Faults diagnosis using self-organizing maps: A case study on the DAMADICS benchmark problem

Andrzej Katunin, Marcin Amarowicz, Paweł Chrzanowski
Silesian University of Technology,
Institute of Fundamentals of Machinery Design
Konarskiego 18A Street, 44-100 Gliwice, Poland
Email: {andrzej.katunin,marcin.amarowicz,pawel.chrzanowski}@polsl.pl

Abstract—This paper deals with a method of faults detection and identification based on the clusterization of the multiple diagnostic signals. Various types of faults and character of their occurrence were simulated using DAMADICS Benchmark Process Control System. A great advantage of the applied approach based on self-organizing (Kohonen) maps is that even the smallest differences in signals allow for detection, isolation and identification of type of occurred faults with respect to the healthy condition of the investigated system based on the unsupervised learning. It was shown that in some cases the faults, which are undetectable during monitoring of simple heuristic and statistical parameters and other previously applied methods, are recognizable when the approach based on self-organizing maps is applied. The case studies presented in this paper show the faults detection procedure as well as clusterization of types and successful classification of almost all the unique faulty states of the investigated system.

I. Introduction

AN increasing usage of control and automation systems in industrial applications influences on development of novel methods of fault diagnosis, particularly their detection, isolation and identification. Moreover, the industrial requirements for such procedures are to be sensitive even for the early stage of development of faults and should ensure the implementation of them in on-line monitoring systems. The specificity of the physical nature of some faults requires the application of new methods in order to detect and classify them with the lowest possible false alarms.

Recent studies in the area of the fault diagnosis are based on many different approaches. They can be classified, in general, into the quantitative and qualitative approaches. The most simple quantitative approach is an application of statistical measures to the diagnostic signals, observation and inference about a fault presence basing on these measures [1]. However, there is a large amount of research done with application of qualitative fault diagnosis, primarily using soft computing methods. Several authors used such methods in the fault diagnosis, in particular in [2], [3] the genetic programming for observer-based fault diagnosis and evolutionary learning of fuzzy models were used, while the authors of [4] used a fuzzy qualitative reasoning approach for the fault detection problem. A lot of studies are concerned with artificial neural networks

(ANN) application to such a class of problems. In [5] the authors applied the group method of data handling (GMDH) neural networks, while in [6], [7] the authors used a neuro-fuzzy approach and ANN with decision trees, in [8] the hidden Markov model for the fault diagnosis was applied.

One of the novel and promising diagnostic tools is the group of approaches based on Kohonen self-organizing maps (SOM). SOM are very often used in problems of the analysis of large data structures e.g. in the problems of clustering or classification [9], [10], [11], [12], image processing [13], [14], [15], robotics [16], [17], time series forecasting [18], [19], [20] and faults detection and identification [21], [22], [23].

In the presented paper the authors analyzed and discussed the overall analysis of all faults using SOM-based approach, which diagnostic signals are able to be generated using the DAMADICS benchmark. The only previous study [24], which was carried out on DAMADICS data using the SOM-based approach, presented a possibility of application of SOM to the faults clusterization, but the results are limited to the analysis of three selected faults. In the present study the process signals were simulated basing on real measurements of diagnostic signals on the actuator system during the evaporisation process in the Lublin sugar factory (Poland) [25], [26]. In order to detect, isolate and identify the faults simulated using DAMADICS benchmark actuator system the SOM-based approach was applied.

II. Description of the problem and the methodology

A. The DAMADICS benchmark

The investigated problem of the faults detection and identification was based on sets of diagnostic signals, generated using the DAMADICS benchmark actuator system, which simulated various types of possible faults. The scheme of the investigated system is presented in Fig. 1 [27]. The actuator system consists of the pneumatic servo-motor S, a control valve V and a positioner P. These three parts of a system are composed by a set of measured diagnostic signals: external controller output CV, flow sensor measurement F, input $P1$ and output $P2$ valve pressure, medium temperature T and the rod displacement X.

Using the DABLib Simulink®-Matlab® library 19 fault types with a variable number of fault intensities (three stages

TABLE I: Types of faults simulated in the DAMADICS benchmark

Symbol	Fault
$f1$	Valve clogging
$f2$	Valve sedimentation
$f3$	Valve erosion
$f4$	Increased of valve or bushing friction
$f5$	External leakage
$f6$	Internal leakage
$f7$	Medium evaporation or critical flow
$f8$	Twisted servo-motor's piston rod
$f9$	Servo-motor's housing or terminals tightness
$f10$	Servo-motor's diaphragm perforation
$f11$	Servo-motor's spring fault
$f12$	Electro-pneumatic transducer fault
$f13$	Rod displacement sensor fault
$f14$	Pressure sensor fault
$f15$	Positioner supply pressure drop
$f16$	Increase of pressure on valve inlet
$f17$	Pressure drop on valve inlet
$f18$	Fully or partly opened bypass valves
$f19$	Flow rate sensor fault

Fig. 1: The scheme of actuator system

of abrupt faults and the incipient fault) can be modelled, which gives total of 45 cases including the case with no faults occurrence. One should consider that not every fault simulation has physical foundation of existence, thus for some types of faults a limited number of intensity subcases existed. The complimentary list of faults that can be modelled and which were used in further analyses was presented with a description in Table I and Table II [27].

TABLE II: Faults able for simulation in the DAMADICS benchmark

No.	Small*	Medium*	Large*	Incipient
$f1$	X	X	X	
$f2$			X	X
$f3$				X
$f4$				X
$f5$				X
$f6$				X
$f7$	X	X	X	
$f8$	X	X	X	
$f9$				X
$f10$	X	X	X	
$f11$			X	X
$f12$	X	X	X	
$f13$	X	X	X	X
$f14$	X	X	X	
$f15$			X	
$f16$	X	X	X	
$f17$			X	X
$f18$	X	X	X	X
$f19$	X	X	X	

* Abrupt faults

B. Previous studies related to diagnosis using DAMADICS benchmark

The problem of faults identification based on DAMADICS benchmark has been investigated by several scientific teams. The first studies were performed by the authors of the benchmark using various methods of faults detection, isolation
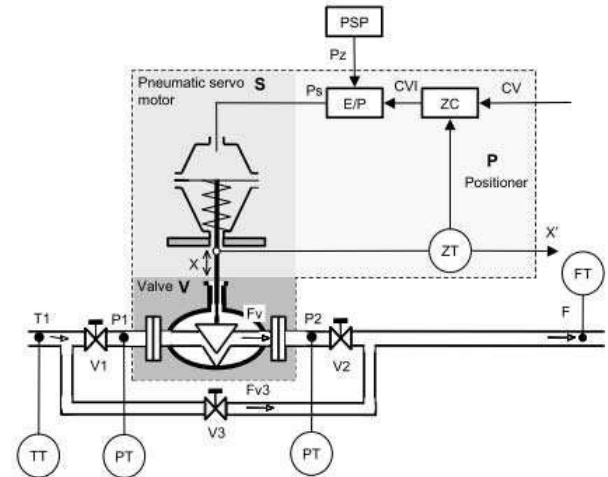
and identification. One can notice the general tendency in these papers where the authors take into consideration only selected cases of faults. The authors of [28] performed the study on faults detection and isolation based on the analysis of residua using binary diagnostic matrices for single sets of signals for a given faults case as well as the pairs and triplets of such cases. The authors of [29] used timed automata approach and considered three selected faults in their identification procedure. The authors of [30] focused on the abrupt large faults available in the DAMADICS benchmark for faults detection using a spectral estimation approach. Other approaches were to consider GMDH neural networks [5], interval observers [31] fuzzy classifiers for fault detection and isolation [32], structural analysis [33] in order to evaluate fault isolability, etc. The only study, which considered the whole set of faults possible to simulate in the DAMADICS benchmark, was performed by the authors of [34]. They used an approach of qualitative reasoning coupled with fuzzy neural networks and basing on their results only two cases were not detected and isolated.

The main goal of the studies on the DAMADICS benchmark was to find an appropriate methodology, which allows for detection, isolation and identification of all types of faults available in the benchmark and relations between them, which can be carried out within a single analysis procedure using all signals of all available faulty cases.

C. Data preparation

The authors of [35] stated that the faults $f8$ and $f12$ have theoretical behaviour, i.e. the methods applied by the authors were not able to distinguish these faults from the healthy state. Results obtained by the authors of [34] show that the small type fault $f16$ and incipient type fault $f18$ are not detectable, while the authors of [33] stated that using their method the faults $f9$ and $f16$ are not detectable. The authors of [5] did not detected some types of faults: $f5$, $f8$, $f9$, $f12$ and $f14$ using the GMDH neural network-based approach. In

the presented study all of the faulty cases were considered in the analysis. Considering initial analysis of signals generated from the DAMADICS benchmark, the signal $P2$ was removed from every set of tested cases due to the fact that this signal remained insensitive to the faults of a system. Moreover, following the previous studies in the fault detection problem on the DAMADICS benchmark data [4], [35], it is difficult to detect some of the faults due to their weak intensities (i.e. very small changes in the diagnostic signals) and/or very slow development. However, these faults cases were considered in further studies in order to investigate their detectability using the proposed methodology.

The sets of signals of each case listed in Table II were generated using the DABLib library. In each case the fault occurred after the 900th second of the simulation. Following this, the duration of the most of the generated sets of signals was limited to 2000 seconds. In some cases of slowly propagating incipient faults the duration of simulation was extended to 4600 seconds ($f17$) and to 86000 seconds ($f2$, $f3$, $f5$, $f6$, $f9$, $f11$ and $f18$) in order to achieve the full propagation history of a given fault until the fault index reaches the unity value.

During the performed researches four classes of faults: small, medium and large of abrupt and incipient faults, were considered. For the clarity of results presentation the numbers from 1 to 4 were assigned to the classes as the subscripts in the form: $fault_{class}$, e.g. the notation $(f2, f14)_4$ denotes that the faults $f2$ and $f14$ of incipient type, whereas $f3_{2,3}$ denotes that the fault $f3$ of medium and large abrupt types is considered.

D. Simple statistics-based direct diagnostics

In the diagnostics of processes, the often applied approach for detection and isolation of faults is the direct diagnostics, which is based on simple signal processing and control of constraints. For this purpose the current values of the features of process variables such as mean value, root mean square, maximal and minimal values, variance and others, are determined. In the case of diagnosing of a system basing on the alarm thresholds the analysis is performed on the basis of a comparison of current signal value of a process with the assumed minimal and maximal values, between which the signal can change in the case of non-faulty condition. In the case of statistical features the estimators in the specified window are calculated and then compared with the appropriate values determined for the non-faulty condition.

In the investigated problem both approaches were applied. In the case of control of constraints the evaluation was performed basing on comparison of current values of signals of processes with maximal and minimal values of the signals representing non-faulty condition of a system. In the case of the analysis of statistical features the mean value and a variance were calculated for comparative procedures. For the comparison purpose of current signals with the signals for the non-fault condition the threshold of 5% of variability range was assumed. When the values of signals exceed the threshold

the change is recognized as significant one, which indicates the damage. The results of damage detection using the above-described approaches were stored in Tables III, IV, V and VI. Each table refers to the intensity of a damage and the presented values denote the ability of detection of a given fault using the described approaches: 0 for the case of undetectable fault and 1 for the case of detectable fault.

Obtained results show that the detection of the following faults was not possible: small faults $(f10, f14, f16, f8)_1$, medium faults $(f10, f14, f16, f8)_3$, large faults $(f14, f8)_3$ and incipient faults $(f11, f2, f3, f4, f5, f6, f9)_4$. As can be noticed, the application of the described approaches for faults detection is possible only for a limited group of cases. The isolation of faults using these approaches is not possible in direct manner, but could be possible only after application of the algorithm, which allows for distinguish the faults. One of such extensions, which allows for distinguishing the faults, is the binary diagnostic matrix, however in numerous cases distinguishing of the faults was not possible. Therefore, it is necessary to develop appropriate approaches that would be sensitive to changes in the diagnostic signals and allow for their distinguishing and diagnostic inference about faults occurrence.

E. Motivation

Considering the results of faults detection presented above it can be noticed that the simple statistical methods often remain insensitive to changes in signals and thus the faults could not be detected. In order to detect and identify the faults occurred during the investigated industrial process the more advanced techniques should be applied. The SOM seems to be an effective tool both for the faults detection and identification and has an adequate sensitivity to recognize even the tiny changes in the analysed signals. Moreover, the application of SOM allows for the analysis and comparison of sets of multiple signals, i.e. it is not necessary to preprocess the measurement data in order to find the signal containing information about the occurred fault.

III. SOM AND METHODOLOGY

A. Fundamentals on the self-organizing maps

The Kohonen SOM are a type of ANN with unsupervised learning. Two layers, input and output, can be distinguished in the structure of this network (see Fig. 2). The input layer is a vector of neurons, while the output layer is a multidimensional representation of neurons, commonly the two-dimensional (2D) or three-dimensional (3D) one. In the case of 2D representation a rectangular or hexagonal map can be used. A toroidal or cylindrical maps could be used in case of 3D representation. The number of neurons in the input layer is equal to the number of attributes that describe each input data. In case of output layer $5\sqrt{N}$ neurons, where N is the number of input samples, was usually used [36]. But the practical size of this layer depends on nature of input samples. For diversified cases of input data bigger maps should be applied. Each neuron from input layer is connected with each neuron

TABLE III: Results of statistics-based diagnostics for abrupt small faults

type	$f10_1$	$f12_1$	$f13_1$	$f14_1$	$f16_1$	$f18_1$	$f19_1$	$f1_1$	$f7_1$	$f8_1$
min	0	0	0	0	0	0	0	0	0	0
max	0	0	0	0	0	0	0	0	1	0
mean	0	0	1	0	0	1	1	0	1	0
var	0	1	1	0	0	1	1	1	1	0

TABLE IV: Results of statistics-based diagnostics for abrupt medium faults

type	$f10_2$	$f12_2$	$f13_2$	$f14_2$	$f16_2$	$f18_2$	$f19_2$	$f1_2$	$f7_2$	$f8_2$
min	0	0	0	0	0	0	0	0	0	0
max	0	0	0	0	0	0	0	0	1	0
mean	0	0	1	0	0	1	1	1	1	0
var	0	1	1	0	0	1	1	1	1	0

TABLE V: Results of statistics-based diagnostics for abrupt large faults

type	$f10_3$ $f17_3$	$f11_3$ $f18_3$	$f12_3$ $f19_3$	$f13_3$ $f1_3$	$f14_3$ $f2_3$	$f15_3$ $f7_3$	$f16_3$ $f8_3$
min	0	0	0	0	0	1	0
	0	0	0	0	1	0	0
max	0	0	0	0	0	1	0
	0	0	0	0	0	1	0
mean	1	0	0	1	0	1	1
	1	1	1	1	1	1	0
var	1	1	1	1	0	1	1
	1	1	1	1	1	1	0

TABLE VI: Results of statistics-based diagnostics for incipient faults

type	$f11_4$	$f13_4$	$f17_4$	$f18_4$	$f2_4$	$f3_4$	$f4_4$	$f5_4$	$f6_4$	$f9_4$
min	0	0	0	0	0	0	0	0	0	0
max	0	0	0	0	0	0	0	0	0	0
mean	0	1	1	0	0	0	0	0	0	0
var	0	1	1	1	0	0	0	0	0	0
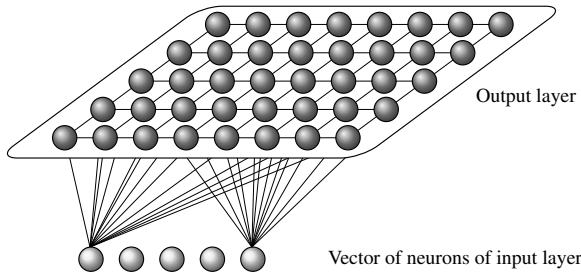
Fig. 2: The structure of the self-organizing map

of output layer [37]. For an arbitrary neuron n_i from output layer it is possible to define some neighbourhood. For this purpose the radius r is used. Based on this radius some area (called a neighbourhood) around considered neuron can be selected. The neighbourhood of the neuron n_i is described by a function $h_{ci}[\bullet] \in [0; 1]_R$ called a neighbourhood's function. The values of this function are distributed from neuron n_i along the radius r.

During the learning process of SOM for particular input data the best neuron from output layer is selected. In the literature this neuron is called a winner neuron or best matching unit (BMU). For the purpose of such selection some of the distance metrics e.g. Euclidean, Chebyshev, Manhattan, etc. can be used. In a learning process the weights of chosen neurons

are updated according to the formula:

$$w_i(k + 1) = w_i(k) + \eta(k)h_{ci}[x(k) - w_i(k)], \qquad (1)$$

where k is the number of iterations of the learning process, $\eta(k)$ is the learning coefficient, h_{ci} is the value of the neighbourhood function of the neuron n_i with respect to the neighbourhood of the winner neuron n_c and $x(k)$ is the input vector inputted to the network during k-th iteration. Generally the learning process consists of two stages. The first of them is called Winner Takes Most (WTM). During this stage the weights of a winner neuron and its neighbours neurons are updated. Second stage is called Winner Takes All (WTA). In this case the radius $r = 0$, therefore the value of coefficient $h_{ci} = 1$, and only the weights of the winner neuron are updated.

B. The methodology

For the purpose of carrying out the research described in this paper, the SOM Toolbox for Matlab® environment [38] was used. The input data was prepared basing on values of particular signals measures during occurrence of some faults. From each of considered signals (five signals without "$P2$, see the list in Sec. II-A) N consecutive values were considered (e.g. from 100 s to $(100 + N)$ s). Based on these values the vectors of $5N$ elements that describe particular faults were created. Finally, the input matrix has a size of $M \times 5N$, where

M is a number of considered fault cases plus one non-faulty exemplary case. During the conducted study different numbers of considered cases were being taken into consideration. The main settings of the SOM algorithm were as follows:

- type of output layer – 2D hexagonal maps of neurons,
- size of maps – from 10×10 for the simplest example to 100×100 for the most complicated examples,
- data normalization – variance was normalized to one,
- training algorithm – batch algorithm,
- number of training (WTM,WTA) iterations – from $(10, 10)$ to $(100, 80)$, depending on complexity of data,
- neighbourhood function – Gaussian function,
- distance measure – Euclidean metric.

For particular cases of input data the U-matrix maps were determined. An inspection of the obtained maps and quantization error were used for verification and evaluation of the quality of obtained results. For selected neurons of these maps the labels of considered faults or non-faulty case were assigned. Based on determined in this way maps the relevant conclusions about the possibility of detection, isolation and identification of particular faults were formulated.

In accordance with [39] the fault is detected when the difference in the set of signals with respect to an initial (healthy) state can be observed. As isolation stage we assume that the faults can be distinguished from each other and the identification stage allows us to identify a class and a type of a given fault. In this context we assume that basing on obtained SOM maps the fault is detected, isolated and identified when is located in separate cluster of map according to clusters related to the adequate non-faulty state or other fault cases. According to this assumption in Fig. 3 (subscript s and b in labels of faults denotes the small and big cases of fault, respectively) one can see that all faults are detectable. Faults $f2_s$ and $f2_b$ are detectable, but because they are located in the same cluster there are not isolated. Faults $f8_b$ and $f11_s$ are detectable and isolable as well since they are located in separate clusters. Finally, faults $f1_b$ and $f1_s$ are both detectable, isolable and identifiable. Additionally, based on obtained maps it is possible to conclude that some faults e.g. $f3_b$ could be weakly detectable.

In the most cases 1100 samples of each signal, starting from 901-th second (the onset of fault), were taken into account during the study. For several cases the number of considered samples was enlarged due to slow propagation of some of incipient faults. The analysis was realized following the standard two-step approach used during application of ANN. Firstly, the SOM was trained on standard data obtained from the generator following the description above. Then, three additional sets of data with modified character of noise distribution in signals were generated in order to validate the proposed approach during the testing stage.

IV. RESULTS

A. Diagnosing process

At the beginning, all of the cases of faults were considered in order to evaluate their distinguishability. The re-
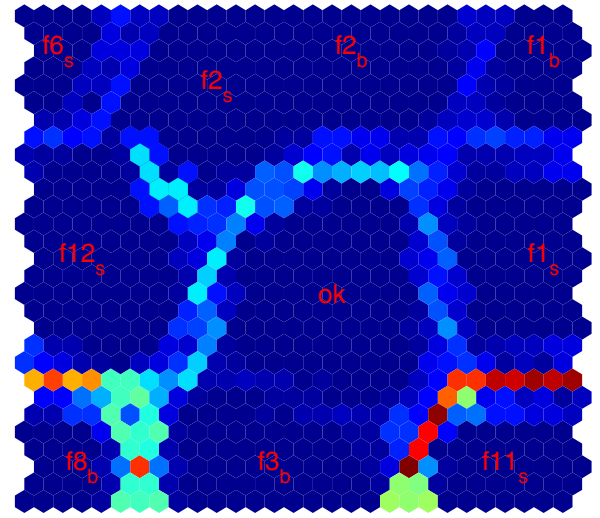
Fig. 3: Exemplary maps with associated faults for selected clusters

sulted exemplary maps for this study are shown in Fig. 4. Two groups of faults: weakly and very well detected can be distinguished. The set of weakly detectable faults (the same cluster on the obtained map) contains the following faults: $(f8, f14, f16)_1$, $(f8, f14, f16)_2$, $(f8, f14)_3$, $(f2, f3, f5, f6, f9, f11, f18)_4$. On the other hand, the subset of very well detected faults contains the subsequent items: $f7_1$, $(f1, f7)_2$, $(f1, f2, f7, f15, f16)_3$, $(f13, f17)_4$.

In the next stage of the analysis the faults in the context of their particular classes (three types of abrupt and one type of incipient faults) were considered. Examples of obtained maps are shown in Fig. 5. Based on the obtained results two groups of faults can be distinguished. In the first of them the weakly detectable faults such as $(f8, f14)_1$, $(f8, f14)_2$, $(f8, f14)_3$, $(f5, f6, f18)_4$ are assigned. A special case of assignment was obtained for $f14$, for which the same neuron as for non-faulty condition was assigned. Further analysis (search for the differences between the non-faulty and $f14$ signals) allows to conclude that the process signals of $f14$ were identical to the signals from the non-faulty condition of a system. The second group of faults contains the faults which are well detectable while taking into account the individual classes. The faults $(f1, f7, f13, f18)_1$, $(f1, f7, f13, f18)_2$, $(f2, f7, f15)_3$ and $(f4, f13, 17)_4$ can be assigned to this subset. Furthermore, it can be seen that some faults are very similar to each other, e.g. $f13_2$ and $f18_2$ as well as $f13_3$ and $f18_3$.

The consideration of the weakly detectable faults only, such as

$(f2, f3, f5, f6, f9, f11, f18)_4$, $f8_{1,2,3}$, (the faults $f14_{1,2,3}$ are not further considered), which were excluded during the previous analysis, allows to conclude that all of them are very well isolated with respect to the non-faulty case (see Fig. 6). Only the fault $f5_4$ is located in the same cluster as non-faulty case, thus this fault is very similar to this state.

Fig. 4: Results obtained for all considered faults: (a) map size: 20×20, learning steps (WTM, WTA): (10, 10); (b) map size: 45×45, learning steps (WTM, WTA): (50, 30)

Fig. 5: Maps obtained for single classes of faults, map size: 15x15, learning steps (WTM, WTA): (30, 10): (a) large abrupt faults; (b) incipient faults

Further studies were concentrated on the sensitivity analysis with taking into consideration the shorter realizations of signals for the particular groups of faults. Two main groups: weakly detectable (listed above) and the remaining faults were considered. Exemplary maps are shown in Fig. 7. Based on the obtained results for these analyses, it is possible to conclude that the reduction of a length of signals influences on the quality of the results. The minimal period of time that makes possible the detection of selected faults equals 1 s. A vector

of five elements (five signals were considered) describes each fault in these cases. Selected groups of faults (usually the intensities of particular faults) became indistinguishable from each other, e.g. $f7_{1,2,3}$, $f10_{2,3}$, $f18_{1,2,3}$.

Final summary of faults detection with division to the possible types is presented in Table VII.

B. Validation of a methodology

Obtained SOM maps can be considered as a diagnostic pattern, based on which it is possible to conclude about technical

Fig. 6: Obtained maps for weakly detectable faults, map size: 20×20, learning steps (WTM, WTA): (20, 15)

TABLE VII: Results assigned to the possible classes of faults

Type	Faults 1100 s
undetectable	$f14_{1,2,3}$
weakly detectable (inside the classes of faults)	$f8_{1,2,3}$ $(f5, f6, f18)_4$
weakly detectable (overall)	$f5_4$
pair of not distinguishable faults	$(f13, f18)_2$ $(f13, f18)_3$
very well detectable	$f1, f7, f13, f18)_{1,2}$ $(f2, f7, f15)_3$ $(f4, f13, f17)_4$

condition of the considered system. In order to validate this approach the new three sets of data for considered 44 faults were prepared (small modification of standard DAMADICS benchmark simulator was made). Obtained data was divided into five groups. The faults, according to the classes: small, medium and large abrupt and incipient, are included in the first four of them. The fifth class includes faults, which were obtained during previous analysis and assigned as the weakly detectable (see Figure 6). From the collection of numerous maps, which were generated previously, five maps were selected according to the particular groups of faults. Basing on these maps, for subsequent cases of new data the best neuron (BMU) from suitable map was determined. The considered fault was assigned as recognizable if the assigned neuron coincided with a neuron corresponding to the same fault from the basic map, or the neuron was located in the same cluster of a map. An example of assigned labels of large abrupt faults is shown in Figure 8, while a full summary of the results is posted in Table VIII. In the last three columns, the numbers of recognized cases of faults for individual groups are presented. It can be observed that the faults $f8_{1,2,3}$ i $f14_{1,2,3}$ are weakly recognizable. The best results are achieved for large abrupt faults, where for the most cases of considered faults the same neurons are assigned similarly as in the basic map.

TABLE VIII: Results of validation of methodology

Groups of faults	Faults	1st set	2nd set	3rd set
Small abrupt	10	8	8	8
Medium abrupt	10	8	8	8
Large abrupt	14	12	12	12
Incipient	10	2	2	3
Weakly detected	10	4	5	7

V. CONCLUSIONS

In the presented paper the novel approach to the process diagnosis problem based on self-organizing maps was presented. The authors used the DAMADICS benchmark for simulation of faults of various types and intensities, which was the basis of the investigated fault diagnosis problem. Using SOM-based approach and the simulation data it was shown that the faults can be not only precisely detected, but also isolated and identified basing on resulted BMU maps. As it was noticed, applied approach is very sensitive to changes between the signals, which allow to detect, isolate and identify almost all unique cases of faults and provides better results of faults detection and isolation than the previously applied methods for the DAMADICS benchmark problem. Moreover, using SOM it is possible to compare multiple to multiple signals, which improves much the methodology of fault diagnosis due to the lack of necessity to carry out preprocessing procedures of the process signals. In order to validate the presented approach additional datasets were generated from the DAMADICS benchmark simulator, which were slightly different than those generated for the main analysis. The validation procedures show that in the most cases the faults were appropriately classified to the particular clusters.

Presented studies is a part of on-going research. Several open problems, which were planned to solve in future, should be highlighted. The improvement of the distinguishability of faults is expected during the additional preprocessing of input diagnostic signals, e.g. their normalization. Additionally, the qualitative measure (e.g. a threshold), based on which the value of a neuron on the BMU map can be used for the identification of a fault, should be developed.

The application of SOM-based algorithm in process diagnostic systems has a great potential due to the very high sensitivity to the differences in signals and datasets and could be automated and successfully adapted for the industrial installations. Since the presented algorithm is time-consuming, it cannot be applied for real-time diagnosis, however further attempts will be made to optimize the SOM-based fault diagnosis algorithm in order to increase its effectiveness.

REFERENCES

[1] M. Basseville and I. Nikiforov, "Detection of Abrupt Changes: Theory and Applications", Prentice Hall Information and Systems Science Series, 1993.
[2] M. Witczak, R.J. Patton and J. Korbicz, "Fault Detection with Observers and Genetic Programming: Application to the DAMADICS Benchmark Problem", *Proc. 5th IFAC Symp. on Fault Detection, Supervision and Safety of Technical Processes – SAFEPROCESS*, Washington, 2003, pp. 1203-1208.

Fig. 7: Maps for different times of analyzed signals, (a) well detectable faults; period of time: 1000 s, map size: 25x25, learning steps (WTM, WTA): (35, 15); (b) well detectable faults; period of time: 1 s, map size: 25×25, learning steps (WTM, WTA): (30, 15); (c) weakly detectable faults; period of time: 750 s, map size: 20x20, learning steps (WTM, WTA): (25, 15); (d) weakly detectable faults; period of time: 1 s, map size: 20×20, learning steps (WTM, WTA): (25, 10)

Fig. 8: Exemplary map of validation of proposed approach.

[3] A. Lipnickas and J. Korbicz, "Evolutionary Learning in Identification of Fuzzy Models: Application to DAMADICS Benchmark", *Proc. 6th Domestic Conf. "Diagnostics of Industrial Processes" DPP'03*, Władysławowo, 2003.

[4] J.M.F. Calado, F.P.N.F. Carreira, M.J.G.C. Mendes, J.M.G. Sá da Costa

and M. Bartyś, "Fault Detection Approach Based on Fuzzy Qualitative Reasoning Applied to the DAMADICS Benchmark Problem", *Proc. 5th IFAC Symp. on Fault Detection, Supervision and Safety of Technical Processes – SAFEPROCESS*, Washington, 2003, pp. 1179-1184.

[5] M. Witczak, J. Korbicz, M. Mrugalski and R.J. Patton, "A GMDH Neural Network-Based Approach to Robust Fault Diagnosis: Application to the DAMADICS Benchmark Problem", *Control Eng. Pract.*, vol. 14, 2006, pp. 671-683, http://dx.doi.org/10.1016/j.conengprac.2005.04.007.

[6] Y. Kourd, N. Guersi and D. Lefebvre, "Neuro-Fuzzy Approach for Default Diagnosis: Application to the DAMADICS", *Proc. 4th IEEE Conf. on Digital Ecosystems and Technologies*, Dubai, 2010, pp. 107-111, http://dx.doi.org/10.1109/DEST.2010.5610663.

[7] Y. Kourd, D. Lefebvre and N. Guersi, "Fault Diagnosis Based on Neural Networks and Decision Trees: Application to DAMADICS", *Int. J. Innov. Comput. I.*, vol. 9, 2013, pp. 3185-3195.

[8] G.M. de Almeida and S.W. Park, "Fault Detection and Diagnosis in the DAMADICS Benchmark Actuator System – a Hidden Markov Model Approach", *Proc. 17th World Congress of the IFAC*, Seoul, 2008, pp. 12419-12424, http://dx.doi.org/10.3182/20080706-5-KR-1001.2573.

[9] S. Openshaw and I. Turton, "A Parallel Kohonen Algorithm for the Classification of Large Spatial Datasets", *Comput. Geosci.*, vol. 22, 1996, pp. 1019-1026, http://dx.doi.org/10.1016/S0098-3004(96)00040-4.

[10] W. Melssen, R. Wehrens and L. Buydens, "Supervised Kohonen Networks for Classification Problems", *Chemometr. Intell. Lab.*, vol. 83, 2006, pp. 99-113, http://dx.doi.org/10.1016/j.chemolab.2006.02.003.

[11] D. Bianchi, R.Calogero and B. Tirozzi, "Kohonen Neural Networks and Genetic Classification", *Math. Comput. Model.*, vol. 45, 2007, pp. 34-60, http://dx.doi.org/10.1016/j.mcm.2006.04.004.

[12] M. Amarowicz and A. Katunin, "Clustering of Delaminations in Composite Rotors Using Self-Organizing Maps", *Intelligent Systems in Technical and Medical Diagnostics*, J. Korbicz and M. Kowal, Eds., *Advances in Intelligent Systems and Computing*, vol. 230, Berlin-Heidelberg, Springer, 2014, pp. 149–159, http://dx.doi.org/10.1007/978-3-642-39881-0_12.

[13] S.M. Bhandarkar, J. Koh and M. Suk, "Multiscale Image Segmentation Using a Hierarchical Self-Organizing Map", *Neurocomputing*, vol. 14, 1997, pp. 241-272, http://dx.doi.org/10.1016/S0925-2312(96)00048-3.

[14] C. Amerijckx, J.D. Legat and M. Verleysen, "Image Compression by Self-Organized Kohonen Map", *Syst. Anal. Model. Sim.*, vol. 43, 2003, pp. 1529-1543, http://dx.doi.org/10.1080/0232929032000115182.

[15] W.G. Teng and P.L. Chang, "Identifying Regions of Interest in Medical Images Using Self-Organizing Maps", *J. Med. Syst.*, vol. 36, 2012, pp. 2761-2768, http://dx.doi.org/10.1007/s10916-011-9752-8.

[16] A.G. de Barreto, A.F.R. Araújo and H.J. Ritter, "Self-Organizing Feature Maps for Modeling and Control of Robotic Manipulators", *J. Intell. Robot. Syst.*, vol. 36, 2003, pp. 407-450, http://dx.doi.org/10.1023/A:1023641801514.

[17] M. Johnsson and C. Balkenius, "Sense of Touch in Robots with Self-Organizing Maps", *IEEE T. Robot.*, vol. 27, 2011, pp. 498-507, http://dx.doi.org/10.1109/TRO.2011.2130090.

[18] A. Lendasse, J. Lee, V. Wertz and M. Verleysen, "Forecasting Electricity Consumption Using Nonlinear Projection and Self-Organizing Maps", *Neurocomputing*, vol. 48, 2002, pp. 299-311, http://dx.doi.org/10.1016/S0925-2312(01)00646-4.

[19] G. Simon, A. Lendasse, M. Cottrell, J.C. Fort and M. Verleysen, "Time Series Forecasting: Obtaining Long Term Trends with Self-Organizing Maps", *Pattern Recogn. Lett.*, vol. 26, 2005, pp. 1795-1808, http://dx.doi.org/10.1016/j.patrec.2005.03.002.

[20] C.M. Hsu, "A Hybrid Procedure for Stock Price Prediction by Integrating Self-Organizing Map and Genetic Programming", *Expert Syst. Appl.*, vol. 38, 2011, pp. 14026-14036, http://dx.doi.org/10.1016/j.eswa.2011.04.210.

[21] C.W. Chan, H. Jin, K.C. Cheung and H.Y. Zhang, "Fault Detection of Systems with Redundant Sensors Using Constrained Kohonen Networks", *Automatica*, vol. 37, 2001, pp. 1671-1676, http://dx.doi.org/10.1016/S0005-1098(01)00126-1.

[22] S.L. Jämsä-Jounela, M. Vermasvuori, P. Endén and S. Haavisto, "A Process Monitoring System Based on the Kohonen Self-Organizing Maps", *Control Eng. Pract.*, vol. 11, 2003, pp. 83-92, http://dx.doi.org/10.1016/S0967-0661(02)00141-7.

[23] M. Seera, C.P. Lim, D. Ishak and H. Singh, "Offline and Online Fault Detection and Diagnosis of Induction Motors Using a Hybrid Soft Computing Model", *Appl. Soft Comput.*, vol. 13, 2013, pp. 4493-4507, http://dx.doi.org/10.1016/j.asoc.2013.08.002.

[24] T. Chopra and J. Vajpai, "Classification of Faults in DAMADICS Benchmark Process Control System Using Self Organizing Maps", *Int. J. Soft Comput. Eng.*, vol. 1, 2011, pp. 85-90.

[25] M. Syfert, R. Patton, M. Bartyś and J. Quevedo, "Development and Application of Methods for Actuator Diagnosis in Industrial Control Systems (Damadics): A Benchmark Study", *Proc. 5th IFAC Symp. Fault Detection, Supervision and Safety of Technical Processes – SAFEPROCESS*, Washington, 2003, pp. 939-950.

[26] M. Bartyś, R. Patton, M. Syfert, S. de las Herras and J. Quevedo, "Introduction to the DAMADICS Actuator FDI Benchmark Study", *Control Eng. Pract.*, vol. 14, 2006, pp. 577-596, http://dx.doi.org/10.1016/j.conengprac.2005.06.015.

[27] DAMADICS, "Website of the Research Training Network on Development and Application of Methods for Actuator Diagnosis in Industrial Control Systems" [online], Institute of Automatic Control and Robotics, Warsaw University of Technology, 2004 [viewed: 2015-04-11]. Available from: http://diag.mchtr.pw.edu.pl/damadics.

[28] M. Kościelny, M. Bartyś, P. Rzepiejewski and J. Sá da Costa, "Actuator Fault Distinguishability Study for the DAMADICS Benchmark Problem", *Control Eng. Pract.*, vol. 14, 2006, pp. 645-652, http://dx.doi.org/10.1016/j.conengprac.2005.06.014.

[29] P. Supavatanakul, J. Lunze, V. Puig and J. Quevedo, "Diagnosis of Timed Automata: Theory and Application to the DAMADICS Actuator Benchmark Problem", *Control Eng. Pract.*, vol. 14, 2006, pp. 609-619, http://dx.doi.org/10.1016/j.conengprac.2005.03.028.

[30] F. Previdi and T. Parisini, "Model-Free Actuator Fault Detection Using a Spectral Estimation Approach: the Case of the DAMADICS Benchmark Problem", *Control Eng. Pract.*, vol. 14, 2006, pp. 635-644, http://dx.doi.org/10.1016/j.conengprac.2005.04.001.

[31] V. Puig, A. Stancu, T. Escobet, F. Nejjari, J. Quevedo and R.J. Patton, "Passive Robust Fault Detection Using Interval Observers: Application to the DAMADICS Benchmark Problem", *Control Eng. Pract.*, vol. 14, 2006, pp. 621–633, http://dx.doi.org/10.1016/j.conengprac.2005.03.016.

[32] C.M. Bocaniala and J.M.G. Sá da Costa, "Application of a Novel Fuzzy Classifier to Fault Detection and Isolation of the DAMADICS Benchmark Problem", *Control Eng. Pract.*, vol. 14, 2006, pp. 653-669, http://dx.doi.org/10.1016/j.conengprac.2005.06.008.

[33] D. Düştegör, E. Frisk, V. Cocquempot, M. Krysander and M. Staroświecki, "Structural Analysis of Fault Isolability in the DAMADICS Benchmark", *Control Eng. Pract.*, vol. 14, 2006, pp. 597-608, http://dx.doi.org/10.1016/j.conengprac.2005.04.008.

[34] J.M.F. Calado, J.M.G. Sá de Costa, M. Bartyś and J. Korbicz, "FDI Approach to the DAMADICS Benchmark Problem Based on Qualitative Reasoning Coupled with Fuzzy Neural Networks", *Control Eng. Pract.*, vol. 14, 2006, pp. 685-698, http://dx.doi.org/10.1016/j.conengprac.2005.03.025.

[35] A.R.C. Oliveira and J.M.G. Sá da Costa, "Hierarchic Fault Diagnosis by Pattern-Recognition Approaches Applied to DAMADICS Benchmark", *Proc. 18th IFAC World Congress*, vol. 18, Milano, 2011, pp. 7737-7742, http://dx.doi.org/10.3182/20110828-6-IT-1002.03638.

[36] J. Vesanto, "SOM Implementation in SOM Toolbox, SOM Toolbox Online Help" [Online], Laboratory of Computer and Information Science, 2005 [viewed: 2015-04-11]. Available from: http://www.cis.hut.fi/projects/somtoolbox/documentation/somalg.shtml.

[37] T. Kohonen, "Self-Organizing Maps", Springer Series in Information Sciences, vol. 30, Berlin-Heielberg, Springer, 2001.

[38] J. Vesanto, J. Himberg, E. Alhoniemi and J. Parhankangas, "SOM Toolbox for Matlab 5" [Online], Helsinki University of Technology, 2000 [viewed: 2015-04-11], Research report. Available from: http://www.cis.hut.fi/somtoolbox/package/papers/techrep.pdf.

[39] R. Isemann and P. Ballé, "Trends in the Application of Model-Based Fault Detection and Diagnosis of Technical Processes", *Control Eng. Pract.*, vol. 5, 1997, pp. 709–719, http://dx.doi.org/10.1016/S0967-0661(97)00053-1.