FACTA UNIVERSITATIS (NIŠ)
Series: Electronics and Energetics vol.13, No.2, August 2000, 143–155

# THE IMPACT OF PROCESS MEASUREMENT ON INDUSTRIAL DIAGNOSTICS

**Gordana Bojković and Vera Bajović**

**Abstract.** There is one important term in the intersection of process measurement and industrial diagnostics: automatization of both processes by application of computers. This paper describes the role and significance of measurement for the choice of corresponding diagnostic method as well as for the final diagnostic efficiency. Assumption is that both, measurements and diagnosis, are the part of computer based system for automated control and regulation of industrial processes.

The fact is that the lack of *a-priory* diagnostic information is most often caused by poor or inadequate measurement data. New methods that must have been developed for global industrial diagnostics in such situations can also be successfully implemented in fault diagnostics of measurement instrumentation and systems. There is an obvious feedback from diagnostics to measurement. We can also speak about the impact of diagnostics improvement on improvement of measurement. New diagnostic methods developed for not well instrumented processes (mostly knowledge based methods), give especialy good results when implemented in systems with, from aspect of diagnosis, enough and adequate measurements. In this paper, we also give some of our results in field of diagnostics in order to support previous assertions.

**Key words:** Industrial diagnostics, automatization, proces measurement.

## 1. Introduction

A computer based system for automated supervision, control and regulation of industrial processes, often refereed as information and management systems (IMS), could be decomposed to following functional components:

- Data Acquisition (DAQ)

Manuscript received March 24, 2000.

Authors are with University of Novi Sad, Faculty of Engineering Sciences, Fruškogorska 11, 21000 Novi Sad (e-mails: [gocab, bajovic]@uns.ns.ac.yu).

- Simulation Model (SM)
- Computer Aided Control (CAC)

Process measurements are the important part of DAQ, while diagnostics is the part of CAC system and is closely related to it. Each lower level component of the IMS is the base for the higher level one. If enough measurements are installed in DAQ, then it is possible to generate a good SM, and if a good SM and enough measurement data are disposable, then it is possible to make good CAC and, within its limits, a satisfactory diagnostic system.

Fault diagnosis of industrial and other technological systems, known as technical or more often industrial diagnostics, is the problem of determining the causes of the unusual manifestations from a set of observable. These so-called "set of observable" is composed of test and measurement results of physical quantities relevant to diagnosis and could be regarded as a set of *basic diagnostic values* (BDV). Adequate definition of industrial diagnostics is given in reference [1]: "Industrial diagnostics is the interpretation of information gleaned from sensors in order to assess the condition of a process or to provide an estimate variable".

There are four steps in implementation of any industrial diagnostic strategy, [1], and these are:

- Selection of appropriate sensors ,
- Extraction of features from the data acquired by sensors,
- Comparison of features with defined standards and
- Determination of reliable decision process

The DAQ component of IMS provides conditions for the first stage of diagnostic scheme - selection of appropriate sensors. Trough selection of appropriate sensors, actually, a set of BDV is formed. If there are enough adequate measurements and the domain's structure, functional relationships and all required parameters are completely known, then it is possible to make diagnostic algorithm, i.e., to use classical, algorithmic approach to diagnosis.

However, the most common problem, which arises in practice, is that the decision to implement a diagnostic scheme is taken when the plant is already operational. Measurement sensors, which are already installed for operational purposes, are often unsuitable for diagnostics or they are not installed at the best location. In this case the knowledge-based approach to both development and implementation of a diagnostic procedure is the only solution.

This paper addresses the impact of process measurement on industrial diagnostics. More about the choice of diagnostic method, which also related, among other things, to the data provided by measurement, is discussed in the following section. Then, some aspects of knowledge based diagnostics will be presented and, some of our results.

## 2. Data Acquisition and Diagnostic Method Choosing

In Figure 1, a proposal for the classification of automated, computer-aided diagnostics is presented. The classification is made according to the applied methodology. A detailed analysis of the problem is given in [2] and [3]. The analysis shows that, which approach, i.e., which methodology would be implemented depends primarily on available knowledge bases and databases. At the other side, <u>both mentioned bases mainly rely on measurement data</u>.

As it is said in Introduction, there are two basic groups of diagnostic approaches, i.e., two basic classes of diagnosis: *algorithmic* and *knowledge* based. Both approaches use more or less all of four diagnostic steps mentioned before.

### 2.1 Algorithmic diagnostics (AD)

In *algorithmic* approach, it is possible to write an algorithm for computer program, which supports diagnostics. <u>All necessary diagnostic values (BDV) provided by measurements and relationships between them must be known</u>. It is recommended to use AD whenever it is possible. The main reason for that is it's great certainty, called "algorithmic certainty". AD can be roughly divided in two groups, denoted as *signal-based and model-based* in Figure 1.

*Signal-based* approach consists of variety of methods (estimation, filtering, statistical methods, analysis of alarm and trend of BDV etc.). What all these methods have in common, is the possibility to be taken under algorithm. <u>The main prerequisite for most of them is the abundance of measurements</u>.

An essential prerequisite for *model-based* AD is an *early process fault detection.* Whereas previously diagnostic methods (denoted as others) permitted recognition only when limit values of BDV had already transgressed, *model-based approach* is used for detecting the faults earlier, even while BDV

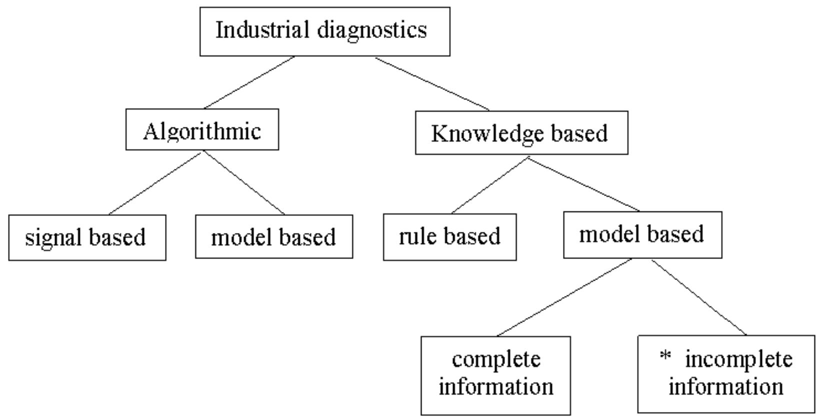


*Fig. 1. One proposal for classification of automated diagnostics*

are in permitted limits. This is possible for a range of process faults by the application of *process models* and *signal models* using process computers.

At this point we suppose that it is necessary to give some explanations in connection with the term "model" in diagnostic meaning. This is necessary because the model based diagnostics could be both, algorithmic and knowledge based. A detailed theoretical treatment is presented in [4]. Starting assumptions are the same for both basic types of diagnostics. A general structure of model based diagnostics, according to [4], is based on three types of models:

(i) a model of *normal* process, which assumes the state without faults;

(ii) a model of *observed process*, whose parameters are real data;

(iii) many models of the *faulty* process, where each represents the state with the known fault.

All of these models are based on *a priori* knowledge and experience of the real process. The values of the observed process are compared to the values of the *normal process*. This responds to the third step of diagnostic scheme ("Comparison of features with defined standards"), mentioned in Introduction. The values of generated differences, known as *error signals or residuals*, are compared to the values of *faulty* model in order to *detect* the faults. The models of the faulty process show the effects of the faults on the analyzed quantities. These effects are called *fault signatures*. By comparing residuals to the fault signatures, which responds to the fourth step of diagnostic scheme ("Determination of reliable decision process"), the task known as fault *diagnosis*, i.e., determination of the *fault location*, fault *size* and *cause of the fault* is performed.

In algorithmic approach of diagnosis, all of three mentioned models must be analytically precise, i.e., they must be *mathematical* or *quantitative*. The quantities, which are observed, are mainly measurable quantities - BDV. In other words, *model-based* AD is possible only in well defined processes with enough measurement data end enough causal knowledge about the process. If there isn't possibility for AD, second type of automated diagnostics must be used - knowledge based diagnostics. It should be emphasized here that important aspiration of any knowledge based diagnostic system is to yield algorithmic certainty.

### 2.2 Knowledge based diagnostic (KBD)

*Knowledge based diagnostic systems*, as knowledge based (expert) system at all, attempt to solve those classes of the problems which are not naturally amenable to numerical representation or which can be more efficiently represented by heuristics. There is two basic type of KBD: *rule based and model-based.* This division is made mainly according to the type of knowledge they use.

Rule based diagnostic system use simple production rules to provide a mapping between the possible causes and inputs (often only BDV, provided by DAQ) of a system and the possible faults. The basic proposition for such system is that a rule exists to conclude each possible fault or malfunction of the system. In algorithmic programming rules can appear in the form of conditional statements, typically IF...THEN... statements. They are executed in a sequence predefined by algorithm. Rules in knowledge based systems also have IF...THEN...syntax. However, there is a significant difference. They can't be executed in a defined order. Based on the matching the data and the IF-part of the rules, the *inference engine* determines which rule to be executed next. In the phase of knowledge acquisition, the "rules" are elicited from domain experts. The programmer's (knowledge engineer's) responsibility is: firstly, to provide a consistent and structured set of rules for the current problem and secondly, to provide a software, called *inference engine*, which controls the rule chaining sequence and, in essence, represents the solution model. The knowledge in the rule base is essentially a compilation of experiences, restricted to a given process, and is called compiled or more precisely, *experiential* knowledge. The fact is that experience is based on tracing relations between BDV and malfunctions of the system. Therefore, —undrebarexperience depends largely on available measurements. *Rule based* diagnostic systems are characterized by great efficiency.

For large, complex applications, many problems arise according to rule based system. For example, there is a problem of knowledge acquisition. It encompasses both: eliciting enough information from the expert and classifying it once it has been acquired. Even though the basic theory of a domain may be known, it is difficult to structure the heuristics which an expert uses in practice. This is valid especially for cases that rarely occur (known as *unusual* malfunctions), or for cases that have not occurred yet, and the expert may not have a clear structure for his heuristic knowledge. There is a need for so called *deep, fundamental or causal* knowledge. Often, this type of knowledge is implicitly contained in simulation models.

Models, whether behavioral, functional, or causal, are among the central mechanisms for organizing more powerful diagnostic systems, *model-based* systems. Supposing that the domain can be characterized as a system with distinct inputs and outputs, then relations, that can describe the model based diagnostic procedure, are:

$$\{\text{inputs}\} \Rightarrow (\text{expected state of the systems} \Rightarrow \{\text{expected outputs}\} \tag{1}$$

and

$$\{\text{expected outputs}\} \Rightarrow \{\text{outputs}\} \Rightarrow \text{state of the systems} \tag{2}$$

Relations (1) are generated from the real process and from the SM of normal process. Relations (2) are related to diagnostic decision process and are connected to the third and fourth step of diagnostic scheme, given in Introduction.

If the state of the system can always be derived from expected outputs, i.e., if both relations, relation (1) and relation (2) exist, the system is denoted as the system with *complete diagnostic information.* These are systems which can have some heuristic knowledge, and which have complete fundamental diagnostic knowledge. Diagnostic system of those types can also have the form of *rule-based* expert system, but the knowledge base is differently made. Very important feature of these KBD systems is that they could attain both *algorithmic certainty and efficiency of rule-based* systems'.

If the state of the system can not always (or can never) be derived from expected outputs, then it is the matter of the system with *incomplete* *diagnostic information*. The main problem, which must be solved in such situation, is the problem of knowledge acquisition, or, more precisely, the problem of extracting relevant diagnostic information from available data.

In accordance with above mentioned, model based KBD systems are devided on systems with *complete diagnostic information* and systems with *incomplete diagnostic information*.

The last are denoted with star ( "*") on the Figure 1 and are regarded as the most complicated diagnostic tasks at all. The methods for knowledge extraction, structuring and presentation in the form appropriate for encoding in diagnostic system, that must have been developed for this type of diagnostics, could also be applied in systems with enough *a priori* information. For this reason, more about this type of KBD is presented in the following subsection.

### 2.3 Model-based KBD with incomplete *a priori* diagnostic information

As it was said, the most critical situation for industrial diagnostic problem solving is <u>when there is neither enough measurement data nor enough *a priori* knowledge</u> (either *deep* or *experiential*). The first task in such situation is to extract more information from all available means, or, more precisely, to acquire in some way the necessary diagnostic knowledge.

According to very informative text in reference [5] (a kind of detailed survey), the techniques used in knowledge acquisition can be roughly divided into two categories:

- elicitation and
- machine induction.

*Knowledge elicitation*, either manual or automatic, is possible only when *a priory* diagnostic knowledge in some form (books, *expert* knowledge itself etc.) already exists. This category is used in rule-based KBD systems.

*Machine induction* (MI) is a special case of machine learning, which encompasses heuristic for generalising data types, methods for generating decision trees and rule sets, function induction and procedure synthesis. The fact is that clear methodological possibilities for automatic synthesis of knowledge for classification and diagnostic purposes exist in two emerging areas of research: *machine learning* and *neural networks*, [6]. Both approaches require training sample set of real examples (of correctly classified instances). An example is a vector of features' values, labeled to a particular class. Sets of such examples are then analysed by neural network or inductive algorithm and rules are generated automatically from these examples. From diagnostics point of view machine *inductive* learning has some important advantages related to neural networks. Firstly, it offers a reasonable explanation during classification, which enables the user to check the line of reasoning. Secondly, it provides an insight into laws of the domain: the obtained set of classification rules can be viewed as a new representation of domain knowledge. The neural networks cannot explicitly explain their results. This limitation disqualifies them for applications such as diagnosis where one usually wants to know answer on the question *why*. When example set does not exist, which is the situation we deal with, the simulation is the only means to provide data for training examples generation. However, in contrast with examples generated from real data which can be classify correctly, classification of examples with simulated data requires confirmation. This is another very important fact for using MI techniques instead of neural networks. The explanation during classification, which gives machine induction, makes possible both feature extraction and representative set of examples generation. Taking in consideration that MI represents learning concept in the form of decision trees or production rules, which in both cases can be diagnostic system, MI is the best method for diagnostic system building in situations with incomplete *a priori* information. In the following section is presented more about this very important kind of diagnostics.

## 3. The Inductive Learning Framework for Diagnostic System Building

A schematic description of the overall framework for diagnostic system development is shown in Fig. 2. The development has two phases: data preprocessing and diagnostic knowledge extraction. The essential elements of first phase are:

- *Archive.* Historical (measured) data, acquired by DAQ system in long period.
- *Simulation.* Simulation of the faults unrecorded in the archive that is often the only way to obtain information of the effects associated with occurrence of certain faults.

The second phase encompasses diagnostic knowledge extraction and structuring. It is an iterative process, which consists of: features identification and selection, training examples generation, diagnostic rules induction, rules evaluation. This process will continue to be performed until certain

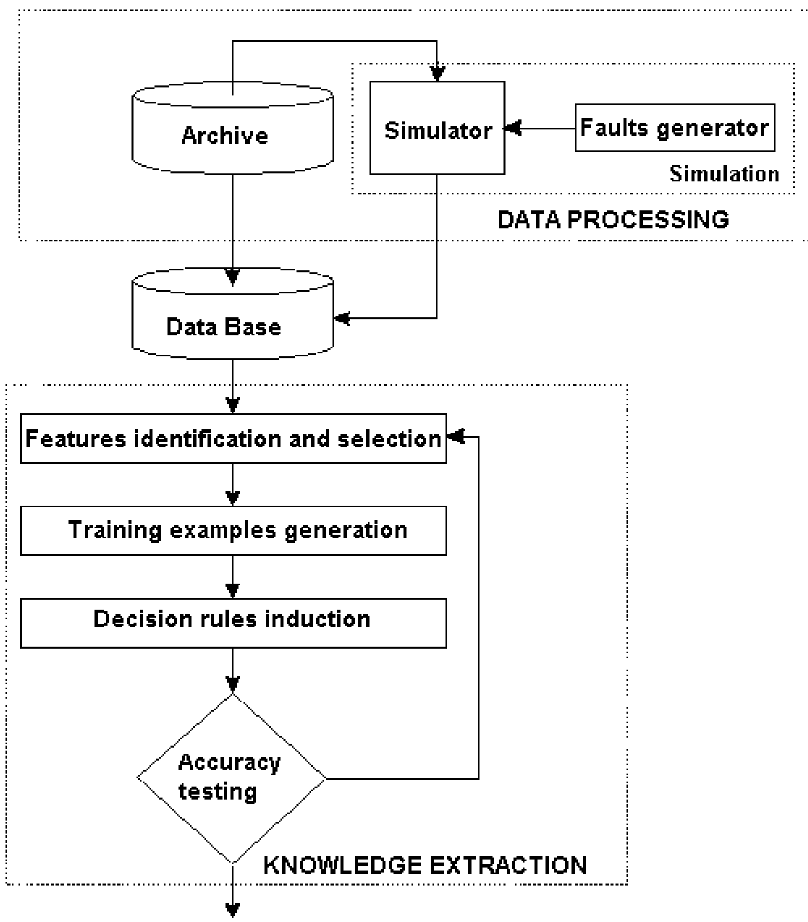


*Fig. 2. Framework for diagnostic system building*

criteria are satisfied, and then a prototype of the diagnostic system can be build. Short comments on each subprocess are as follows:

- *Features identification and selection.* Second phase starts with a thorough search trough: measured quantities, outputs of the simulation model and various derived functions or characteristics. To perform the search and to derive functions and characteristics we can use any of available tools for digital signal processing and time series analysis (for example MATLAB). The goal is to select a small set of features, which are the most informative ones in discriminate sense.

- *Training examples generation.* After selecting a set of relevant features, we can generate required set of training examples. In context of technical diagnostics, training example describes a state of the system under diagnosis, which is classified as normal or faulty.
- *Diagnostic rules induction.* When the learning system, chosen in advance, receives the set of training examples, it draws inductive inference from this set and produces a set of decision rules for correct classification of new (unseen) examples.
- *Decision rules evaluation.* Standard criteria for evaluation of decision rules are: accuracy, transparency and complexity, but usually only accuracy is examined during development. There are several strategies for testing accuracy of decision rules. The most reliable is testing on specially generated test examples (not involved in learning phase).

## 4. Some Experimental Results

a) The framework has been applied to the "NIS-GAS" network for natural gas transmission and distribution, in order to develop a system for leakage detection and location. It was at our disposal the archive of real, measured data, the stationary model of the network and Assistant Professional - a system for inductive learning from examples, [7]. In this paper, we present some results of our research project, as illustration of power of applied methods and techniques to extract knowledge from unstructured and mostly numeric data.

Figure 3 illustrates the process of features extraction using sophisticate data processing techniques. It shows time variation of two features selected for leakage detection problem. The first feature (Fig. 3a) is algebraic sum of pressure deviation over measurement sites (denoted for short as ASD). The second feature (Fig. 3b), derived from the first, represents cumulative sum of ASD, calculated over limited interval (cumul_ASD, for short).

Even simple visual comparison between Fig. 3a and Fig 3b shows that the second feature is far more informative in discriminate sense, as it is easier to separate its values that correspond to different classes.

The sample of the training set, generated for pipeline section, which supplies the area of Novi Sad, is shown in Table 1. In Table 1 each column represents one example (i. e., vector of 10 numeric elements). The complete set contains more than 1200 examples. Decision tree, shown in Figura 4, is

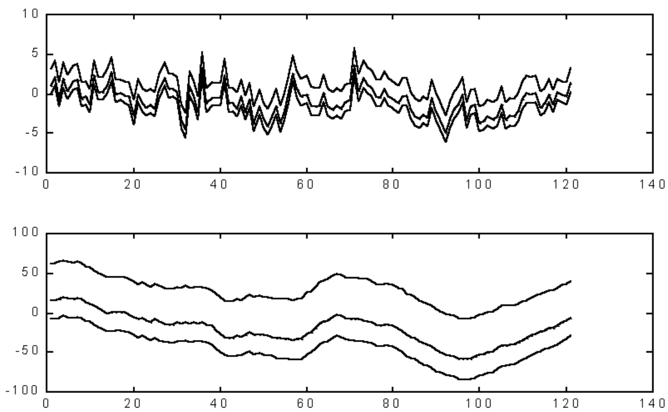

*Fig. 3. Time variation of two feature s: ASD and cumul_ASD.*

a result of inductive learning from complete set of examples performed by Assistant Professional. The tree correctly classifies 96.33 % of test examples.

Table 1. A subset of training examples

| class label | B1 | B1 | NO | NO | B3 | B3 | A1 |
|---|---|---|---|---|---|---|---|
| pressure | 30.41 | 30.41 | 30.39 | 30.43 | 30.43 | 30.39 | 30.43 |
| flow | 18171.8 | 17903.2 | 17968.7 | 18062.5 | 12446.8 | 12578.1 | 16003.1 |
| press_dev | 0.36 | 0.31 | 0.34 | 0.38 | 0.87 | 0.86 | 0.55 |
| ASD | −6.83 | −6.45 | 0.26 | −0.59 | 4.09 | 3.19 | 2.17 |
| line pack | 3147 | 3084 | 3053 | 3146 | 3138 | 3046 | 3143 |
| slope_1 | −0.00127 | −0.00082 | −0.00018 | 0.00072 | 0 | −0.00018 | 0 |
| slope_2 | −51.8378 | −80.2783 | 80.1139 | 57.5288 | 64.5305 | 56.0797 | 82.9685 |
| coher_1 | 0.9997 | 0.9997 | 0.9993 | 0.9996 | 0.9987 | 0.9993 | 0.9987 |
| cumul_dev | 8.78 | 8.83 | 8.56 | 8.52 | 20.08 | 20.13 | 12.67 |
| cumul_ASD | −62.1 | −66.46 | −35.01 | −36.03 | 30.67 | 29.79 | −12.19 |

b) We applied described methodology for faulty components detection and identification in production testing of analog electronic boards. In reference [8] such type of knowledge-based system is proposed and described. Its main parts are: the guided measuring probe, for data (voltages' values) acquisition and diagnostic expert system developed by using inductive machine learning technique for diagnostic rules acquisition. An example is vector of voltages' values at defined points of the boards, labeled to a particular class of faults.

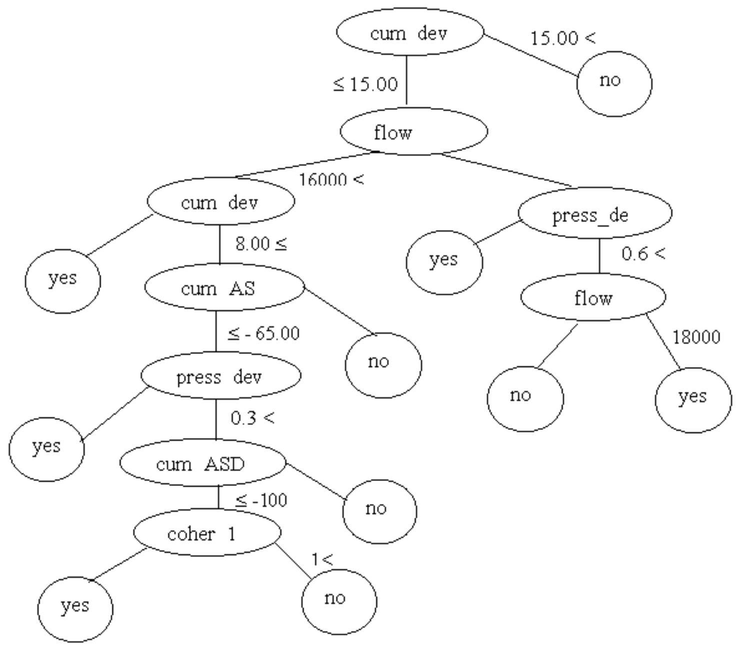


*Fig. 4. Decision tree for leakage detection on pipeline section "Novi Sad".*

Another similar example of application of described methodology based on MI for diagnostic system building is presented in reference [9].

Both examples point to the fact that diagnostic procedure, which must have been developed because the lack of measurement, can very successfully be implemented in industrial test and measurement procedures which obligatory include diagnostics.

## 5. Conclusion

The role of measurement in contemporary industrial diagnostics is to provide all basic diagnostic values. This is necessary requirement for realization of algorithmic diagnostic system which is characterized by great certainty. If this basic requirement is not satisfied, or if the system under diagnostics is not amenable to only numerical representation, then knowledge based diagnostic must be used. The choice of diagnostic approach in KBD depends on available measured data, too. In this paper, we have presented classification of industrial diagnostics in accordance with available measurement data. Much attention is focused on a framework for diagnostic system building based on inductive machine learning. The experimental results have shown that by using proposed framework it is possible to solve complex diagnostic problems, such as leakage detection in a network for natural gas transmission. We have attempted to point out to the paradox that the lack of diagnostically important measurement data has influence on development more sophisticated, knowledge based diagnostic technologies, which can be very successfully implemented in situation with enough measurement.

## REFERENCES

1. T.M. Romberg, J.L.Black, T.J. Ledwidge: *Signal Processing for Industrial Diagnostics.* John Wiley& Sons, England, 1996.
2. G.Bojković, V.Bajović: *Classification of computer based technical diagnostics.* Tehnika - Elektrotehnica, Vol. 46, No 5/6, 1997, pp. E1-E5 (in Serbian).
3. G. Bojković, V. Bajović, V. Kovačević: *Knowledge based diagnostics in gas distribution network with poor instrumentation.* XII International Conference of System Science, Wroclav, Poland, September 1995, Proc. Vol. III, pp. 189-186.
4. R. Isermann: *Process fault detection based on modeling and estimation methods - a survey.* Automatica, Vol 20, No 44, 1984, pp. 387-404.
5. G. Johannsen, J. Alty: *Knowledge engineering for industrial expert Systems.* Automatica, Vol. 27, No.1, 1991, pp. 97-114.
6. I. Konvalinka, V. Kovačević, V. Bajović, G. Bojković: *Decission trees development for leak detection on gas transmission system using stationary model and machine learning from examples.* First Internationale Conference on Intelligent Systems Engineering, U.K., Edinburgh, 1992.
7. V. Bajović, G. Bojković: *Inductive learning based framework for diagnostic system building.* 3rd International Symposium Interdisciplinary Regional Research, , Novi Sad, Sep. 1998, pp. 17-20.
8. V. Bajović, G. Bojković, V. Kovačević: *Knowledge based system for faulty components detection in production testing of electronic device.* IEEE International Symposium on Electronics and the Environment, San Francisco, California, May 1997, pp. 257-260.
9. V. Ghaemi, R. Godberson, W. Schwetlick, D. Filbert: *Some aspects of knowledge based fault diagnosis in electronic devices.* Measurement, vol. 10, no.1, Jan-Mar 1992, pp. 2-7.