
2. Metoda analizy rozptywu energii obciążeń dynamicznych w systemach mechanicznych umożliwia wyznaczanie charakterystyk procesu degradacji technicznej obiektów mechanicznych i opis ich stanu technicznego.
3. Badania metodą testu impulsowego pozwalają wyznaczyć obciążenia inicjujące proces uszkodzenia struktury konstrukcji.

LITERATURA

- [1] Cempel C., Tabaszewski M.: *Teoria degradacji maszyn i urządzeń jako ewolucyjnych systemów otwartych z ograniczoną dyssypacją* -

- sprawozdanie z badań KBN 0858/P4/93*. Politechnika Poznańska (IMS), Poznań 1996.
- [2] Cempel C., *Teoria inżynierii systemów*, skrypt, Zakład Dynamiki - Wibroakustyki Systemów, Politechnika Poznańska, Poznań 2000.
- [3] Cempel C., *Modele systemów przetwarzania energii w teorii i inżynierii systemów*. Promocja na Doktora Honoris Causa Politechniki Szczecińskiej, Politechnika Szczecińska, 1995.
- [4] Cempel C., *Modele diagnostyki wibroakustycznej*, Diagnostyka Maszyn Roboczych i Pojazdów, Bydgoszcz 1994.
- [5] Cieślak J., *Natężenie dźwięków strukturalnych w diagnostyce nieciągłości elementów konstrukcyjnych*, 3rd International Congress of Technical Diagnostics, Poznań 2004.
- [6] Kaźmierczak H., *Analiza rozkładu mocy obciążeń dynamicznych w systemach mechanicznych*, Rozprawy Nr 363, Wydawnictwo Politechniki Poznańskiej, Poznań 2001.
- [7] Kaźmierczak H., *Dynamic load power distribution in mechanical systems*, Zagadnienia Eksploatacji Maszyn, Zeszyt 3(127), 127-141, ITE Radom 2003.
- [8] Kaźmierczak H., *Badania trwałości zmęczeniowej maszyn metodą analizy rozkładu mocy obciążeń dynamicznych*, Diagnostyka vol.26, 133-142, 2002, PTDT przy Wydziale Nauk Technicznych PAN.
- [9] Kaźmierczak H., *Energetic description of the destruction process of machine structural nodes*, Machine Dynamics Problem, Vol. 27, No 3, 113-123, Warszawa 2003.
- [10] Kaźmierczak H., *Analiza destrukcji maszyny metodą rozkładu mocy obciążeń dynamicznych*, DIAGNOSTICS'2004 – 3rd International Congress of Technical Diagnostics.
- [11] Kaźmierczak H., Kromulski J., Pawłowski T., *Energetyczne charakterystyki degradacji przyczepy*, Diagnostyka vol.33, 2005, PTDT przy Wydziale Nauk Technicznych PAN.
- [12] Kaźmierczak H., Kromulski J., Cempel C., Barczewski R., *Energetic description of the destruction process of steel concrete structures*, COST Action 534 New Materials and Systems for Prestressed Concrete Structures, Workshop of COST on NTD Assessment and New Systems in Prestressed Concrete Structures, Radom 2005.
- [13] Kaźmierczak H., Pawłowski T., Kromulski J., Barczewski R., *Vibration characteristics of the technical degradation process of an object*, 13 International Congress on Sound and Vibration 2–6 July 2006 Wiedeń.
- [14] Łagoda T., *Energetyczne modele oceny trwałości zmęczeniowej materiałów konstrukcyjnych w warunkach jednoosiowych i wieloosiowych obciążeń losowych*. Studia i monografie, z. 121. Politechnika Opolska, Opole 2001.
- [15] Radkowski S., *Wibroakustyczna diagnostyka uszkodzeń niskoenergetycznych*, Wydawnictwo Instytutu Technologii Eksploatacji, Warszawa – Radom 2002.
- [16] Szala J., *Zmęczeniowe pękanie materiałów i konstrukcji – rozwój nauki i zastosowań praktycznych*, Zagadnienia Eksploatacji Maszyn, zeszyt 2 (126), vol. 36, Instytut Technologii Eksploatacji, Radom 2001.
- [17] Szala J., Szala G., *Dwuparametryczne charakterystyki zmęczeniowe – sformułowanie problemu*, Problemy Eksploatacji Nr 3/2001(42), Instytut Technologii Eksploatacji, Radom 2001.
- [18] Wegner T., *Metody energetyczne w wytrzymałości materiałów, Hipoteza wytrzymałościowa stateczności równowagi wewnętrznej*. Wydawnictwo Politechniki Poznańskiej, 81 stron, Poznań 1999.



Doc dr hab. **Henryk KAŹMIERCZAK** – absolwent Wydziału Mat. Fiz. Chem. Uniwersytetu im. Adama Mickiewicza w Poznaniu, stopień doktora nauk technicznych uzyskał w 1977r. na Wydziale Budowy Maszyn Politechniki Poznańskiej. Stopień doktora habilitowanego nauk technicznych z dziedziny mechanika uzyskał w 2002r. na Wydziale Budowy Maszyn i Zarządzania Politechniki Poznańskiej. Jest autorem ponad 250 publikacji naukowych. Zajmuje się zagadnieniami z dziedziny dynamiki maszyn, diagnostyki technicznej, identyfikacji własności dynamicznych maszyn, w tym metodami analizy modalnej. Ostatnie publikacje dotyczą energetycznego modelowania obciążeń w maszynach oraz zastosowań metody analizy rozkładu mocy obciążeń dynamicznych do badań procesów degradacji i trwałości maszyn. Jest członkiem krajowych i zagranicznych towarzystw naukowych.

ZAGADNIENIE BILANSOWANIA OBIEGU CIEPLNEGO Z ZASTOSOWANIEM RACHUNKU WYRÓWNAWCZEGO W DIAGNOSTYCE OKRĘTOWYCH TURBINOWYCH SILNIKÓW SPALINOWYCH

Marek DZIDA*, Andrzej ADAMKIEWICZ**

* Politechnika Gdańska, Wydział Oceanotechniki i Okrętownictwa
ul. G. Narutowicza 11/12, 80-952 Gdańsk-Wrzeszcz
e-mail: dzida@pg.gda.pl; tel: (058) 34 72 135

** Akademia Morska w Szczecinie, Wydział Mechaniczny
ul. Wały Chrobrego 1-2, 70-500 Szczecin
e-mail: andrzej.adamkiewicz@am.szczecin.pl; tel: (091) 48 09 384

Streszczenie

Referat przedstawia metodę określania niedostępnych pomiarowo parametrów stosowanych w diagnozowaniu okrętowych turbinowych silników spalinowych. Na przykładzie silnika jednowirnikowego pokazano parametry sygnału diagnostycznego. Przedstawiono wstępne szacowanie parametrów w bilansowaniu obiegu jednowirnikowego turbinowego silnika spalinowego. Pokazano przykładowe wyniki obliczeń sprawności termicznej komory spalania z wykorzystaniem uwiarygodnionych za pomocą rachunku wyrównawczego zarejestrowanych danych pomiarowych. Przedstawiono dyskusję otrzymanych wyników dla prezentowanej metody.

Słowa kluczowe: turbinowy silnik spalinowy, diagnoza, parametr pracy, rachunek wyrównawczy.

A PROBLEM OF BALANCING THERMAL CYCLE USING RECONCILIATION CALCULATIONS IN THE DIAGNOSTICS OF MARINE GAS TURBINE ENGINES

Summary

This paper presents a method of determination of experimentally inaccessible parameters for diagnosing marine gas turbine engines. The applied parameters of the diagnostic signal were shown for a one-rotor engine, chosen as an example. Initial estimations of parameters for balancing the cycle of a one-rotor gas turbine engine were presented. Example results of calculating thermal efficiency of the combustion chamber, verified by reconciliation calculations of the recorded values, were given. A discussion of the obtained results for the studied method was presented.

Keywords: marine gas turbine engine, diagnosis, operation parameter, reconciliation calculations.

1. WPROWADZENIE

Specyfika eksploatacji silników turbinowych w warunkach okrętowych powoduje, iż diagnozowane stany silników reprezentowane są przez ciągi dyskretnych sygnałów diagnostycznych, rejestrowanych przez okrętową sieć pomiarową w niejednakowych odcinkach czasu. Pewnymi wskaźnikami stanu całego zespołu prądotwórczego z turbinowym silnikiem spalinowym jak i jego podzespołów są sprawności wewnętrzne: - sprężarki η_{is} , turbiny η_{iT} oraz sprawność termiczna komory spalania, definiowana zależnością w postaci:

$$\begin{aligned}\eta_{tKS} &= \frac{\dot{Q}_{spal3} - \dot{Q}_{pow2}}{\dot{Q}_{pal}} = \\ &= \frac{\dot{m}_{spal3} i_{spal3} - \dot{m}_{pow2} i_{pow2}}{\dot{m}_{pal} W_d} = \\ &= \frac{\dot{m}_{spal3} c_{p,spal} T_3 - \dot{m}_{pow2} c_{p,pow} T_2}{\dot{m}_{pal} W_d}\end{aligned}\quad (1)$$

gdzie strumień masy spalin wypływających z komory spalania

$$\dot{m}_{spal3} = \dot{m}_{pow2} + \dot{m}_{pal} \quad (2)$$

$i_3(T_3)$ – entalpia właściwa spalin na wylocie z komory spalania,

$i_2(T_2)$ – entalpia właściwa powietrza za sprężarką,

$\dot{Q}_{pal}(W_d)$ – strumień ciepła dostarczanego do komory spalania w postaci energii chemicznej paliwa o wartości opałowej dolnej W_d ,

W pracy przedstawiono wyniki obliczeń sprawności termicznej komory spalania z wykorzystaniem uwiarygodnionych za pomocą rachunku wyrównawczego wyników pomiarów parametrów pracy jednowirnikowego turbinowego silnika spalinowego GTU-6A o mocy nominalnej 600 kW. Schemat zdekomponowanego silnika z oznaczonymi strumieniami energii przedstawiono na rys. 1.

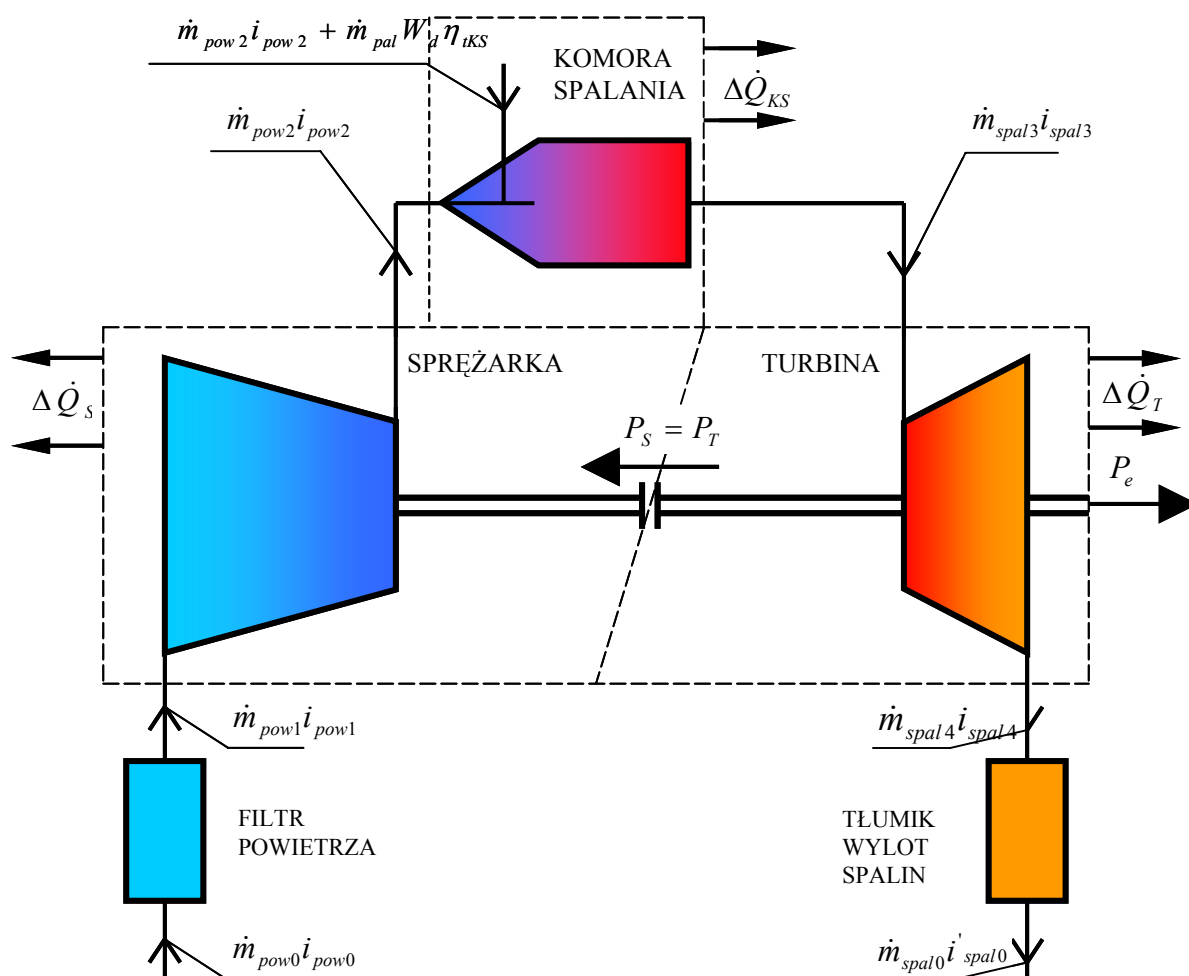
Strumień masy paliwa $\dot{m}_{pal}$ zużywanego przez silnik turbinowy jest wielkością mierzoną bezpośrednio, natomiast strumień ciepła $\dot{Q}_{spal3}$ i $\dot{Q}_{pow2}$ oblicza się w oparciu o wyniki pomiarów

wielkości reprezentujących niewiadome strumienie energii: potencjalnej – ciśnienie powietrza w końcu procesu sprężania p_2 ; zasilającej silnik – chemicznej, zdeterminowany wartością opałową dolną paliwa W_d ; energii

powietrza zasysanego przez silnik $\dot{Q}_{pow1}$ określony przez ciśnienie p_1 i temperaturę T_1 ; energii cieplnej spalin której miarą jest temperatura przed turbiną T_3 lub zastępczo za turbiną T_4 oraz energię powietrza z temperaturą za sprężarką T_2 .

Do określenia wielkości nieznanych i niedostępnych pomiarowo rozpatrzono więzy energetyczne wyznaczone przez bilanse strumieni energii (np. moc, strumień ciepła), a nie ilość energii (moment obrotowy ciepło).

W praktyce każdy opis stanu jest ograniczony dostępnością pomiarową i jest modelem budowanym z uwzględnieniem przyjętych założeń upraszczających.



Rys. 1. Schemat przepływu strumieni energii przez podzespoły jednowirnikowego turbinowego silnika spalinowego

2. SZACOWANIE WIELKOŚCI NIEDOSTĘPNYCH POMIAROWO

Wielkościami trudnymi pomiarowo, a często niewiadomymi w opisie procesów w silnikach turbinowych jest temperatura spalin przed turbiną i strumień masy powietrza podawanego do komory spalania. Wielkości te obliczane na podstawie równań bilansowych, uzupełnia zbiór bezpośrednio mierzonych parametrów sygnału diagnostycznego,

a dokładność pomiaru ma wpływ na niepewność wyznaczenia sprawności maszyn wirnikowych silnika i komory spalania [1].

Uwiarygodnienie pomiarów i wielkości nieznanych w bilansach masy i energii umożliwia zastosowanie metod rachunku wyrównawczego. Jego algorytm wymaga wstępnego oszacowania wielkości niewiadomych. Stosując klasyczną metodę uzgadniania wartości wielkości niewiadome obliczane są z wybranych równań bilansowych zakładając, że dokładność wstępnego szacowania nie jest znana [1, 3]. Przy dużych niepewnościach wielkości mierzonych oraz pewnej strukturze równań wartości poprawek wielkości niewiadomych mogą powodować, że estymaty tych wielkości znajdują się w przestrzeni rozwiązań niedopuszczalnych z punktu widzenia fizyki zagadnienia. W klasycznej metodzie uzgadniania liczba niezależnych równań modelu musi być większa od liczby wielkości niewiadomych.

Na podstawie dostępnych danych producenta oraz wyników prób gwarancyjnych, (wykonanych poza systemem energetycznym okrętu, z większą dokładnością pomiarów niż wykonywane mogłyby być z pomocą okrętowych układów pomiarowych), wstępnie oszacowano wartości wielkości niewiadomych. Przyjęto dlatego, że dokładność tego szacowania była znana. Uogólniona metoda rachunku wyrównawczego wykorzystuje takie informacje o dokładności wstępnego szacowania niewiadomych.

W bilansowaniu obiegu silnika turbinowego wielkością zamykającą jest strumień entalpii spalin wypływających z komory spalania:

$$\dot{m}_{spal3} i_{spal3} = \dot{m}_{spal4} i_{spal4} + P_T + \Delta \dot{Q}_T \quad (3)$$

przy czym dla komory spalania obowiązuje równość w postaci:

$$\dot{m}_{pal} i_{pal} + \dot{m}_{pal} W_d \eta_{KS} + \dot{m}_{pow2} i_{pow2} = \dot{m}_{spal3} i_{spal3} + \Delta \dot{Q}_{KS} \quad (4)$$

a dla strumienia powietrza opuszczającego sprężarkę:

$$\dot{m}_{pow1} i_{pow1} + P_S = \dot{m}_{pow2} i_{pow2} + \Delta \dot{Q}_S \quad (5)$$

gdzie w równaniach (3) – (5)

$$\Delta \dot{Q}_T, \Delta \dot{Q}_{KS}, \Delta \dot{Q}_S$$

– strumienie energii rozpraszanej odpowiednio przez

turbinę, komorę spalania i sprężarkę do otoczenia.

Uwzględniając fakt, iż moc rozwijana przez turbinę zużywana jest na napęd sprężarki oraz przekazywana odbiornikowi energii jako moc efektywna, równa jest ona sumie mocy:

$$P_T = P_S + P_e \quad (6)$$

równanie (3) dla turbiny można zapisać w postaci:

$$\dot{m}_{spal3} i_{spal3} = \dot{m}_{spal4} i_{spal4} + P_S + P_e + \Delta \dot{Q}_T \quad (7)$$

Równanie bilansowe silnika otrzymano składając równania bilansowe dla jego podzespołów (3) – (5):

$$\begin{aligned} \dot{m}_{pow1} i_{pow1} + \dot{m}_{pal} i_{pal} + \dot{m}_{pal} W_d \eta_{KS} = \\ = \dot{m}_{spal4} i_{spal4} + P_e + \Delta \dot{Q}_\Sigma \end{aligned} \quad (8)$$

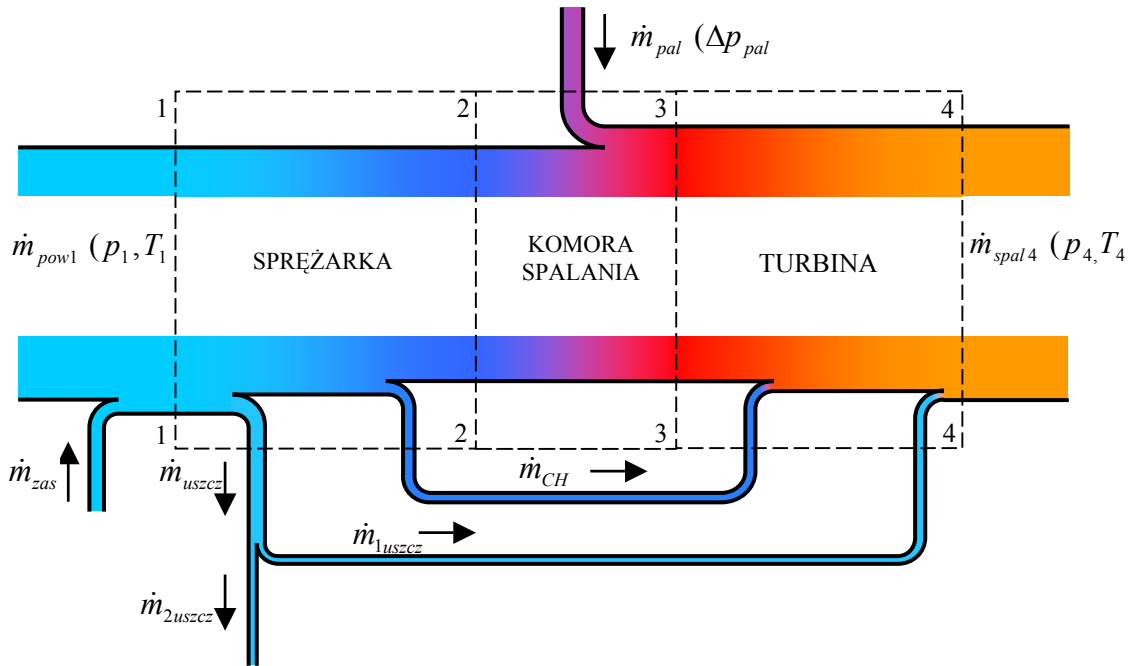
gdzie:

$$\Delta \dot{Q}_\Sigma = \Delta \dot{Q}_S + \Delta \dot{Q}_{KS} + \Delta \dot{Q}_T \quad (9)$$

w którym pozostały tylko wyrażenia uwzględniające połączenie strumieni energii z otoczeniem. Chociaż równanie (8) powstało z równań bilansowych maszyn silnika, to nie zawiera ono nowej informacji. Jest ono wykorzystywane w miejsce równania bilansowego jednej z maszyn silnika, w zależności od dostępności pomiarowej parametrów pracy.

Strumień spalin wylotowych wynika z bilansu strumieni powietrza roboczego, zasysanego, chłodzącego, uszczelniającego i strumienia paliwa. Model bilansu strumieni masy czynników roboczych w silniku przedstawiono na rysunku 2.

Wśród zmiennych stanu występujących w równaniach modelu opisującego przetwarzanie energii w silniku [1] są takie, których bezpośredni pomiar można dokonać z możliwą do zaakceptowania dokładnością, ale jest utrudniony technicznie lub praktycznie niemożliwy. Rezygnując z parametrów trudnych do zmierzenia na przyjętym poziomie oprzyrządowania układu pomiarowego silnika można je obliczyć, uzyskując dokładność znacznie wyższą, niż otrzymana przy bezpośrednim pomiarze.



Rys. 2. Model bilansu strumieni masy głównego czynnika roboczego w silniku jednowirnikowym

Metoda ta nie nakłada ograniczeń na liczbę równań bilansowych. Jednak należy zachować ostrożność przy wstępnym szacowaniu wartości i macierzy kowariancji wielkości niewiadomych. Optymalny dobór tych wielkości jest zagadnieniem złożonym, zależnym od specyfiki rozwiązywanego problemu.

Stwierdzenie to szczególnie dotyczy wstępnego oszacowania wartości i niepewności oszacowania niewiadomej entalpii spalin opuszczających komorę spalania:

$$i_{spal3} = \frac{1}{\dot{m}_{spal3}} (\dot{m}_{spal4} i_{spal4} + P_S + P_e + \Delta \dot{Q}_T) \quad (10)$$

Z występujących w równaniach bilansowych parametrów największe trudności z określeniem, z zadowalająco wysokim stopniem dokładności, wykazuje moc zapotrzebowana przez sprężarkę. Do jej określania zastosowano równanie bilansowe sprężarki (5) w postaci:

$$P_S = \dot{m}_{pow2} i_{pow2} - \dot{m}_{pow1} i_{pow1} + \Delta \dot{Q}_S \quad (11)$$

Strumień powietrza zasysanego przez sprężarkę jest pomniejszony za sprężarką o strumień powietrza pobieranego do chłodzenia gorących detali silnika

$$\dot{m}_{pow2} = \dot{m}_{pow1} - \dot{m}_{pow.CH} \quad (12)$$

to definiując wskaźnik udziału powietrza chłodzącego gorące detale silnika w strumieniu powietrza zasysanego przez sprężarkę w postaci

$$\gamma = \frac{\dot{m}_{pow.CH}}{\dot{m}_{pow1}} \quad (13)$$

i podstawiając zależności (12) i (13) do równania (11), po przekształceniu otrzymano

$$P_S = \dot{m}_{pow1} [(1 - \gamma) i_{pow2} - i_{pow1}] + \Delta \dot{Q}_S \quad (14)$$

Do obliczenia sprawności termicznej komory spalania posłużono się równaniem dla całego silnika (8), otrzymując po przekształceniu wyrażenie (1) w postaci:

$$\eta_{KS} = \frac{1}{\dot{m}_{pal} W_d} (\dot{m}_{spal4} i_{spal4} + P_e + \dot{Q}_\Sigma - \dot{m}_{pow1} i_1 - \dot{m}_{pal} i_{pal}) \quad (15)$$

Identyfikacja eksperymentalna parametrów prawej strony równania (15) powinna być dokonana

z minimalną, osiągalną niepewnością. Dotyczy to pomiaru strumienia paliwa lub ciśnienia paliwa za pompą paliwową, strumienia powietrza i temperatury spalin do określenia ich entalpii według zależności (10) oraz mocy efektywnej silnika.

3. DOSTĘPNOŚĆ I DOKŁADNOŚĆ POMIAROWA PARAMETRÓW SYGNAŁU DIAGNOSTYCZNEGO

Kierując się wymaganiami przydatności w diagnozowaniu oraz uwzględniając:

- reprezentację strumieni energii użytecznej i zasilające silnik,
- występowanie w zależnościach opisujących przetwarzanie energii w silnikach,

- doświadczalnie potwierdzoną eksploatacyjną pewność pomiarów,
- spełnienie podstawowych wymagań producenta pod względem bezpieczeństwa pomiarów,
- zagwarantowania prawidłowości funkcjonowania silnika, wiarygodności pomiarowej i możliwości automatycznej rejestracji,

pomiarom poddano następujące parametry sygnału diagnostycznego:

$$D_{mk} = \{P_{ek}, p_{2k}, T_{2k}, p_{palk}, T_{4k}, p_0, T_0\} \quad (16)$$

gdzie: D_{mk} – zbiór wielkości zmierzonych podczas k – tej obserwacji;

P_{ek} – moc efektywna silnika;

p_{2k}, T_{2k} – ciśnienie i temperatura powietrza za sprężarką;

p_{palk} – ciśnienie paliwa za pompą;

T_{4k} – temperatura spalin za turbiną;

p_{0k}, T_{0k} – ciśnienie i temperatura powietrza atmosferycznego.

Rozwój technik pomiarowo-informatycznych umożliwia stosowanie różnorodnych procedur obliczeniowych wykorzystujących wyniki pomiarów [1, 3] z rozproszonych układów pomiarowo-rejestrujących. Klasa dokładności okrętowych układów pomiarowych nie zapewnia jednak uzyskania wiarygodności obliczeń sprawności maszyn na poziomie pomiarów zdawczych czy gwarancyjnych. Dla potrzeb pracy zbudowano specjalny układ rejestrujący parametry sygnału diagnostycznego podczas niestacjonarnego testu eksploatacyjnego [1]. Posłużył on do określenia wartości średnich parametrów przy różnych mocach silników. Zastosowano tu czujniki okrętowego układu pomiarowego. Z zarejestrowanych przebiegów parametrów wyselekcjonowano serie parametrów pracy potraktowane jako sekwencje czasowe stanów pracy ustalonej silników. Dla poszczególnych serii pomiarowych obliczono

wartości średnie parametrów $\bar{d}_i$ oraz eksperymentalne odchylenia standardowe σ_{d_i} .

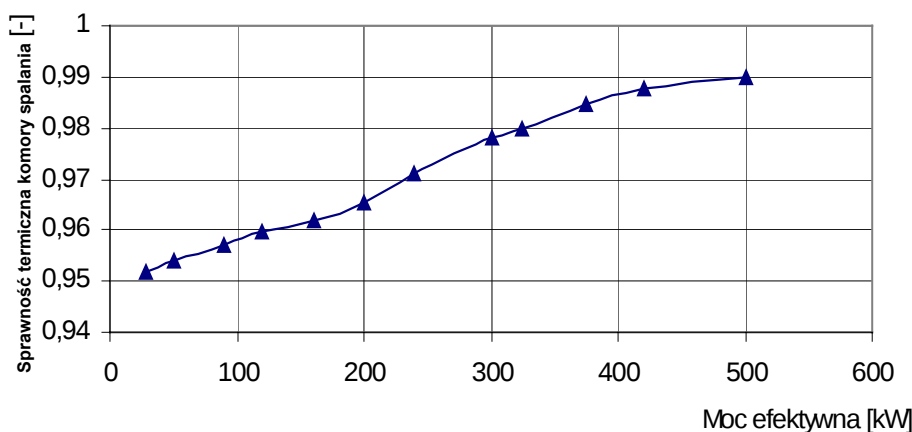
Na ich podstawie obliczono średnie błędy względne serii pomiarów, zgodnie z zależnością:

$$\delta_{d_i} = \frac{\sigma_{d_i}}{\bar{d}_i} \quad (17)$$

Dla otrzymanych populacji par: *zmierzony parametr diagnostyczny – moc efektywna silnika* obliczano wielkości niedostępne pomiarowo, jako funkcje mocy efektywnej silnika. Zastosowano w tym celu modele regresyjne z funkcjami opisującymi w postaci wykładniczej. W oparciu o przyjęte funkcje opisujące, dla mocy wyróżnianych w dokumentacji silnika, estymowano wartości wielkości niedostępnych pomiarowo. Przykładowe wyniki dla sprawności termicznej komory spalania η_{tKS} , określonej na podstawie zależności (15), w funkcji mocy silnika przedstawiono na rysunku 3.

Dla porównania dokładności wyników pomiarów szacowanych wielkości w tabeli 1 podano dopuszczalne wartości niepewności względnych parametru d_i okrętowego układu pomiarowego silnika typu GTU – 6A δ_{OUPd_i}

ustalone na podstawie dokumentacji technicznej silnika oraz maksymalne niepewności względne uzyskane podczas rejestrowanych obserwacji parametrów. Dla temperatury spalin, ciśnienia powietrza za sprężarką i ciśnienia paliwa za pompą niepewności względne metody pomiarowej są mniejsze od dopuszczalnych niepewności okrętowych układów pomiarowych. Jedynie względne niepewności pomiarów ciśnienia paliwa i strumienia masy paliwa okazały się większe od dopuszczalnych niepewności względnych okrętowych układów pomiarowych.



Rys. 3. Przykład zależności sprawności termicznej komory spalania jako funkcji mocy silnika na podstawie estymowanej wartości temperatury T_3

Tabela 1. Niepewności metody pomiarowej i układu pomiarowego

Parametr diagnostyczny	Niepewność względna metody pomiarowej	Niepewność względna okrętowych układów pomiarowych
$\bar{d}_i$	$\delta_{d_i} [\%]$	$\delta_{OUPd_i} [\%]$
Moc silnika	0,3814	1,0
Temperatura spalin za turbiną	0,2448	1,5 - 2,8
Ciśnienie powietrza za sprężarką	0,2485	0,6 - 1,3
Ciśnienie paliwa za pompą	1,1104	1,0
Strumień masy paliwa	1,3725	1,0
Prędkość obrotowa wirnika	0,1105	0,3 - 1,5

4. UWIARYGODNIENIE WYNIKÓW POMIARÓW I WIELKOŚCI NIE MIERZONYCH ZA POMOCĄ RACHUNKU WYRÓWNAWCZEGO

Rachunek wyrównawczy zastosowano w celu skorygowania wielkości zmierzonych i wstępnie obliczonych lub oszacowanych niewiadomych, tutaj np. sprawności komory spalania. W rachunku tym równania zachowania składające się na model matematyczny były spełnione przy równoczesnej minimalizacji odchyłek wyników pomiarów od wartości skorygowanych. Zastosowano tu zasadę maksymalizacji funkcji wiarygodności w n -wymiarowej przestrzeni błędów [2]. Wynikający z niej estymator największej wiarygodności jest równoważny estymatorowi o minimalnej wartości błędów estymacji w metodzie najmniejszych kwadratów Gaussa. W uogólnionej metodzie uzgadniania, zgodnie z zasadą estymacji średniokwadratowej, ważoną funkcję ryzyka można zapisać następująco:

$$(\bar{D}_m - D_c)^T M^{-1} (\bar{D}_m - D_c) \rightarrow \min \quad (18)$$

gdzie:

$\bar{D}_m$ – wektor wielkości mierzonych niewiadomych,

D_c – wektor wyników pomiaru i wielkości niewiadomych, wstępnie oszacowanych,

M – macierz kowariancji wielkości mierzonych i oszacowanych niewiadomych.

Ekstremum jest warunkowe, ponieważ wielkości mierzone i niewiadome są ze sobą

wzajemnie powiązane równaniami modelu matematycznego analizowanego procesu:

$$F(\bar{d}) = 0 \quad (19)$$

Nieliniowy układ równań modelu (19) stanowi równania warunków w procedurze uzgadniania, polegającego na wyznaczeniu wartości poprawek i wyrównanych ocen wielkości mierzonych oraz niewiadomych w celu spełnienia równań modelu matematycznego.

5. PODSUMOWANIE

Jeżeli pod pojęciem równania bilansowego rozumiane jest matematyczne wyrażenie praw zachowania energii i praw zachowania substancji dla maszyn, to dla ustalonych warunków pracy silnika sprowadzają się one do bilansu energii doprowadzonej i odprowadzonej od maszyny wirnikowej i komory spalania oraz ilości substancji odprowadzonej od nich do innych maszyn silnika lub do otoczenia. W żadnej z maszyn silnika nie zachodzi wówczas akumulacja lub wydzielanie energii, czy też substancji.

Można wskazać dwie korzyści wykorzystania równań bilansowych i rachunku wyrównawczego:

1. Korektę wartości danych pomiarowych parametrów występujących w równaniach bilansowych
2. Określenie wielkości niedostępnych pomiarowo i nieznanymi, poprzez pomiar z zadowalającą dokładnością ograniczonej liczby parametrów diagnostycznych i parametryzację drogą obliczeń, pozostałych zmiennych wchodzących w skład równań.

Wśród zmiennych stanu występujących w równaniach modelu opisującego przetwarzanie energii w silniku [1] są takie, których bezpośredni pomiar można dokonać z możliwą do zaakceptowania dokładnością, ale jest utrudniony technicznie lub praktycznie niemożliwy. Rezygnując z parametrów trudnych do zmierzenia na przyjętym poziomie oprzyrządowania układu pomiarowego silnika można je obliczyć, uzyskując dokładność znacznie wyższą, niż otrzymana przy bezpośrednim pomiarze.

Do obliczenia sprawności komory spalania zgodnie z zależnością (15), niezbędne jest określenie mocy zapotrzebowanej przez sprężarkę wg (14) do czego niezbędna jest znajomość strumienia masy oraz ciśnienia i temperatury powietrza przed i za sprężarką, jako zmiennych stanu entalpii. Wówczas można uniknąć jednego z najtrudniejszych pomiarów – średniej temperatury spalin za komorą spalania, a przed turbiną, potrzebnej do określenia entalpii spalin.

Obliczenia pokazały, że sprawność komory spalania obniża się podczas pracy silnika przy

częściowych obciążeniach i biegu jałowym, chociaż jest zawsze niewiele mniejsza od jedności.

Prezentowana metoda, przy ograniczonej dostępności pomiarowej, jak to ma miejsce w przypadku okrętowych silników turbinowych, staje się metodą weryfikacyjną, stwarzającą możliwość lokalizowania niesprawności w podzespołach silnika. W tym celu przyjmuje się hipotezę, że wielkości charakteryzujące podstawowe podzespoły przyjmują wartości w przedziałach zmian dopuszczalnych, natomiast pewne jako niesprawne. Na podstawie wyników pomiarów można weryfikować przyjętą hipotezę dla różnych kombinacji parametrów pracy. Uzyskanie pozytywnego wyniku weryfikacji jest jednoznaczne ze zlokalizowaniem elementów niesprawnych lub niezdatnych.

W sytuacji małej dostępności pomiarowej silnika turbinowego, gdy liczba równań bilansowych, które można ułożyć wykorzystując dostępne wyniki pomiarów, jest mniejsza od liczby niewiadomych, należy stosować metodę optymalizacyjną z wykorzystaniem rachunku wyrównawczego. Wówczas układ równań jest nieoznaczony i istnieje nieskończenie wiele jego rozwiązań. Cechą charakterystyczną tak postawionego problemu jest fakt, że liczba wielkości charakteryzujących elementy niesprawne lub niezdatne nie różni się znacznie od ilości parametrów diagnostycznych.

Przedstawiona metoda oceny stanu silników turbinowych nie jest uniwersalna, a jej wady i zalety uwidaczniają się podczas diagnozowania silników o większym stopniu złożoności. Jednak jest to jedyna możliwa droga praktycznej oceny zdolności silników do funkcjonowania w eksploatacji.

LITERATURA

- [1] Adamkiewicz A.: *Application of Balance Equations for Determination of Experimentally Inaccessible Quantities in the Diagnostics of Marine Gas Turbine Engines*. The Third Scientific International Scientific Symposium on Technical COMPOWER'04, Poland, Gdańsk, 2 – 3 December, 2004. Economic and Environmental Aspects of Combined Cycle Power Plants, Gdańsk University of Technology, Alstom Power Ltd., Gdańsk 2004. Monografia pod red. prof. Z. Domachowskiego, w rozdziale: Turbine Plant Diagnostics, ss.123 – 133.
- [2] Sanecki J.: *Elementy rachunku wyrównawczego*. Fundacja Rozwoju Wyższej Szkoły Morskiej w Szczecinie, Szczecin 2004.
- [3] Szargut J. (red.): *Rachunek wyrównawczy w technice cieplnej*. Ossolineum, Wrocław 1985.

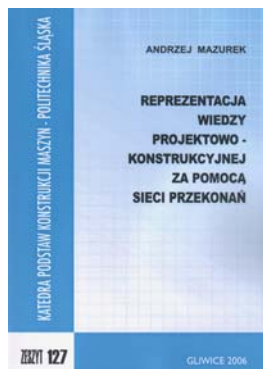
Marek DZIDA, dr hab. inż., prof. ndzw. Politechniki Gdańskiej, pracownik naukowo-dydaktyczny Katedry Automatyki Okrętowej i Napędów Turbinowych. Zainteresowania naukowe: siłownie okrętowe, turbiny parowe i gazowe. Dziekan Wydziału Oceanotechniki i Okrętownictwa PG.



Andrzej ADAMKIEWICZ, dr hab. inż., prof. ndzw. Akademii Morskiej w Szczecinie, pracownik naukowo - dydaktyczny Instytutu Technicznej Eksploatacji Siłowni Okrętowych Wydziału Mechanicznego Akademii Morskiej. Doktoryzował się w Instytucie Maszyn Przepływowych PAN w Gdańsku w 1991 r. (promotor prof. dr hab. inż. Jerzy Krzyżanowski). Stopień doktora habilitowanego uzyskał na podstawie kolokwium przeprowadzonego na Wydziale Maszyn Roboczych i Transportu Politechniki Poznańskiej w 2005 r. Zainteresowania naukowe autora obejmują zagadnienia: eksploatacji siłowni okrętowych, turbin parowych i gazowych, utylizacji ciepła w statkowych systemach odzyskiwania energii odpadowej, cieplnych maszyn wirnikowych oraz diagnostykę maszyn i urządzeń okrętowych. Prowadzi zajęcia dydaktyczne z: maszyn przepływowych, turbin parowych i gazowych, okrętowych napędów parowych, kotłów okrętowych i projektowania siłowni okrętowych.

Zeszyty naukowe – Katedry Podstaw Konstrukcji Maszyn – Politechniki Śląskiej

Poniżej prezentujemy Państwu najnowsze zeszyty naukowe wydawane przez Katedrę Podstaw Konstrukcji Maszyn – Politechniki Śląskiej.



Andrzej MAZUREK „Reprezentacja wiedzy projektowo-konstrukcyjnej za pomocą sieci przekonań” Zeszyt 127

Praca stanowi próbę oceny przydatności zastosowania sieci przekonań do reprezentacji wiedzy przybliżonej w dziedzinie projektowania i konstruowania maszyn. Opracowano odpowiednią metodologię oraz określono sposób definiowania modelu projektowanego obiektu. Zwrócono między innymi uwagę na konieczność podziału złożonego zadania na podzadania, które można rozpatrywać niezależnie oraz omówiono proces łączenia poszczególnych – powstałych w ten sposób – podsieci, umożliwiające budowę modelu globalnego.

Opracowaną metodę zweryfikowano dla wybranego zadania projektowego dotyczącego łożyskowania wałów maszynowych, a uzyskany model posłużył do opracowania systemu doradczego działającego w środowisku oprogramowania do komputerowo wspomaganego projektowania.

Wspomaganie procesu projektowo-konstrukcyjnego traktowanego jako proces intelektualny wiąże się z koniecznością zastosowania odpowiednich metod umożliwiających pozyskiwanie, gromadzenie i wykorzystanie wiedzy inżynierskiej. Obecnie dużym uznaniem cieszą się rozwiązania wykorzystujące metody i techniki sztucznej inteligencji. W odniesieniu do procesu rozumowania projektanta można stwierdzić, że całość działań związanych z realizacją określonych etapów projektu opiera się na wnioskowaniu przybliżonym i niedokładnym. Wykorzystując podejście zaproponowane przez Prof. W. Cholewę, polegające na zastosowaniu sieci przekonań do reprezentacji wiedzy przybliżonej, podjęto badania mające na celu ocenę przydatności takiego rozwiązania w dziedzinie projektowania i konstruowania maszyn. Realizacja badań związana była z procesem pozyskiwania, gromadzenia i stosowania wiedzy inżynierskiej z zastosowaniem sieci przekonań. Opracowując odpowiednią metodologię postępowania, wskazano sposób definiowania modelu projektowanego

obiektu. Zwrócono uwagę na konieczność podziału złożonego zadania na mniejsze podzadania, które można rozpatrywać niezależnie. Wskazano poszczególne elementy sieci przekonań, ich związki z opisem dziedziny zastosowania. Przedstawiono wpływ struktury sieci na jej działanie oraz możliwość jej modyfikacji. Omówiono proces łączenia poszczególnych podsieci umożliwiające budowę modelu globalnego. Na przykładzie wybranego zadania projektowego dotyczącego łożyskowania wałów maszynowych przeprowadzono weryfikację opracowanej metody. Na podstawie pozyskanych przykładów uczących przeprowadzono proces uczenia sieci. Tak opracowany model wykorzystano do wykonania systemu doradczego działającego w środowisku oprogramowania do komputerowo wspomaganego projektowania.



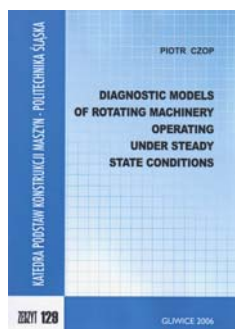
Jacek WOJTUSIK „Wielomodelne diagnostyczne maszyn wirnikowych” Zeszyt 128

W pracy przedstawiono wady i zalety wielomodeli diagnostycznych. Dokonano przeglądu i systematyki modeli stosowanych w diagnostyce technicznej wraz z omówieniem koncepcji rozpatrywania obiektu w diagnostyce technicznej w postaci „czarnej skrzynki”. Wprowadzono definicje wielomodeli i wyróżniono, ze względu na ich strukturę, dwie klasy wielomodeli: modele wielowarstwowe i modele wielostopniowe. Przedstawiono również pole możliwych rozwiązań metod wyznaczania jednego wyjścia modeli wielowarstwowych na podstawie wyjść modeli składowych. Omówiono metody agregacji i uzgadniania wyjść.

Zaprezentowano przykład metody identyfikacji wielomodelu diagnostycznego do rozpoznawania stanu technicznego maszyny wirnikowej w zakresie przemieszczeń podpór łożyskowych. Wielomodel będący fragmentem większego modelu, służył do rozpoznawania klasy położenia podpór łożyskowych na podstawie klas obciążenia łożysk. Stanowił drugi stopień modelu dwustopniowego, w którym w odróżnieniu od jednostopniowych klasyfikatorów uwzględniono dodatkową informację o obiekcie badań poprzez zdefiniowanie przestrzeni klas obciążenia łożysk, spinającą obydwa stopnie modelu. W opisywanych badaniach zastosowano model wielowarstwowy, w którym wyjście było wynikiem agregacji wyjść modeli składowych. Modelami składowymi były klasyfikatory bazujące na sieci neuronalnej, interpolacyjnych metodach odwrotnych, sieci

Bayesa oraz metodzie *fuzzy k-means*. Metodę tą zweryfikowano na podstawie danych uczących pochodzących z symulacji udostępnionych przez Instytut Maszyn Przepływowych Polskiej Akademii Nauk (IMP PAN) w Gdańsku.

Modele trenowane były na podstawie przykładów uczących, dotyczących zachowania obiektu w warunkach przemieszczenia jednej, wybranej podpory. Testowanie modeli i wielomodelu przeprowadzono na podstawie przykładów stanów złożonych, w których przemieszczane były dwie podpory jednocześnie. Na podstawie przeprowadzonych badań stwierdzono, że rozpoznawanie klas położenia podpór z zastosowaniem wielomodelu daje lepsze wyniki niż osobne stosowanie modeli składowych tego wielomodelu jako modeli globalnych.



Piotr CZOP
**"Diagnostic models of
 rotating machinery
 operating under steady
 state conditions"**
Zeszyt 129

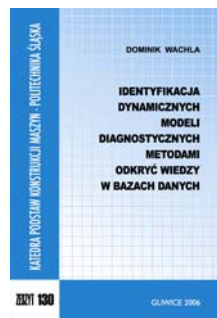
The present monograph deals with the development of techniques designed for the experimental vibration analysis of rotating machinery from the viewpoint of input and/or output system analysis. Such experimental vibration has been used to describe the dynamic behavior of the rotor supported by hydrodynamic bearings in terms of natural frequencies, damping factors and phase shifts. The main advantage of the implementation of automatic diagnostic tests and procedures into practice is an increase of a confidence level resulting in more suitable decisions being made automatically.

The aim of the monograph is to provide a class of diagnostic symptoms based on a parametric approach in the scope of parameter-based monitoring intended for rotating machinery. Such diagnostic symptoms may prove to be powerful tool for a new generation of decision-supporting systems.

The method applied in research is parametric system identification with use of AR, ARX models.

In the monograph several aspects of testing methods for rotating machines under steady state conditions have been discussed. A mathematical and physical model of the rotating machinery has been utilized to generate displacement responses under normal and malfunction conditions. Unbalance, preload, rub, excessive clearance, and oil instability were described as typical malfunctions of rotating machinery. On the base of the research it was stated that parameter-based symptoms can replace actually use non parametric-

based symptoms for diagnosis of rotating machinery under steady state operating conditions.



Dominik WACHLA
**„Identyfikacja
 dynamicznych modeli
 diagnostycznych metodami
 odkrywania wiedzy w bazach
 danych”.**
Zeszyt 130

Przedmiotem pracy jest metoda identyfikacji modeli autoregresyjnych opisujących

dynamikę procesów zachodzących i/lub realizowanych przez różne systemy techniczne. Metoda pozwala identyfikować modele procesów w sposób automatyczny z użyciem danych gromadzonych przez przemysłowe systemy SCADA. Jest oparta na wybranych metodach odkrywania wiedzy w bazach danych (ang. *Knowledge Discovery in Databases, KDD*). Jednym z głównych elementów metody jest transformacja (przekształcenie) danych procesowych będących szeregiem czasowym do postaci tzw. przykładów uczących, która jest wymagana przy prowadzeniu analiz metodami KDD. Drugi etap metody obejmuje identyfikację modelu procesu za pomocą wybranej metody KDD. Identyfikacja modelu jest połączona z jednoczesną selekcją zmiennych niezależnych. Wynik selekcji zależy od jakości zidentyfikowanego modelu, który oceniany jest za pomocą kryterium heurystycznego opartego na regułach selekcji modeli stosowanych w dziedzinie identyfikacji systemów.

Weryfikację metody przeprowadzono z użyciem danych wygenerowanych na podstawie modeli matematycznych hipotetycznych systemów dynamicznych oraz danych zgromadzonych przez przemysłowy system SCADA, który współpracuje z pompownią wód głębinowych. Do identyfikacji modeli obserwowanych procesów zastosowano metodę wektorów wspomagających (ang. *Support Vector Machines, SVM*), natomiast w procesie selekcji zmiennych niezależnych został użyty algorytm genetyczny. W pracy zamieszczono wybrane wyniki badań weryfikacyjnych opracowanej metody, które w sposób indukcyjny potwierdzają słuszność postawionych tez.

Recenzenci publikowanych prac:

prof. dr hab. inż. Lesław BĘDKOWSKI
dr hab. inż. Piotr BIELAWSKI, prof. AM
prof. dr hab. inż. Tadeusz BURCZYŃSKI
dr hab. inż. Tadeusz DĄBROWSKI, prof. WAT
dr hab. inż. Piotr FEDELIŃSKI, prof. PŚ
prof. dr hab. inż. Jerzy GIRTLER
prof. dr hab. inż. Jerzy JAŻWIŃSKI
dr hab. Henryk KAŻMIERCZAK
prof. dr hab. Witold KOSIŃSKI
dr hab. inż. Piotr KRZYWORZEKA

dr hab. inż. Wiesław MICZULSKI, prof. UZ
prof. dr hab. inż. Wojciech MOCZULSKI
dr hab. inż. Andrzej OBUCHOWICZ, prof. UZ
dr hab. inż. Andrzej PIECZYŃSKI, prof. UZ
prof. dr hab. inż. Janusz SZPYTKO
prof. dr hab. inż. Tadeusz UHL
dr hab. inż. Teodor WINKLER, prof. PŚ
dr hab. inż. Zygmunt WRÓBEL, prof. UŚ
prof. dr hab. inż. Bogdan ŻÓŁTOWSKI

Call for Papers

DAMAS 2007

**7th INTERNATIONAL
CONFERENCE
on
DAMAGE ASSESSMENT
OF STRUCTURES**

25th – 27th June 2007

Hosted by
Politecnico di Torino
at
**Centro Congressi
dell'Unione Industriale
Torino, Italy**

<http://www.polito.it/DAMAS-2007>





PHOTO D'ORTIZIO



**21st Canadian Congress of
Applied Mechanics**

RYERSON UNIVERSITY
Department of Mechanical and
Industrial Engineering
Toronto, Canada
June 3-7, 2007

CALL FOR PAPERS

The Canadian Congress of Applied Mechanics (CANCAM) is a biennial conference with a forty-year history. It is a premiere Canadian conference that attracts researchers in mechanical engineering, applied mathematics, civil engineering, and aerospace engineering from around the world.

Format of the Conference
Papers accepted for presentation will be published in the Congress Proceedings. The official languages of CANCAM07 are English and French. Short papers (two-pages) will be accepted in both languages.

Goals

- To provide an international forum for communicating the most recent advances in the field of applied mechanics.
- To provide a forum for information exchange among delegates from industry, research labs, and academia.

Keynote Speakers
A. Bejan (Duke) - Thermofluids
J. Hansen (UTIAS) - Structures
R. Sauve (AECL) - Nonlinear Mechanics
M. Pettigrew (École Polytechnique Montréal) - Flow-Induced Vibration

Organizing Committee
Honorary Congress Chair:
W. L. Cleghorn (University of Toronto)
Congress Chair: D. Naylor (Ryerson University)
Congress Co-Chair: S. D. Yu (Ryerson University)
Conference Website: www.ryerson.ca/cancam07

Important Dates:
Short Paper (2 page) Submission: January 15, 2007
Notification of Acceptance: March 1, 2007

Wszystkie opublikowane w czasopiśmie artykuły uzyskały pozytywne recenzje, wykonane przez niezależnych recenzentów.

Redakcja zastrzega sobie prawo korekty nadesłanych artykułów.

Kolejność umieszczenia prac w czasopiśmie zależy od terminu ich nadesłania i otrzymania ostatecznej, pozytywnej recenzji.

Wytyczne do publikowania w DIAGNOSTYCE można znaleźć na stronie internetowej:

<http://www.uwm.edu.pl/wnt/diagnostyka>

Redakcja informuje, że istnieje możliwość zamieszczania w DIAGNOSTYCE ogłoszeń i reklam.

Jednocześnie prosimy czytelników o nadsyłanie uwag i propozycji dotyczących formy i treści naszego czasopisma.

Zachęcamy również wszystkich do czynnego udziału w jego kształtowaniu poprzez nadsyłanie własnych opracowań związanych z problematyką diagnostyki technicznej. Zwracamy się z prośbą o nadsyłanie informacji o wydanych własnych pracach nt. diagnostyki technicznej oraz innych pracach wartych przeczytania, dostępnych zarówno w kraju jak i zagranicą.