

Dobór funkcji aktywacji sieci neuronowej realizującej odtwarzanie wielkości wejściowej przetwornika pomiarowego

Piotr Makowski
Jerzy Roj*

W artykule przedstawiono wyniki badań wybranych struktur sieci neuronowych realizujących odtwarzanie wielkości wejściowych czujników dwuparametrowych. Poddano analizie dokładność realizacji procedury odtwarzania przez sieć neuronową w zależności od rodzaju zastosowanych funkcji aktywacji neuronów. Podjęto próbę praktycznej weryfikacji uzyskanych rezultatów wykorzystując do tego celu układ służący do pomiaru stężenia gazów.

Selection of transfer function of neural network realising reconstruction of measuring transducer input quantity. The paper presents investigation results of the selected structures of a neural network realising reconstruction of two-parameter sensor input quantities. The accuracy of reconstruction realisation by a neural network is analysed according to the type of the neuron transfer function used. The obtained results are practically verified on the example of the system for gas concentration measurement.

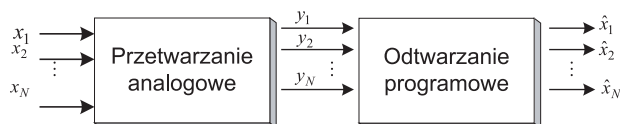
Wprowadzenie

Jednym z obszarów zastosowania sieci neuronowych może być odtwarzanie wielkości wejściowych czujników wieloparametrowych. W tego rodzaju czujnikach dokonywane jest jednoczesne przetwarzanie wielu wielkości mierzonych x_1, x_2, \dots, x_N na wielkości wyjściowe y_1, y_2, \dots, y_N , przy czym każda z wielkości wyjściowych jest zależna od wszystkich wielkości mierzonych. Modelem właściwości statycznych czujnika wieloparametrowego jest z reguły układ równań:

$$y_i = f_i(x_1, x_2, \dots, x_N), \quad (1)$$

gdzie $i = 1, 2, \dots, N$.

W celu wyznaczenia wartości wielkości mierzonych na podstawie uzyskanych wielkości wyjściowych, najczęściej stosuje się odpowiednie algorytmy przetwarzania cyfrowego, realizujące tzw. odtwarzanie wielkości wejściowych [1, 2], polegające w ogólnym przypadku na rozwiązywaniu funkcji odwrotnych do równań opisujących przetwarzanie analogowe (1). W takim przypadku proces pomiaru można przedstawić w sposób pokazany na rys.1.



Rys. 1. Proces pomiaru realizowany na zasadzie odtwarzania dla przetwornika wieloparametrowego

Dokładność realizacji pomiaru na zasadzie odtwarzania uwarunkowana jest głównie przez dokładność identyfikacji modelu matematycznego czujnika, oraz jego modelu odwrotnego. W praktyce uzyskanie odpowiednio dokładnych równań modelu czujnika jest zadaniem trudnym i wymagającym najczęściej przeprowadzenia żmudnego procesu wzorcowania. Również złożoność numeryczna algorytmów odtwarzania rośnie wraz ze wzrostem liczby wielkości mierzonych. W tej sytuacji wydaje się, że realizacja odtwarzania przez sieć neuronową w opisanych warunkach może być zasadna. Oczywiście jest, że zastosowanie sieci neuronowej nie eliminuje konieczności wzorcowania czujnika, jednak pozwala uniknąć budowania modelu odwrotnego oraz programowej implementacji jego równań. Właściwości aproksymacyjne sieci neuronowych pozwalają również przypuszczać, że pro-

* Mgr inż. Piotr Makowski, dr inż. Jerzy Roj – Politechnika Śląska, Instytut Metrologii i Automatyki Elektrotechnicznej

ces wzorcowania czujnika może obejmować mniejszą liczbę punktów pomiarowych, niż w przypadku zastosowania klasycznych algorytmów odtwarzania, na założonym poziomie dokładności.

W ogólnym przypadku trudno wskazać jednoznaczna metodę doboru struktury sieci neuronowej do realizacji określonego zadania pomiarowego [3, 4]. Podobnie rzecz się ma z doбором funkcji aktywacji poszczególnych warstw neuronów sieci. Można podjąć próbę zbadania dokładności realizacji procedury odtwarzania przez określonego typu sieć neuronową dla różnych funkcji aktywacji przy założeniu, że model matematyczny właściwości statycznych czujnika wieloparametrowego stanowi pewna klasa funkcji. Z tego względu dalsze rozważania ograniczono do czujnika dwuparametrowego przy założeniu, że jego model matematyczny w warunkach statycznych stanowi następujący przykładowy układ równań wykładniczych:

$$\begin{cases} y_1 = 0,8 \cdot e^{-5x_1} + 0,2 \cdot e^{-x_2} \\ y_2 = 0,3 \cdot e^{-2x_1} + 0,7 \cdot e^{-4x_2} \end{cases} \quad (2)$$

Badania wybranych struktur sieci neuronowych

Badania przeprowadzono dla sieci neuronowej jednokierunkowej, uczonej za pomocą algorytmu Levenberga-Marquardta, będącego modyfikacją algorytmu wstecznej propagacji błędów [4, 5, 6]. Programową realizację sieci wykonano w oparciu o program MatLab wykorzystując bibliotekę Neural Network Toolbox. Dla tego typu sieci program umożliwia zastosowanie następujących funkcji aktywacji:

- sigmoidalnej unipolarnej:

$$U = f(\zeta) = \frac{1}{1 + e^{-\zeta}} \quad (3)$$

- sigmoidalnej bipolarnej:

$$B = f(\zeta) = \frac{2}{1 + e^{-2\zeta}} - 1 \quad (4)$$

- liniowej:

$$L = f(\zeta) = \zeta \quad (5)$$

W pierwszym etapie badań, w oparciu o twierdzenie Kołmogorowa [6], przyjęto sieć o 5 neuronach w warstwie ukrytej, dla której przyjęto oznaczenie 2-5-2. W procesie uczenia sieci wykorzystano 1000-elementowy zbiór wzorców uzyskany na podstawie równań (2) modelu czujnika. Jako miarę stopnia wytrenowania sieci wykorzystano wartość sumarycznego błędu na jej wyjściu, obliczoną z zależności:

$$\delta_{SSE} = \sum_{i=1}^{i=W} \sum_{j=1}^{j=M} (\hat{x}_j^i - x_j^i)^2 \quad (6)$$

gdzie \hat{x}_j^i jest wartością wyjściową sieci, x_j^i - wartością poprawną, W - liczbą elementów zbioru wzorców, M - liczbą neuronów warstwy wyjściowej.

Należy zaznaczyć, że skuteczność uczenia sieci zależy m.in. od wartości początkowych współczynników wagowych, które zostają zainicjowane w sposób losowy. Zatem dla tej samej sieci neuronowej i zbioru uczącego uzyskuje się różne wartości błędu δ_{SSE} na jej wyjściu. W związku z tym kilkunastokrotnie powtarzano proces uczenia sieci, w wyniku czego otrzymano zbiór wartości błędów, posiadający powtarzalne minimum, które ostatecznie przyjęto do oceny jakości nauczania sieci (proces uczenia przerywano w momencie, gdy wartość błędu na wyjściu sieci przez dłuższy czas nie ulegała zmianie).

W tabeli 1 zestawiono wartości błędu δ_{SSE} na wyjściu sieci neuronowej w zależności od rodzaju zastosowanych funkcji aktywacji. Pierwsza z dwóch liter symbolizujących funkcje aktywacji sieci neuronowej dotyczy neuronów warstwy ukrytej, druga - neuronów warstwy wyjściowej.

Na podstawie wyników zestawionych w tabeli 1 można jedynie wnioskować o dokładności, z jaką sieć realizuje odtwarzanie dla zbioru uczącego, w zależności od zastosowanych funkcji aktywacji. Uzyskane wyniki wskazują, że oprócz przypadków, w których zastosowano liniowe funkcje aktywacji w warstwie środkowej, praktycznie nie ma znaczenia rodzaj funk-

Tab. 1. Wartości błędu δ_{SSE} na wyjściu sieci neuronowej o strukturze 2-5-2, uzyskane w procesie uczenia, w zależności od rodzaju zastosowanych funkcji aktywacji

Funkcje aktywacji	L-L	L-U	L-B	U-L	U-U	U-B	B-L	B-U	B-B
δ_{SSE} [1]	43,69	-	-	5,33	5,29	5,85	5,33	5,29	5,85

Tab. 2. Wartości błędu δ_{SSE} na wyjściu sieci neuronowej o strukturze 2-5-2, uzyskane w procesie testowania, dla różnych funkcji aktywacji

Funkcje aktywacji	U-L	U-U	U-B	B-L	B-U	B-B
δ_{SSE} [1]	0,36	0,34	0,39	0,36	0,34	0,39

cji aktywacji wykorzystanych w tego rodzaju sieci ze względu na skuteczność procesu nauki sieci. Powyższe rozważania potwierdziły wyniki uzyskane podczas testowania sieci, polegającego na podawaniu na jej wejścia wartości ze 100-elementowego zbioru testującego, które zestawiono w tabeli 2.

W kolejnym etapie poddano badaniom sieć neuronową o dwóch warstwach ukrytych składających się z 9 neuronów każda. Dla tej struktury sieci przyjęto oznaczenie 2-9-9-2. Ucznienie sieci przeprowadzono w sposób analogiczny jak dla sieci o strukturze 2-5-2, w szczególności wykorzystano te same zbiory uczący i testujący. W tabeli 3 zestawiono wartości błędów δ_{SSE} sieci 2-9-9-2 uzyskanych w procesie jej uczenia dla różnych funkcji aktywacji w poszczególnych warstwach. Pewne kombinacje funkcji aktywacji (funkcje liniowe w warstwach ukrytych) uniemożliwiały proces uczenia sieci, stąd nie zostały ujęte w tabeli.

Z danych zestawionych w tabeli 2 wynika, że najlepsze efekty uczenia uzyskano dla sigmoidalnych unipolarnych funkcji aktywacji w warstwach ukrytych oraz liniowej funkcji aktywacji w warstwie wyjściowej (U-U-L), natomiast najgorsze dla sigmoidalnych unipolarnych funkcji aktywacji w pierwszej warstwie ukrytej i warstwie wyjściowej oraz bipolarnej sigmoidalnej funkcji aktywacji w drugiej warstwie ukrytej (U-B-U). Skuteczność uczenia sieci dla wymienionych przypadków potwierdziły badania testowe, w wyniku których otrzymano:

$\delta_{SSE} = 0,0033$ dla funkcji aktywacji U-U-L oraz $\delta_{SSE} = 0,0078$ dla funkcji U-B-U.

Uzyskane rezultaty badań poddano praktycznej weryfikacji wykorzystując do tego celu układ pomiarowy służący do wyznaczania stężenia gazów.

Układ pomiaru stężenia gazów

W układzie pomiarowym służącym do wyznaczania stężenia gazów zastosowano rezystancyjne czujniki stężenia gazów oparte na dwutlenku cyny SnO₂ [7, 8]. Ich cechą charakterystyczną jest ograniczona selektywność, tzn. jednoczesne reagowanie na kilka gazów, przy czym czułości czujnika na każdą z rozpo-

znawanych substancji są różne. Na rys. 2 przedstawiono charakterystyki statyczne czujnika typu TGS813, służącego do pomiaru stężenia metanu.

W przypadku tego czujnika, obecność tlenu węgla wpływa w sposób zakłócający na wynik pomiaru. Analogicznie obecność metanu jest czynnikiem zakłócającym w przypadku pomiarów stężeń tlenu węgla realizowanych przy użyciu czujnika o oznaczeniu TGS822. W związku z powyższym przyjęto koncepcję realizacji układu pomiarowego stężenia gazów (CH₄, CO) przedstawioną na rys.3.

Dokładność realizacji odtwarzania stężenia gazów przez sieć neuronową w zależności od funkcji aktywacji neuronów

Przyjęto realizację fazy odtwarzania przez jednokierunkową sieć neuronową o strukturze 2-5-2. W celu przeprowadzenia stosownych badań przygotowano zbiór 131 wzorcowych mieszanin gazowych, złożonych z powietrza, do którego dozowano jednocześnie dwa polutanty: metan oraz tlenek węgla. Drogą losowania wybrano 96 mieszanin, tworząc tym samym zbiór uczący, zaś pozostałych 35 mieszanin utworzyło zbiór testowy. W tabeli 4 zestawiono wartości sumarycznego błędu uczenia sieci δ_{SSE} w zależności od zastosowanych funkcji aktywacji.

Dane z tabeli 4 wskazują, że podobnie jak w przypadku badań modelu czujnika opisanego układem równań (2) dla sieci neuronowej o strukturze 2-5-2 praktycznie nie ma znaczenia rodzaj zastosowanych funkcji aktywacji biorąc pod uwagę skuteczność uczenia sieci. Dokonując następnie procesu testowania sieci wyznaczono następujące wartości:

- średni błąd kwadratowy oznaczeń poszczególnych składników mieszanin testowych [9] wyznaczony zgodnie z zależnością

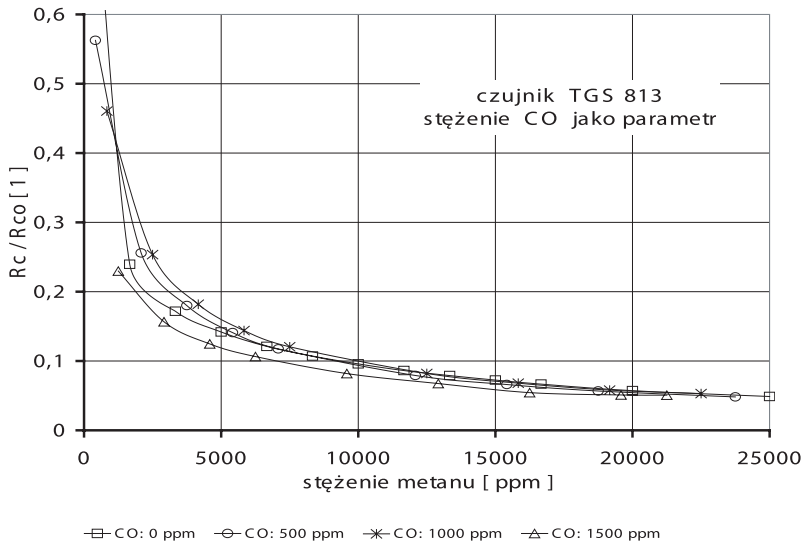
$$\delta_{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^{i=N} (c_{pi} - c_{ri})^2} \tag{7}$$

Tab.3. Wartości błędu δ_{SSE} na wyjściu sieci neuronowej o strukturze 2-9-9-2, uzyskane w procesie uczenia, w zależności od rodzaju zastosowanych funkcji aktywacji

Funkcje aktywacji	U-U-U	U-U-B	U-U-L	U-B-U	U-B-B	U-B-L	B-B-B	B-B-U	B-B-L	B-U-U	B-U-B	B-U-L
$\delta_{SSE}[1]$	0,18	0,15	0,12	0,21	0,20	0,16	0,15	0,13	0,15	0,15	0,20	0,13

Tab. 4. Wartości błędu δ_{SSE} na wyjściu sieci neuronowej o strukturze 2-5-2, uzyskane w procesie uczenia, dla różnych funkcji aktywacji.

Funkcje aktywacji	U-L	U-U	U-B	B-L	B-U	B-B
$\delta_{SSE}[1]$	0,55	0,62	0,66	0,53	0,58	0,60



Rys. 2. Rodzina charakterystyk statycznych czujnika TGS813, przy czym: R_c – rezystancja czujnika w powietrzu skażonym, R_{co} – rezystancja czujnika w powietrzu czystym

gdzie: c_{pi} – rozpoznawana przez sieć koncentracja polutantu, c_{ri} – rzeczywista koncentracja polutantu w mieszaninie, $N = 35$, oraz – wartość unormowaną błęd δ_{RMS} w oparciu o wyrażenie:

$$\delta_{RMS}^* = \frac{\delta_{RMS}}{c_{gazmax}} \quad (8)$$

gdzie: c_{gazmax} – zakres badanych stężeń danego polutantu, tj.: 25000 ppm dla metanu oraz 1500 ppm w przypadku tlenku węgla.

Uzyskane wyniki, potwierdzające wcześniejsze spostrzeżenia, zamieszczono w tabeli 5.

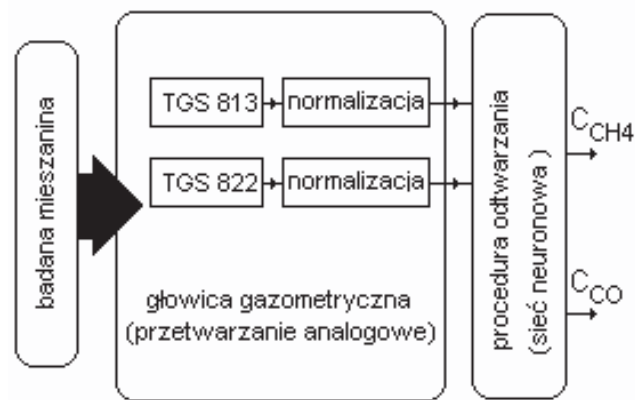
Na rys. 4 i rys. 5. przedstawiono przykładowe wartości błędów odtwarzania przez sieć neuronową o strukturze 2-5-2, odpowiednio stężenia metanu oraz tlenku węgla.

Z przeprowadzonych badań wynika, iż w rozpatrywanym przypadku wybór funkcji aktywacji nie ma istotnego wpływu na właściwości sieci. Zarówno błędy uczenia, jak i testowania sieci mają porównywalne wartości. Uznając dokładność rozpoznawania wzorców testujących za niewystarczającą, zmodyfikowano strukturę sieci do postaci 2-9-9-2, dla której przeprowadzono analogiczne badania. Ich wyniki wskazały na nadmiarowość założonej struktury sieci – wprawdzie otrzymano błędy uczenia o wartościach o rząd mniejszych niż w przypadku sieci 2-5-2, jednak błędy testowania

przyjmowały wartości wyższe o rząd i więcej. Przyczyn takiego stanu rzeczy należy się dopatrywać w zbyt małej liczbie wzorców uczących. Sieć o strukturze 2-9-9-2 „lepiej” dostosowywała swoje wagi do zbioru uczącego, jednak w procesie testowania wykazywała większe błędy niż strukturalnie prostsza sieć 2-5-2. W związku z powyższym praktyczna weryfikacja dokładności realizacji odtwarzania stężenia gazów przez sieć o strukturze 2-9-9-2 zostanie przeprowadzona po zrealizowaniu dodatkowych pomiarów mających na celu zwiększenie liczebności zbioru uczącego.

Uwagi końcowe

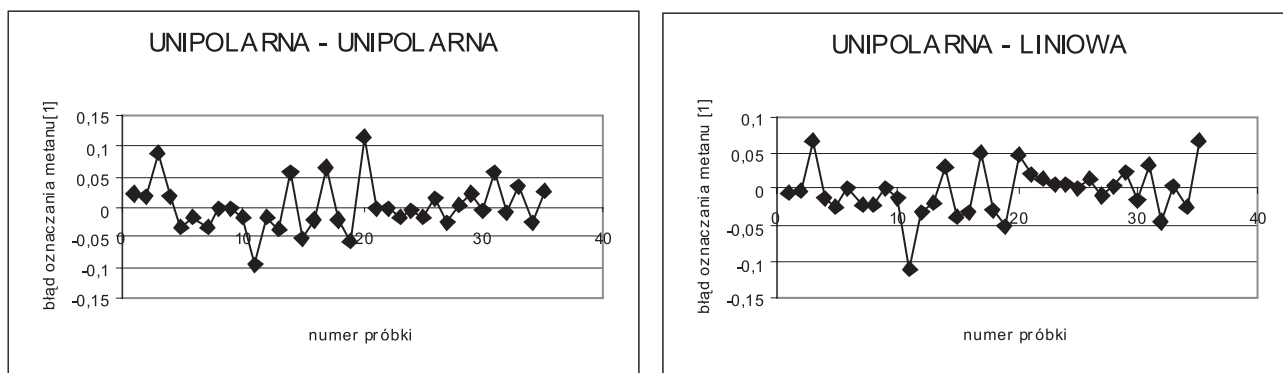
W oparciu o przeprowadzone badania dotyczące zastosowania sieci neuronowej w przypadku modelu czujnika dwuparametrowego opisanego układem równań (2) można sformułować następujące spostrzeżenia. Dla prostej struktury sieci z jedną warstwą ukrytą o 5 neuronach, praktycznie nie ma znaczenia rodzaj zastosowanych funkcji aktywacji ze względu na dokładność realizacji odtwarzania wielkości wejściowych. Pomijając przypadki funkcji liniowych w warstwie ukrytej, które uniemożliwiały proces uczenia, uzyskiwano porównywalne wartości sumarycznego błęd sieci zarówno w procesie jej uczenia jak i testowania. Natomiast w przypadku bardziej złożonej struktury sieci, o dwóch warstwach



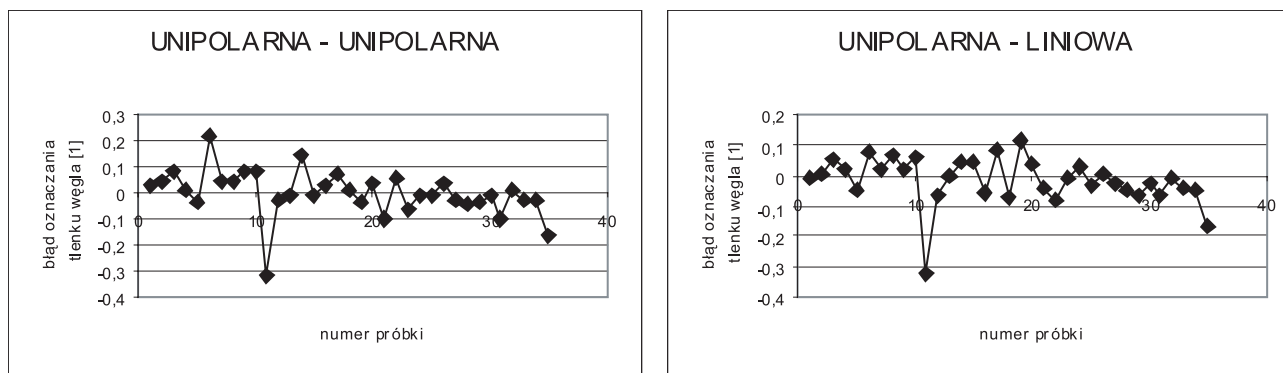
Rys. 3. Koncepcja układu pomiarowego stężenia gazów

Tab. 5. Wartości błędów δ_{RMS} , δ_{RMS}^* na wyjściu sieci neuronowej o strukturze 2-5-2, uzyskane w procesie testowania, dla różnych funkcji aktywacji.

Funkcje aktywacji		U-L	U-U	U-B	B-L	B-U	B-B
δ_{RMS} [ppm]	CH ₄	874,45	1009,72	1198,36	941,93	1000,38	1251,07
	CO	118,89	132,42	117,67	113,69	135,04	114,89
δ_{RMS}^* [1]	CH ₄	0,035	0,041	0,048	0,038	0,041	0,051
	CO	0,079	0,089	0,079	0,076	0,091	0,077



Rys. 4. Błędy odtwarzania stężenia metanu przez sieć o strukturze 2-5-2 dla przykładowych funkcji aktywacji.



Rys. 5. Błędy odtwarzania stężenia tlenku węgla przez sieć o strukturze 2-5-2 dla przykładowych funkcji aktywacji.

ukrytych składających się z 9 neuronów każda, można zauważyć istotne różnice w wartościach sumarycznego błędu sieci w zależności od zastosowanych funkcji aktywacji. Najlepsze efekty zarówno podczas uczenia jak i testowania sieci uzyskano stosując funkcje aktywacji U-U-L, natomiast ok. dwukrotnie większe wartości błędu otrzymano dla funkcji U-B-U. Próba weryfikacji otrzymanych rezultatów przy wykorzystaniu rzeczywistych danych pomiarowych pokazała jak ważna jest wzajemna współzależność liczności zbioru uczącego i złożoności struktury sieci neuronowej. Dysponując mało licznym zbiorem uczącym można „dobrze” nauczyć sieć neuronową o złożonej strukturze rozpoznawania wzorców z tego zbioru, jednak dla wartości spoza niego, sieć może generować błędy grube.

Bibliografia

- [1] Morawski R. Z.: Metody odtwarzania sygnałów pomiarowych. Monografia 1. Metrologia i Systemy Pomiarowe, Warszawa 1989.
- [2] Jakubiec J.: Bieżące programowe odtwarzanie wartości chwilowych dynamicznych przebiegów wejściowych nieliniowych przetworników pomiarowych. Monografia. ZN Pol. Śl. Seria Elektryka, z. 111, Gliwice 1988.
- [3] Hertz J., Krogh A., Palmer R.G.: Wstęp do teorii obliczeń neuronowych. Wyd. II, WNT, Warszawa 1995.
- [4] Korbicz J., Obuchowski A., Uciński D.: Sztuczne sieci neuronowe. Podstawy i zastosowania. Akademicka Oficyna Wydawnicza PLJ, Warszawa, 1994.
- [5] Tadeusiewicz R.: Sieci Neuronowe. Akademicka Oficyna Wydawnicza RM, Warszawa, 1993.
- [6] Osowski S.: Sieci neuronowe w ujęciu algorytmicznym. WNT, Warszawa 1996.
- [7] Torbicz W., Brzózka Z.: Czujniki chemiczne i bioczuJNIKI. Polskie Towarzystwo Techniki Sensorowej, Warszawa 1995.
- [8] Lach M.: Detektor gazu. Elektronika Praktyczna Nr 9, str. 47-50, 1996.
- [9] Brudzewski K.: Rozpoznawanie sygnałów z macrycz czujników z wykorzystaniem sztucznych sieci neuronowych. Problemy Eksploatacji - Zeszyty Naukowe Instytutu Technologii Eksploatacji w Radomiu, zeszyt Nr 1, s. 63-69, 1996.