Model based diagnosis using causal graph

Anna Sztyber

Institute of Automation Control and Robotics, Warsaw University of Technology

Abstract: This paper concerns fault diagnosis of industrial plants and complex systems with special interest in fault diagnosis system design. Scope of research connected with using causal graphs to fault diagnosis is presented. Directed graph is used to describe causal relationships between process variables and faults. New method for finding set of model structures based on causal graph is presented. Model structure is understood as an output variable and set of input variables. Algorithm for determining model sensitivity to faults is described. Method for finding possible ability to detect and isolate each fault given calculated set of models is described. Main ideas are explained on simple example.

Keywords: fault diagnosis, causal graph, model

1. Introduction

In industrial plants faults can lead to large economic losses and cause dangerous situations [5]. This is the reason why fault diagnosis is an important problem. In recent years issues of fault diagnosis system design are attracting a lot of attention.

Causal graphs are useful tools for fault diagnosis system analysis. This topic was first concerned in paper [1]. Graph vertices can represent process variables, system components and events like faults and operator interventions. Graph edges represent causal influences between vertices. Directed graphs can be used to fault symptoms propagation analysis [2, 3] and to find fault signatures [4, 6]. Simulation of fault propagation can be obtained and set of rules for fault discrimination can be built [7]. Another application of causal graphs is multiple fault diagnosis [8, 9].

One of important problems in causal graph analysis is existence of cycles. Methods for dealing with feedback and control loops are considered [10]. Another issue is size of the model for complex systems. An idea of graph partition is presented in paper [10]. Graph modelling real system can be obtained from mathematical description [11], piping and instrumentation diagrams [12] and from archival industrial databases [13, 14].

In most of papers diagnostic signal is understood as crossing of alarm thresholds [8, 15] or as an alarm coming from system component [16, 17]. In case of an alarm threshold crossing diagnostic signal is often described using fuzzy logic [18].

Using causal graph to model based diagnosis was first proposed in paper [19]. This work continues that idea. Main difference in relation to most of previous works is that diagnostic signal is understood as a difference between measured signal and reference value calculated from model. In this context model can be set of algebraic equations, differential algebraic equations, look-up table, neural model, fuzzy model etc.

2. Causal graph

As a model of a process causal graph is used. Vertices represent the process variables, control signals or faults excluding sensor faults. Directed edges represent influences between vertices. Following methods and algorithms will be explained on a simple example of a single tank system presented in fig. 1.

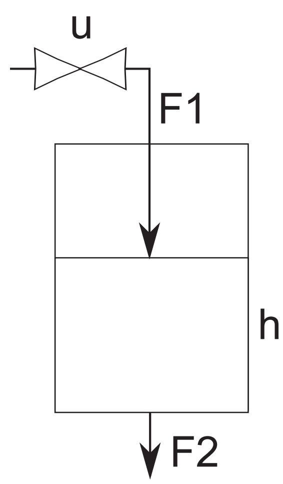

Fig. 1. Single tank system
Rys. 1. Układ zbiornika

List of variables is shown in tab. 2. Considered faults are presented in tab. 1.

Tab. 1. List of variables
Tab. 1. Lista zmiennych

CV	controller output
CVv	control signal received in valve
u	valve position
F1	inflow
h	tank level
F2	outflow

Tab. 2. List of faults

Tab. 2. Lista uszkodzeń

f_1	control circuit fault
f_2	valve fault
f_3	tank leakage
f_4	outlet clogging

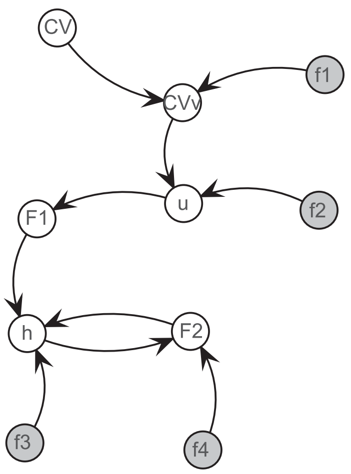

Fig. 2. Casual graph of a single tank system

Rys. 2. Graf przyczynowo-skutkowy układu zbiornika

Causal graph of a single tank system is presented in fig. 2. Graph vertices represent all variables and faults. Edges show causal relationships. Signal from a controller is send to a valve and received signal influences valve position. Degree of valve opening causes changes in inflow. Tank level depends on inflow and outflow. Growth of tank level causes increase of outflow.

Fault in a control circuit disturbs value of a control signal received in a valve. Valve fault influences valve position. Tank leakage causes decrease of tank level. Clogging of an outlet causes decrease of outflow.

This kind of causal graph containing vertices representing faults can be used to find set of possible process models and their sensitivity to faults.

3. Model structures

Model structure is understood as an output variable and set of input variables. Given causal graph of a system all possible model structures can be found. Method for finding model structures is presented in work [19]. This paper presents new method for solving this problem.

3.1. Requirements for model structures

Set of an input variables for a given output variables should fulfil following requirements:

1. For each input variable in a causal graph a path must exists from input variable to output variable.
2. Set of input variables should be complete.

Requirement no. 1 means that each model input should influence modelling variable. For example, there is no point in building model of valve position u with tank level h as an input because tank level have no influence on valve position (influence through control circuit is no considered). Complete set of an inputs means that set of input variables should cut all causal influences between outer variables and model output. If variable v is not influenced by any of model inputs then in a graph should not exist a path from vertex v to model output not containing any of model inputs. For example model with tank level h as an output and outflow $F1$ as an input does not have complete set of inputs because in a given causal graph exist path from inflow $F1$ to h.

Some remarks about causal relations between model inputs

When building model structures some additional requirement related to causal relationships between inputs should be considered. For example, model of a tank level h containing as an inputs valve position u and inflow $F1$ is not a good idea, because given measure of an inflow $F1$ data about valve position u gives no useful additional information. The strictest requirement is to forbid existence of any path between model inputs. That approach was presented in work [19]. In this paper another approach is proposed. The requirement is that in a causal graph must exist a path from each input of a model to an output not containing any other inputs. In other words each input variable has influence on output variable that cannot be described using other input variables. Difference between this two approaches is visible only when in a graph exists some path ramifications.

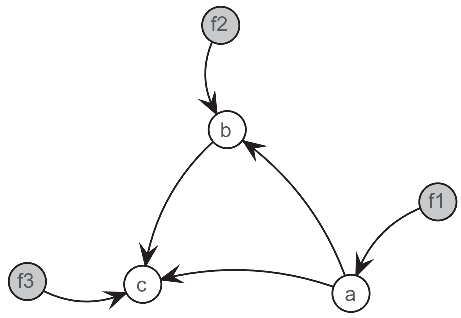

Fig. 3. Example graph G1

Rys. 3. Przykładowy graf G1

Consider example graph $G1$ presented in fig. 3. Vertices a, b and c represent process variables, vertices $f1$, $f2$ and $f3$ represents faults. There are two possible models of variable c: $\hat{c} = f(a)$ and $\hat{c} = f(a, b)$. Model with one input b has not complete set of inputs. Model with one input $\hat{c} = f(a)$ fulfils strict requirement of no causal relations between inputs. This model is disturbed by faults $f2$ and $f3$. Model with two inputs a and b does not fulfil strict requirement but fulfils requirement of existence of path from each input of a model to an output not containing any other inputs. This model is disturbed only by fault $f3$ which means that models with causal relations between variables can be used to improve faults discrimination.

3.2. Finding model structures

Calculation of all possible model structures contains following steps:

1. Finding and merging strongly connected components.
2. Topological sorting of vertices.
3. Building model structures.
4. Finding proper models.

5. Checking for fulfilment of requirement related to causal relations between inputs.
6. Adding models with more than one variable from the same strongly connected component.

Each step will be described in following paragraphs.

Strongly connected components Strongly connected component in a graph G is a set of vertices that for each pair of vertices u and v in a graph G exists path from u to v and from v to u [20]. One strongly connected component contains set of variables influencing each other. Merging strongly connected components of graph G solves problem of dealing with cycles. Strongly connected components of a graph G can be identified using well known algorithm based on a depth-first search [20]. Causal graph of a single tank system with merged strongly connected components is presented in fig. 4. This graph contains only one strongly connected component including vertices h and $F2$. These two vertices were replaced by one vertex named $SN0$.

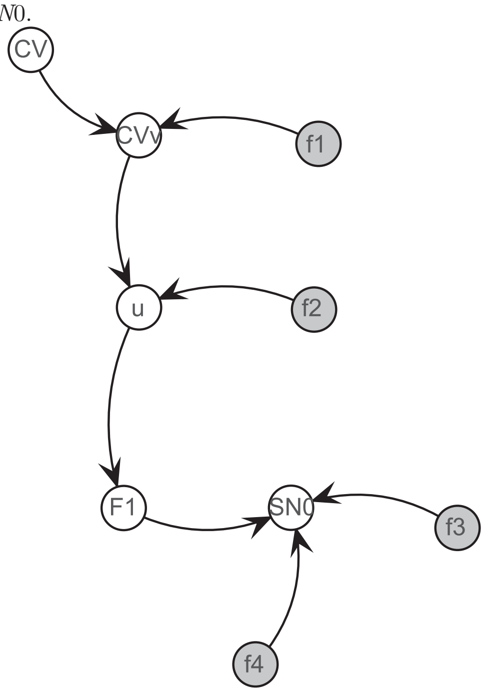

Fig. 4. Graph of a single tank system with merged strongly connected components

Rys. 4. Graf przyczynowo-skutkowy układu zbiornika z połączonymi silnie spójnymi składowymi

Topological sorting Topological sorting of graph provides partial order of vertices. If in a graph G exists a path from vertex u to vertex v, then vertex u precedes vertex v in obtained order. Topological sorting of a graph makes sense only when graph does not contain cycles, so it could be used to process causal graph with merged strongly connected components. Processing graph in a topological order ensures that when vertex u is processed all of its predecessors in a graph G were already processed. Topological sorting of a graph can be obtained using well known algorithm also based on a depth-first search [20]. In a graph of single tank system with merged strongly connected components after topological sorting vertices have order as follows: CV, CVv, u, $F1$ and $SN0$.

Building model structures All possible model structures fulfilling requirements 1 and 2 are found by algorithm 1.

Algorithm 1 ModelStructures(GX, Q)

while $Q \neq 0$ **do**
  $v \leftarrow max(Q)$
  $P \leftarrow$predecessors set of vertex v
  $add(Models,\ (v,\ P))$
  $Sets(P,\ P,\ v)$
end while

G_X is a symbol for causal graph of a system with merged strongly connected components and deleted faults. Vertices representing faults are not useful to finding model structures because fault cannot be a model input or output. Q is a priority queue containing vertices of graph G_X in topological order. Algorithm ModelStructures(G_X, Q) process all vertices v of graph G_X in a topological order. For each vertex v first model contains all v predecessors in a graph G_X as an input set. Then recursive procedure Sets(SP, P, v) is called. Set*Models* is a set of pairs. Each pair contains output variable of a model and set of input variables.

Algorithm 2 Sets(S_p, P, v)

for all $p \in S_p$ **do**
  if $p \in P$ **then**
    for all $M \in Models(p)$ **do**
      $remove(S_p,\ p)$
      $add(S_p,\ in(M))$
      $add(Models,\ (v,\ SP\ ))$
      $Sets(S_p,\ P,\ v)$
    end for
  end if
end for

Procedure Sets(S_p, P, v) extends set of models replacing each of model inputs p by inputs set of some model of p. This could be done because all models of p were found previously because vertices are processed in topological order. SP is a set of model inputs, v is modelled variable, P is a set of v predecessors. $Models(p)$ is a set of variable p models, $in(M)$ means input variables set of model M.

Models calculated for single tank system are listed in tab. 3.

Tab. 3. List of models for single tank system

Tab. 3. Lista modeli dla układu zbiornika

CVv	$[CV]$
u	$[CVv]$
$F1$	$[u]$, $[CV]$, $[CVv]$
$SN0$	$[u]$, $[CV]$, $[CVv]$, $[F1]$

Proper models For fault diagnosis can be used only models containing known variables. This motivates following definition.

Proper model structure is a model structure satisfying following conditions:

1. Model output is measured process variable or a merged strongly connected component containing at least one measured process variable.
2. Set of model inputs contains only control signals or measured process variables or a merged strongly connec-

ted components containing at least one measured process variable.

Assuming that following variables are measured in a single tank system: u, $F1$, h, $F2$ we obtain set of proper models listed in tab.4.

Tab. 4. List of proper models for single tank system
Tab. 4. Lista właściwych modeli dla układu zbiornika

u	$[CV]$
$F1$	$[u]$, $[CV]$
$SN0$	$[u]$, $[CV]$, $[F1]$

Requirement related to causal relationships between model inputs This part of algorithm depends on selected requirement. In case of requirement of existence of path from each input to output not containing other inputs algorithm 3 can be used. Edges in a graph GX are reversed and for all models M depth-first search is started from model output v. When one of model inputs is encountered its successors in graph GT are not added to a queue. If all model inputs were encountered then model M fulfils requirement.

Algorithm 3 CausalRelations($Models, G_X$)

```
create graph G^T by reversing all edges in G_X
    for all M ∈ Models do
        do DFS(v,G^T ) with stopping on M inputs
        if not all in(M) were met then
            remove(Models,M)
        end if
    end for
```

In single tank system example there is no problem with casual relationships between variables.

Additional models Last step of finding all model structures is adding models containing more than one variable from one strongly connected component. Variables in one strongly connected component all influences each other so this additional models could not fulfil requirement of no causal relations between model inputs. Despite this they are worth consideration because they can improve faults discrimination.

New models are added by algorithm 4.

Algorithm 4 $NewModels(G_X)$

```
for all SN ∈ G_X do
    create graph G_SN by splitting SN into vertices
    create graph G^T_SN reversing all edges G_SN
    for all measured v ∈ SN do
        for all measured w ∈ SN do
            delete outgoing edges of w
            DFS(G^T_SN, v) with stopping on measured variables
                    var(v,w) ← set of first measured variables encountered in DFS(G^T_SN, v)
            AddModel(v,w, var(v,w))
        end for
    end for
end for
```

Algorithm search for new models for all strongly connected components SN in a graph G_X. Component SN is divided into vertices and edges in a graph are reversed. All possible model outputs v and all possible additional inputs w from the same component are considered. Set $var(v, w)$ contains measured variables u for which exists path from u to v not containing w. Only measured variables are considered because obtained models should be proper models. In $DFS(G^T_{SN}, v)$ successors of measured variables are not added to a queue, so set $var(v,w)$ contains only variables near v. Models of encountered variables are already obtained so there is no point in further searching. At the end of algorithm recursive procedure $AddModel(v,w, var(v,w))$ is called.

Procedure $AddModel(v, addIn, variables)$ is presented as an algorithm 5. v is modelled variable, $addIn$ is a set of possible additional inputs from the same strongly connected component and $variables$ is a set of variables u for which exist path from u to v not containing any vertex from set $addIn$. If $variables$ is an empty set then set $addIn$ is an complete model and can be add as an new model of variable v. If set $variables$ is included in some input set of model of v then model with input set containing $variables$ and $addIn$ can be added. At the end searching for next additional input from the same component is started.

Algorithm 5 $AddModel(v, addIn, variables)$

```
if variables = 0 then
    add(Models, (v, addIn))
end if
for all m ∈ Models(v) do
    if variables ∈ in(m) then
        add(Models, (v, addIn ∪ variables)
    end if
end for
for all vn ∈ SN(v) do
    if vn ∈ v variables then
    ok ← TRUE
    for all in ∈ addIn do
        if in ∉ var(v, vn) then
            ok ← FALSE
        end if
        remove(variables, vn)
        if ok then
            variables ← variables ∩ var(v, vn)
            add(addIn, vn)
        end if
        AddModel(v, addIn, variables)
    end for
    end if
  end for
```

All models obtained for a single tank example are listed in tab. 5.

Tab. 5. List of all proper models for single tank system
Tab. 5. Lista wszystkich właściwych modeli dla układu zbiornika

Model output	Set of all proper models
u	$[CV]$
$F1$	$[u]$, $[CV]$
h	$[u]$, $[CV]$, $[F1]$, $[F2, CV]$, $[u, F2]$, $[F1, F2]$
$F2$	$[u]$, $[CV]$, $[F1]$, $[h]$

4. Models sensitivity

For models designed for diagnosis system their sensitivity to faults is very important.

Model structure sensitivity to faults is a set of faults which can cause difference between the value calculated from model and measured value of variable.

In a causal graph model structure sensitivity to faults is a set of faults for which in a graph exists path from fault to modelled variable not containing any of input variables. Set of faults disturbing each model is found in an algorithm 6. Graph edges are reversed and for each model outgoing edges of each input variable are deleted. Model is sensitive to faults encountered by depth-first search started from model output v.

Algorithm 6 *Faults*(*Models*, *G*)

```
create graph G^T reversing all edges G
for all M ∈ Models do
    for all w ∈ in(M) do
        delete outgoing edges of w
    end for
    DFS(G^T, v)
    f(M) ← set of faults encountered in DFS(G^T, v)
end for
```

Last step is to add sensitivity to sensor faults. Model is sensible to faults of all sensors measuring input variables and output variable.

Models and their sensitivity to faults for single tank example were listed in tab. 6. Sensor faults are marked by letter *f* and symbol of measured variable.

5. Faults detection and discrimination

Given set of model structures and their sensitivity to faults possible ability of diagnosis system can be obtained. Fault can be detected when exists at least one model sensitive to this faults. Two faults can be distinguished when they can be detected and at least one model sensitive to one of them and not sensitive to another exists.

Results obtained this way are optimistic prognosis because causal graph is a qualitative model of a process. Models good from causal point of view can be impossible to use in practise in case of bad quality of measurements, presence of large disturbances or little fault influence.

Tab. 6. List of all models with sensitivity to faults for single tank system
Tab. 6. Lista wszystkich modeli wraz z ich wrażliwością na uszkodzenia

Model output	inputs	faults
u	$[CV]$	fu, $f1$, $f2$
$F1$	$[CV]$	$fF1$, $f1$, $f2$
$F1$	$[u]$	fu, $fF1$
h	$[u]$	fu, fh, $f3$, $f4$
h	$[F1]$	fh, $f3$, $fF1$, $f4$
h	$[CV]$	fh, $f3$, $f4$, $f1$, $f2$
h	$[F2, CV]$	fh, $f3$, $f1$, $f2$, $fF2$
h	$[u, F2]$	fu, fh, $f3$, $fF2$
h	$[F1, F2]$	fh, $f3$, $fF1$, $fF2$
$F2$	$[u]$	fu, $f3$, $f4$, $fF2$
$F2$	$[F1]$	$f3$, $fF1$, $f4$, $fF2$
$F2$	$[h]$	fh, $f4$, $fF2$
CV	$[h]$	$f3$, $f4$, $f1$, $f2$, $fF2$

In a single tank example all faults can be detected. Faults $f1$ and $f2$ cannot be distinguished.

6. Summary

Applications of directed graphs to fault diagnosis were described. Idea of a causal graph searching application to a model based diagnosis was presented.

Requirements for causal relation between model input and output variables were discussed and new requirement for causal relations between model inputs was proposed which allows generation of additional models and can give better faults discrimination.

New method for finding set of possible model structures was presented. Method differs from algorithm presented in [19]. Problem of dealing with cycles in graph was solved. Need for generation of special tree for each output variable was eliminated. Proposed method allows generation of all models at once using results calculated previously for models of other output variables. Presented algorithms can be easily implemented using well known methods of graph processing.

Method for finding possible ability of diagnosis system based on a calculated set of models was described.

Acknowledgements

The project has been partially financed from funds of National Science Centre granted by decision number DEC-2011/01/B/ST7/06183 and from Grant of Dean of Faculty of Mechatronics.

Bibliography

1. Ira M., Aoki K., O'Shima E., Matsuyama H., *An algorithm for diagnosis of system failures in the chemical process*, "Computers&Chemical Engineering", 3:489–493, 1979.

2. Ulerich N.H., Powers G.J., *On-line hazard aversion and fault diagnosis in chemical processes: The digraph + fault-tree method*. *IEEE Transactions on Reliability*, 171–177, 1988.
3. Fan Yang, Shah L.S., Deyun Xiao, *Sdg modelbased analysis of fault propagation in control systems*. Canadian Conference on Electrical and Computer Engineering, 1152–1157, 2009.
4. Chung-Chien Changt, Cheng-Ching Yu, *Online fault diagnosis using the signed directed graph*. "Industrial & Engineering Chemistry Research", 29:1290–1299, 1990.
5. Blanke M., Kinnaert M., Lunze J., Staroswiecki M., *Diagnosis and fault-tolerant control*. Springer-Verlag, Berlin 2003.
6. Gang Xie, Xiue Wang, Keming Xie, *Sdg-based fault diagnosis and application based on reasoning method of granular computing*, Control and Decision Conference, 1718–1722, 2010.
7. Tarifa E.E., Scenna N.J., *Fault diagnosis, direct graphs, and fuzzy logic*, "Computers & Chemical Engineering", 21:649–654, 1997.
8. Fang T., Pattipati K.R., Deb S., Malepati V.N., *Computationally efficient algorithms for multiple fault diagnosis in large graph-based systems*, "IEEE Transactions on Systems, Man and Cybernetics", Part A: Systems and Humans, 73–85, 2003.
9. Hideo Nakano, Yoshiro Nakanishi, *Graph representation and diagnosis for multiunit faults*, "IEEE Transactions on Reliability", 23(5):320–325, 1974.
10. Maurya M.R., Rengaswamy R., Venkatasubramanian V., *A systematic framework for the development and analysis of signed digraphs for chemical processes*, Industrial & Engineering Chemistry Research, 4789–4827, 2003.
11. Maurya M.R., Rengaswamy R., Venkatasubramanian V., *A signed directed graph and qualitative trend analysis-based framework for incipient fault diagnosis*, "Chemical Engineering Research and Design", 85(29):1407–1422, 2007.
12. Fan Yang, Shah L.S., Deyun Xiao, *Signed directed graph modeling of industrial processes and their validation by data-based methods*. 2010 Conference on Control and Fault-Tolerant Systems (Sys-Tol), 387–392, 2010.
13. Bauer M., Cox J.W., Caveness M.H., Downs J.J., Thornhill N.F., *Finding the direction of disturbance propagation in a chemical process using transfer entropy*, "IEEE Transactions on Control Systems Technology", 15(1), 12–21, 2007.
14. Bauer M., Thornhill N.F., *A practical method for identifying the propagation path of plant-wide disturbances*, "Journal of Process Control", 18:707–719, 2008.
15. Wen-Liang Cao, Bing-ShuWang, Liang-Yu Ma, Ji Zhang, Jian-Qiang Gao. *Fault diagnosis approach based on the integration of qualitative model and quantitative knowledge of signed directed graph*. International Conference on Machine Learning and Cybernetics, 2251–2256, 2005.
16. Lakshmanan K.B., Rosenkrantz D.J., Ravi S.S., *Alarm placement in systems with fault propagation*. "Theoretical Computer Science", 243:1217–1223, 2000.
17. Rao N.S.V., *On parallel algorithms for single-fault diagnosis in fault propagation graph systems*, "IEEE Transactions on Parallel and Distributed Systems", 7(12):1217–1223, 1996.
18. Bingshu Wang, Wenliang Cao, Liangyu Ma, Ji Zhang, *Fault diagnosis approach based on qualitative model of signed directed graph and reasoning rules*. FSKD (2)'05, 339–343, 2005.
19. Ostasz A., *Causal graph and its application to finding residual set and diagnostic relation* (in polish). PhD thesis, Warsaw University of Technology, Warsaw 2006.
20. Cormen T.H., Leiserson C.E., Rivest R., *Introduction to Algorithms*. Massachusetts Institute of Technology, 2009.

Zastosowanie grafu przyczynowo-skutkowego w diagnostyce wykorzystującej modele procesu

Streszczenie: Artykuł dotyczy zagadnień projektowania systemów diagnostyki procesów przemysłowych z wykorzystaniem grafów przyczynowo-skutkowych. Przedstawiono stan badań dotyczących zastosowania grafów w diagnostyce. Graf przyczynowo-skutkowy jest grafem skierowanym zawierającym wierzchołki reprezentujące zmienne i uszkodzenia oraz krawędzie obrazujące wzajemne oddziaływania. Zaprezentowano metodę znajdowana zbioru struktur wszystkich modeli, które mogą zostać wykorzystane w systemie diagnostycznym. Opisany jest sposób określania wrażliwości modeli na uszkodzenia oraz znajdowania możliwej do uzyskania wykrywalności i rozróżnialności uszkodzeń.

Słowa kluczowe: diagnostyka przemysłowa, graf przyczynowo-skutkowy, modele

Anna Sztyber, MSc

Graduated from Automation Control and Robotics at Faculty of Mechatronics of Warsaw University of Technology with excellent results. Now she is a PhD student and research worker of Institute of Automation Control and Robotics. Field of her interest is diagnosis of industrial plants.

e-mail: sztyber.anna@gmail.com