MODELS IN DIAGNOSTICS OF INDUSTRIAL PROCESSES

Jan Maciej KOŚCIELNY, Michał BARTYŚ

(Politechnika Warszawska, Instytut Automatyki i Robotyki, ul. A. Boboli 8, 02-525 Warszawa,
(e-mail: jmk@mchtr.pw.edu.pl, bartys@mchtr.pw.edu.pl)

Summary
In the paper, an overview and classification of modelling approaches used in diagnostics of industrial processes is presented. There are characterised the models used for fault detection and fault isolation. Main features of characterised models are underlined and particular attention was paid on the model practicability. Special attention was paid also on models making use of artificial intelligence and expert knowledge.

Keywords: Modelling, diagnostics of industrial processes, fault detection, fault isolation, artificial intelligence, soft computing

MODELE W DIAGNOSTYCE PROCESÓW PRZEMYSŁOWYCH

Streszczenie
W referacie przedstawiono przegląd i klasyfikację metod modelowania stosowanych w diagnostyce procesów przemysłowych. Przedstawiono zarówno modele wykorzystywane w procesie detekcji jak i lokalizacji uszkodzeń. Podano najważniejsze właściwości modeli, zwracając szczególną uwagę na aspekt ich zastosowań praktycznych. Zwrócono także uwagę na istotny walor aplikacyjności modeli bazujących na metodach sztucznej inteligencji i wiedzy eksperckiej.

Słowa kluczowe: Modelowanie, diagnostyka procesów przemysłowych, detekcja uszkodzń, lokalizacja uszkodzeń, sztuczna inteligencja

1. INTRODUCTION

Diagnostics of industrial processes is a part of diagnostics particularly dealing with the on-line, real-time, fault detection, isolation and identification of process faults. Therefore, the industrial diagnostics is focused on components of technological installation, its equipment (actuators and instrumentation) as well as on the process itself. Diagram of the scheme of diagnostics of industrial processes is given on Fig. 1. Installations in chemical, petrochemical, power, metal, pharmaceutical, food, paper, gas, oil industries, as well as pipelines, rockets, pumps, engines turbines etc. are typical examples of diagnosed systems.

Generally, three phases [2,5,6,9] are distinguished in the industrial diagnostics: detection, isolation and identification. In fact, the identification phase is often omitted or sometimes is integrated with isolation phase. Typically, diagnostic process is carried out in two phases: detection and isolation (FDI). Fault symptoms are formulated in detection phase. Fault symptoms are generated by means of system models or system knowledge and process data. Process faults are pointed out in the isolation phase based on patterns of detected symptoms.

Fault detection may be carried out with application of system models or without the models. In the first case, fault detection consists of model based residual generation, residual evaluation and finally on decision-making procedures classifying detected fault symptoms. General model based diagnostic system scheme is given on Fig. 2.

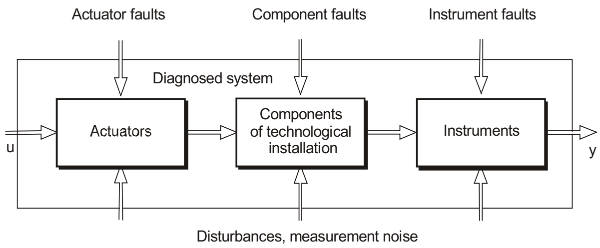

Fig.1. Diagram of diagnostics of industrial systems

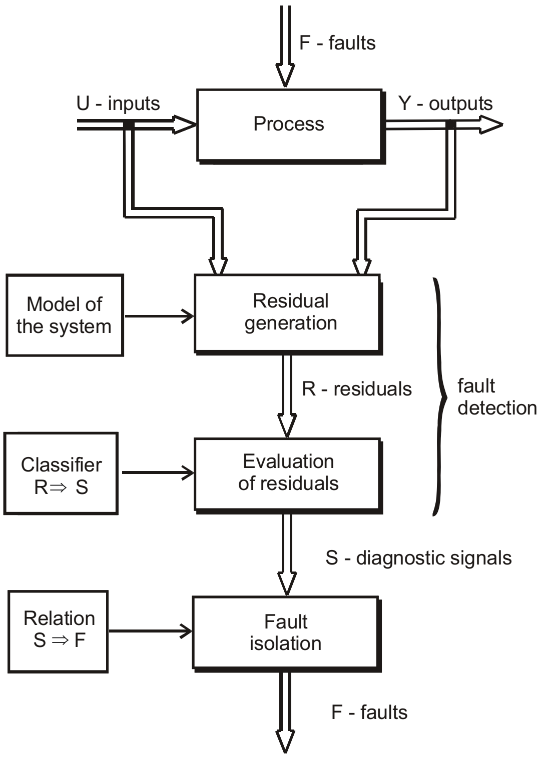

Fig.2. Model based diagnostics

In case of lack of full model or in case of extraordinary model complexity, the limits control, signal analysis or checking of relations between process variables are used for fault detection purposes. However, fault diagnostics of industrial processes based on models takes substantial advantages in comparison to the diagnostics based on methods that do not need the models. Model based fault diagnostics allows early fault detection, gives more accurate fault isolation and detection and allows the recognition of small sized faults. On the other hand, model based fault detection procedures are time consuming particularly in preparation phase and need more computational power during exploitation.

2. GENERAL MODEL OF DIAGNOSED FAULTY SYSTEM

Full description of dynamic system with respect to faults and disturbances [5,6] is as follows:

$$\dot{\mathbf{x}}(t) = \phi[\mathbf{x}(t), \mathbf{u}(t), \mathbf{d}(t), \mathbf{f}(t)] \tag{1}$$

$$\mathbf{y}(t) = \psi[\mathbf{x}(t), \mathbf{u}(t), \mathbf{d}(t), \mathbf{f}(t)] \tag{2}$$

where:

u - signals representing external effects on system. Those signals are called forced inputs or inputs or system excitations. In these group one can distinguish: control signals and known inputs.

$$\mathbf{u}(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ \dots \\ u_p(t) \end{bmatrix} \tag{3}$$

y – output signals representing effects of system acting on environment. Those actions may be interpreted as system responses.

x – state space co-ordinates

$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \\ \dots \\ y_q(t) \end{bmatrix} \tag{4}$$

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \dots \\ x_n(t) \end{bmatrix} \tag{5}$$

State **x**(t) of the system in any time moment t depends on the system state $\mathbf{x}(t_0)$ in the initial moment t_0 and trajectory (history) of input $\mathbf{u}(t_0,t)$ in the time interval (t_0,t). Thus, state vector reflects the portion of information from the past necessary to calculate the actual state change and system output.
d – disturbations, that belong to the subset of input signals with unknown values
f – faults, that belong to the separate subset of inputs having destructive action on system behaviour. Faults may appear as abrupt or incipient.

$$\mathbf{f}(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \\ \dots \\ f_k(t) \end{bmatrix} \tag{6}$$

The scheme of the system with respect to influence of disturbances and faults is presented on Fig. 3.

Let us assume, that the technical state of the diagnosed system $z(t)$ is the function of faults [5,6]:

$$z(t) = z[\mathbf{f}(t)] \tag{7}$$

If we will be able to determine the fault vector **f** based on state space equations (1) and system outputs equations (2) assuming lack of disturbances, then the problem of system diagnostics may be solved.

$$\mathbf{f}(t) = \Psi[\mathbf{y}(t), \mathbf{x}(t), \mathbf{u}(t),]; \quad \mathbf{d} = 0, \tag{8}$$

However, even if equations (1) and (2) are known, not ever is possible to find out the inverse model (8). The equation (8) very often has the entangled form and in reality, the number of faults is greater then the number of equations describing the system.

Models applied in diagnostics of industrial processes are classified into two groups: models used for fault detection and models used for fault isolation. Models used for fault detection describe the relations between the system inputs and outputs $U \Rightarrow Y$ (mainly in the normal system state, without faults) and allow detection of the changes (symptoms), caused by faults. Models used for fault isolation define relations between the diagnostic signals (symptoms) and faults $S \Rightarrow F$.

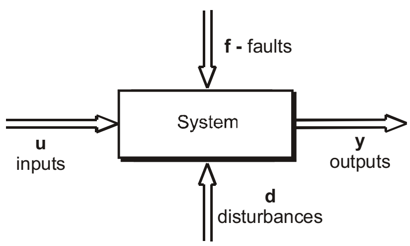

Fig.3. Scheme of diagnosed system

3. MODELS FOR FAULT DETECTION

Fault detection is a process of generation of diagnostic signals *S* based on process variables *X*. Diagnostic signals should reflect the information about the faults, so that detection may be defined as a process of mapping the space of process variables *X* onto the space of diagnostic signals *S* combined with evaluation of diagnostic signals.

Fault detection process consists of two parts (Fig.4). In the first part, residual values are calculated on the basis of the system model, while in second part the evaluation of achieved residual is carried out combined with generation of diagnostic signals. Therefore, fault detection phase need appropriate models.

Typically, models describe the system in normal state (without faults). Therefore, in case of fault occurrence, it is possible to determine the discrepancies between the current system behaviour and expected behaviour in normal state. Residuals can be calculated as:

- the difference between the value of process variable and value calculated from the model (Fig. 4).
- the difference between the left-hand and right-hand of equation describing the system
- the difference between the nominal and estimated values of the parameter of the system

Residual values should be as close to zero as possible in the fault-free state of the system. Residual values significantly different from zero values are defined as the fault symptoms.

A set of models (so called partial models) is used for fault detection of complex systems. These models should “cover” all the system i.e. there should not be lack of unconnected outputs and inputs of all partial models. Diagnostics based on partial models takes advantages over the diagnostics based on global ones [7]. There is possibility to achieve simple models, shorter fault detection times, lower commissioning costs, better flexibility of diagnostic system. Obtaining of analytical, fuzzy or neural models is not possible in some cases. This results mostly from the physical unavailability of input signals in industrial installations. Therefore, it is not possible to obtain also outputs of the models.

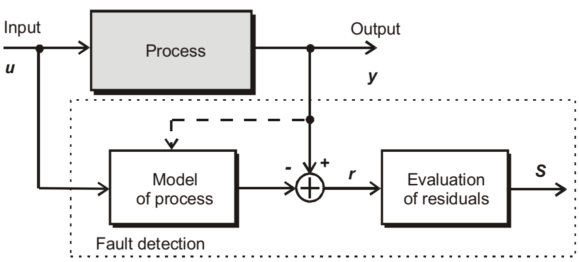

Fig.4. The scheme of fault detection based on models

3.1. Fault detection with application of analytical models

In the group of analytical detection methods one can distinguish [2,3,4,5,6,7,8,9]:

- detection with the use of physical models (e.g., balance models, dynamic equations, etc.),
- detection with the use of linear input-output types of models (parity equations),
- detection with the use of state observers or Kalman filters,
- detection based on on-line identification.

<u>Residual generation based on physical models</u>

Generally, a most complete model of a system can be directly obtained from physical equations [1]. Non-linear static systems are described by following equation:

$$\Psi(y, \mathbf{u}) = 0 \tag{9}$$

This equation describes the system in the state of full aptitude. The relationship given above is not true in case of faults. Therefore, the residual value different from zero may be symptom of fault. Residual can be calculated as:

$$r = \Psi(y, \mathbf{u}) \tag{10}$$

Residual generation based on general non-linear models give an assumption to achieve the most reliable and robust detection technique. However, results are highly dependent on the model accuracy. Models based on physical laws describe most completely relationships existing between process variables. Those models reflect static and dynamic system properties in the whole space of states. Therefore, they enable to detect faults that have also small sizes. Development of models based on physical relationships is extremely difficult or outright impossible in many cases. Moreover, identification of model parameters brings additional difficulties. Thus, application of this method is limited to the systems that are described by relatively simple equations.

<u>Residual generation based on linear models</u>

Linear dynamic systems are commonly described by means of transfer function defined as a ratio of Laplace transformation of output signal $y(s)$ to the Laplace transformation of input signal $u(s)$ by zeroed initial conditions.

$$G(s) = y(s)/u(s) \tag{11}$$

Any *j*-th output is given by:

$$y_j(s) = \mathbf{G}_j(s)\mathbf{u}(s) = G_{j1}(s)u_1(s) + G_{j2}(s)u_2(s) + \ldots + G_{jp}(s)u_p(s) \tag{12}$$

Methods of residual generation and structurisation based on transfer function were developed by Gertler [3]. Descriptions of those methods are given also in [5,6,7,8,9].

State space equations approach is also applied to description of linear systems. Dynamic stationary linear system with *p* inputs and *q* outputs may be described by state and output equations in continuous time.

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \quad (13)$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$$

Method of residual generation based on the state space equations was developed by Chow and Willsky and is presented in many publications, for example in [5,6,8,9].

Observers estimate dynamic state of the system based on the input and output signals. The equation of full observer is following:

$$\hat{\mathbf{x}}(k+1) = \mathbf{A}\hat{\mathbf{x}}(k) + \mathbf{B}\mathbf{u}(k) + \mathbf{H}[\mathbf{y}(k) - \mathbf{C}\hat{\mathbf{x}}(k)] \quad (14)$$
$$\hat{\mathbf{y}}(k) = \mathbf{C}\hat{\mathbf{x}}(k)$$

where: $\hat{\mathbf{x}}$ - state estimate, $\hat{\mathbf{y}}$ - output estimate, $\mathbf{H}$ – observer feedback matrix.

In this case residual vector is given by:

$$\mathbf{r}(k) = \mathbf{y}(k) - \mathbf{C}\hat{\mathbf{x}}(k) \quad (15)$$

Clark, Frank and Patton [2,8,9] developed the methods of residual generation based on Luenberg observer.

Linear models are valid only in the near surrounding of the system nominal operating point i.e. point where parameters of the model of the system were identified. Hence, any change of the operating point may cause the appearance of non-zero residual values, similar to that caused by faults. On the other hand, linear models enable early detection of the small parametric faults. It is however obtained at the cost of necessity of definition of sufficiently precise models what is often very difficult. The residuals have to be adequately sensitive to faults but, on the other hand, they should be sufficiently insensitive to other changes such as natural disturbances existing in the process, measurement noises or modelling errors. This is a main reason for what linear models have been found only limited applications in practical implementations.

Residual generation based on on-line identification

Faults appear not only as changes of values of the system outputs but also as changes of physical coefficients $\boldsymbol{p}$ from system dynamic equations, such as: resistances, capacitance, rigidities, etc. These physical coefficients are main constituents of the parameters $\boldsymbol{\theta}$ of the model of the system. If one determines values of the coefficients based on identification of the system and compares them with its nominal values, i.e., parameter values in the state of full aptitude of the system, then the obtained differences are residuals containing information on faults $\boldsymbol{r}=\boldsymbol{p}_N-\boldsymbol{p}$. Such detection method has been developed by Isermann [4].

Model parameters are understood as constants that are appearing in the mathematical description of relations between the inputs and outputs of the system. One can distinguish system static models:

$$y = \beta_0\Phi_0 + \beta_1\Phi_1(\mathbf{u}) + \beta_2\Phi_2(\mathbf{u}) + \ldots \quad (16)$$

where: $\Phi_0 \equiv 1$, $\Phi_i(\mathbf{u})$ - known function of input vector, (e.g.. $\Phi_i(\mathbf{u}) = u_1 u_2$, $\Phi_i(\mathbf{u}) = u_1^2$)

as well as dynamic models given by set of differential equations linearised in the neighbourhood of the operation point.

$$\frac{d^n y}{dt^n} + a_{n-1}\frac{d^{n-1}y}{dt^{n-1}} + \ldots + a_1\frac{dy}{dt} + a_0 y \quad (17)$$
$$= b_m\frac{d^m u}{dt^m} + b_{m-1}\frac{d^{m-1}u}{dt^{m-1}} + \ldots + b_1\frac{du}{dt} + b_0 x;\ (n \ge m)$$

Model parameters $\boldsymbol{\theta}^T = [\beta_0, \beta_1, \beta_2, \ldots]$ or $\boldsymbol{\theta}^T = [a_{n-1}, \ldots, a_1, a_0; b_m, b_{m-1}, \ldots, b_1, b_0]$ are more or less complex functions of physical coefficients. If the method leads to good results, it is necessary to obtain an adequate model of the system with the help of theoretical modelling (based on physical and chemical laws) as well as reliable identification of the system model's parameters. The identification requires having an adequate excitation of the system, so that signals captured by the identification reflect whole range of its changes during regular system operation. Methods based on the system parameter estimation are suitable particularly for the well-defined systems such as mechanical, electrical and electro-mechanical systems. These methods are rarely applied to heat and chemical processes because of difficulties in obtaining appropriate and sufficient good models. Another disadvantage of the method is the need of engaging of huge computational power, as well as problems with detection of additive faults.

3.2. Fault detection with application of neural and fuzzy models

In many practical cases non-linear analytical models are unknown, and moreover the linear models are impractical because of strong system non-linearity. In such cases, the application of neural and fuzzy models is considered. This is implied by the substantial advantage of neural and fuzzy techniques of modelling non-linear systems and ability of learning from the data samples. Additionally, in case of fuzzy models, the expert knowledge may be inputted into the model. In the automated industrial processes, a sets of current and archive values of process variables are available. This makes the frames of building the system models and make use of availability of current process data and the expert knowledge about the relations between the process variables. Simultaneously, the development of computing technologies have broken the essential limitations related to necessity of absorbing high computational power for model tuning and handling of huge data sets.

In last decade, intensive development of the neural network structures has been profiting in numerous of applications in fault detection domain. Multi-layer unidirectional static and dynamic neural networks, radial and GMDH as well as feedback network structures are intensively investigated [5,6]. Neural networks may be interpreted as "black box" entities. Particular network components and its weights or activation functions haven't any physical relation with the modelled system structure and parameters. The expert knowledge in this case may be useful only in limited sense, to define the sets of model inputs and outputs only. Neural networks take advantage of ability of generalisation of network knowledge and its robustness against disturbances.

More often, the knowledge about the diagnosed system is imprecise. Typically, this knowledge is available in the form of implication rules (*if-then*) containing the linguistic evaluation of process variables (for example: high temperature, low level). In this case the fuzzy models are to be considered. These models are based on theory of fuzzy sets introduced by Zadeh. Typically, the models make use of inference scheme given by Mamdani. Fuzzy-neural networks being the conjunction of the fuzzy modelling techniques [5,6,7] and methods of training of neural networks are very convenient for modelling for residual generation purposes. Fuzzy-neural networks enable to make use of expert knowledge for defining the number of *if-then* rules, give hints for planning the initial shapes of fuzzy membership functions and selection the data for the network learning. Expert knowledge is very useful in the phase of model structure building and in phase of setting initial values of model parameters. Created model is no more the "black box". It is a set of rules that may be interpreted and verified by experts. The number of rules grows rapidly with the increase of number of inputs and outputs and number of fuzzy partitions assigned to particular inputs. This causes limitations of applying fuzzy-neural networks for the relative simple systems. Technique of input aggregation [5,6,7] is very helpful here. This technique leads to the reduction of the number of inputs by replacing the subsets of model inputs by the appropriate chosen functions of those inputs. Inputs aggregation may be carried out simultaneously for a few subsets of inputs. Neural and fuzzy models created and tuned on experimental data are able to model the system in the limited space of inputs and outputs determined by spans of available data. Those models are more practicable when comparing with linear models, because of better model behaviour also far from operating point. However, better and wider system description provide non-linear models based on physical laws describing the phenomena taking place in the system.

4. MODELS FOR FAULT ISOLATION

Fault isolation rarely directly makes use of residual values. Typically, prior to fault isolation, the residual values are appropriately pre-processed. Classifier $R \Rightarrow S$ transforming the continuous residual signals into the binary or multiple-valued diagnostic signals is typically used for this purpose. The subsets of diagnostic signals are called fault symptoms. Fault symptoms are only those diagnostic signals that reflect the fault occurrence. The models for fault isolation should map the space of fault symptoms onto the space of discrete faults or system states. One can easily see that relation $S \Rightarrow F$ define the inverse causal model i.e. model of type: effect-cause.

Following models classes are useful for fault isolation:

a) Models that map the space of binary symptoms onto the space of faults or the system states. Binary diagnostic matrix, binary graphs and diagnostic trees, rules, set of rules, logic functions e.t.c. [5,6,7] are belonging to this class.

b) Models that map the space of multiple-valued fault symptoms into the space of faults or the system states. Information system and rules for multi-valued fault symptoms belong to this class.

c) Models that map the space of residuum into the space of faults or system states. Specific musters of residuals or diagnostic signals are assigned to each fault or system state. Classical methods of pattern recognition, neural networks and fuzzy-neural networks are used for modelling and fault isolation in this case.

Models for fault isolation may be defined by application of:

- learning
- knowledge of redundancies in the system structure
- modelling of influence of faults on the residual values
- expert knowledge.

In first case, there is necessary to have the learning data sets from the state of system full aptitude and the data from all states with faults or at least data for defined classes of states. Obtaining of such data sets is difficult and often impossible in case of diagnostics of industrial processes.

Relation fault-symptoms are easy to obtain in case of redundant systems of type K from N. But, because of economical reasons, such redundant systems are applied relatively seldom.

If equations of residual generation making allowance for fault influence are known, then as a result of fault simulation it is possible to determine the spans of residual values and diagnostic signals for the states with, and without faults. In this case there are available sets of diagnostic signals and symptoms assigned to each fault and state of system full aptitude. This approach is very rational, however it is also relatively complex and labour consuming. Mainly, it is related to the difficulties with obtaining the mathematical description of the system making allowance for influence of faults.

Expert knowledge may be also very useful when building the models for fault isolation. An expert should define diagnostic signal values that are related to particular faults.

The method of acquiring knowledge necessary for fault isolation depends on the diagnosed system specificity. For instance for unique, one-of-the kind systems such as chemical plants, acquisition of learning data for states with faults is not possible in practice. Particular faults appear rarely, their potential number is very high, and the diagnostic system should recognise their first appearance. For complex chemical plants, it is very difficult to work out analytical models that take into account an influence of faults on residual values. Application of the expert's knowledge remains therefore the only way. However, in a case of serially manufactured systems such as for example turbines the physical models and very complex analytical models are mostly known.

Learning data from the states with faults can be obtained based either on physical or analytical models. For serially manufactured systems, it is possible to collect data from measurements carried out in states with faults. In such cases, artificial introduction of faults and even system destruction investigations are applied.

5. FINAL REMARKS

Diagnostic of industrial processes should be carried out in real-time in the system exploitation phase. This implies the application of specific approaches. Typically, diagnostic of industrial processes is composed from two main phases: fault detection and fault isolation. Diagnostic of industrial processes is based mainly on signals available from the process. It is not allowed to disturb the process by introduction of specific excitations into the process. Therefore, most algorithms of diagnostics of industrial processes are based on relations between the process variables. Those relations may have a form of analytical, neural and fuzzy models. More or less, this technical domain is based on methods worked out on the grounds of control theory (modelling and identification of dynamic systems). This is a little bit another approach in comparison to those, based on signal analysis intensively applied for diagnostics of machines.

Diagnostic of industrial processes makes use of methods of artificial intelligence by application of neural networks, fuzzy logic and genetic algorithms. Neural or fuzzy models are build-up by means of archive data available from the industrial process control and supervising systems (DCS, SCADA).

This paper was particularly intended to give an overview of the models used in the diagnostic of industrial processes. Two groups of models were characterised: models for fault detection and models for fault isolation. Classification of models given in paper does not include all the models used in different diagnostic approaches. For example, models of system normal state and models of states with different faults are used [5,6]. Variety of models used in diagnostics of industrial processes reflects different degree of knowledge about the diagnosed system, and different ways of acquiring this knowledge.

REFERENCES

[1] Cannon R.H. (1973). Dynamika układów fizycznych. WNT, Warszawa.
[2] Chen J., Patton R.J. (1999). Robust model based fault diagnosis for dynamic systems. Kluver Academic Publishers, Boston.
[3] Gertler J. (1998). Fault Detection and Diagnosis in Engineering Systems. Marcel Dekker, Inc. New York - Basel - Hong Kong.
[4] Isermann R. (1984). Process fault detection based on modeling and estimation. Methods - a survey. Automatica, 20(4), 387-404.
[5] Korbicz J., Kościelny J.M., Kowalczuk Z., Cholewa W., (2002). Diagnostyka procesów. Modele, metody sztucznej inteligencji, zastosowania. WNT, Warszawa,(1-828),
[6] Korbicz J., Kościelny J.M., Kowalczuk Z., Cholewa W., (2004).Fault Diagnosis. Models, artificial intelligence, application. Springer
[7] Kościelny J.M. (2001). Diagnostyka zautomatyzowanych procesów przemysłowych. Akademicka Oficyna Wydawnicza Exit, Warszawa.
[8] Patton R., Frank P., Clark R. (Eds.) (1989). Fault diagnosis in dynamic systems. Theory and Applications. Prentice Hall, Englewood Cliffs, New York.
[9] Patton R., Frank P., Clark R. (Eds.) (2000). Issues of fault diagnosis for dynamic systems. Springer.

Jan Maciej KOŚCIELNY, prof. Ph D Hab. His research interests are concerning fault diagnosis in mechatronics systems and industrial processes as well as the investigation of fault tolerant control systems. He is the author or co-author of more then 160 papers and 4 books. His engineering activity (35 industrial contracts) are focused on the application on automatic control and fault diagnosis in food, chemical power and automatic control industries.

Michał BARTYŚ, Ph. D. His research interests are concerning automatic control, diagnostics and fuzzy logic. He is the author or co-author of more then 80 papers, 2 books, 4 patents. His engineering activities (41 implementations) are focused on development of smart instrumentation, actuators and control systems.