

Konrad ZAJKOWSKI, Stanisław DUER

BUDOWA SZTUCZNEJ SIĘCI NEURONOWEJ DO IDENTYFIKACJI PARAMETRÓW MODELU ODBIORNIKA ELEKTRYCZNEGO

Streszczenie

W artykule przedstawiono sposób wyznaczania parametrów modelu odbiornika elektrycznego w postaci liniowego, aktywnego dwójnika, bazujący na sztucznej sieci neuronowej. Zadaniem sieci jest wyznaczenie rozwiązania N równań o N niewiadomych. Algorytm ten umożliwi wyznaczenie rozwiązania w sytuacji gdy współczynniki A_i w równaniach obarczone są błędami pomiarowymi. Dla pewnych wartości A_i (gdzie $i = 1, \dots, N$) nie istnieją funkcje odwrotne równań wejściowych. W tym przypadku niemożliwe jest wyznaczenie rozwiązania układu równań metodami klasycznymi.

WSTĘP

Wyznaczając model elektryczny dowolnego liniowego dwójnika przeprowadza się aproksymację szeregu zjawisk fizycznych w tym obiekcie, a wynik przedstawia się zazwyczaj w postaci szeregowego połączenia liniowych dyskretnych elementów elektrycznych. W sytuacji, gdy analizowanym obiektem jest uzwojenie silnika indukcyjnego, to oprócz elementów R , L i C w modelu pojawi się również źródło elektryczne modelujące zachowanie się parametrów elektrycznych po wyłączeniu silnika od źródła zasilania. Sytuację tą opisano w pozycjach [3,4,5]. Do pełnego opisu takiego modelu wymagana jest identyfikacja parametrów x_i tych elementów (1). Najczęściej tą identyfikację przeprowadza się pośrednio poprzez pomiary wielkości elektrycznych A_i na badanym obiekcie, przykładowo stosując układy do pomiarów metodami technicznymi [5]. Poszukiwane parametry wyznacza się wówczas z zależności matematycznych (1) opisujących obiekt.

$$\left. \begin{aligned} A_1 &= f_1(y_1, y_2, \dots, y_i, \dots, y_N) \\ A_2 &= f_2(y_1, y_2, \dots, y_i, \dots, y_N) \\ &\vdots \\ A_i &= f_i(y_1, y_2, \dots, y_i, \dots, y_N) \\ &\vdots \\ A_N &= f_N(y_1, y_2, \dots, y_i, \dots, y_N) \end{aligned} \right\} \quad (1)$$

gdzie: f_i – pewne funkcje zależne od modelu
 A_i – wartości mierzone
 y_i – parametry opisujące badany model

Zazwyczaj rozwiązując równanie (1) wyznacza się funkcje odwrotne g_i do funkcji f_i (2). Niedokładności pomiarowe zawarte w A_i przenoszą się wówczas na wyznaczone wielkości y_i .

$$\left. \begin{aligned} y_1 &= g_1(A_1, A_2, \dots, A_N) \\ y_2 &= g_2(A_1, A_2, \dots, A_N) \\ &\vdots \\ y_N &= g_N(A_1, A_2, \dots, A_N) \end{aligned} \right\} \quad (2)$$

Obszar rozwiązań równania (1) istnieje dla pewnej wielowymiarowej przestrzeni w układzie A_i . Współczynniki A_i obarczone błędami pomiarowymi mogą wychodzić poza tę przestrzeń, powodując, że funkcje g_i (2) nie istnieją bądź są niewyznaczalne. W tym przypadku, pomimo faktu fizycznej realizowalności modelu teoretycznego, praktycznie nie udaje się przeprowadzić parametryzacji modelu [5]. Nie istnieją wówczas funkcje odwrotne g_i dla przyjętych współczynników A_i .

Spełnienie równania (1) oznacza również słuszność następującej zależności:

$$|A_i - f_i(y_1, y_2, \dots, y_N)| = 0 \quad (3)$$

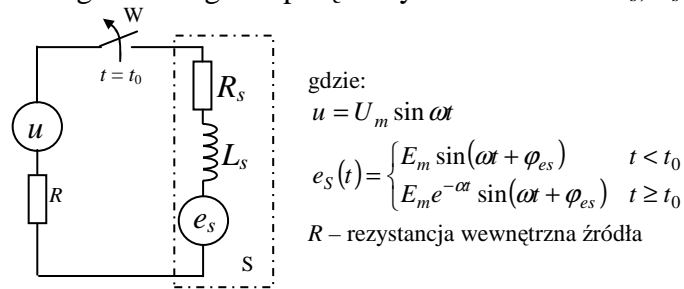
Równania (2) są wyznaczalne dla wartości z otoczenia punktów A_i obarczonych błędami pomiarowymi. W takim przypadku poszukuje się rozwiązań przybliżonych do (3), spełniających zależność (4):

$$|A_i - f_i(y_1, y_2, \dots, y_N)| \approx 0 \quad (4)$$

Oznacza to, że rozwiązanie będzie leżeć w pobliżu wielkości A_i .

1. MODEL JEDNEJ FAZY SILNIKA INDUKCYJNEGO

Do analizy przyjęto jedną fazę silnika indukcyjnego. Celem analizy jest wyznaczenie zależności prądowo-napięciowych na zaciskach jednej fazy silnika. Zależności te wyznaczyć można z modelu zbudowanego z szeregowo połączonych elementów: R_s , L_s i e_s (rys.1).



Rys.1. Model elektryczny silnika i źródła zasilania

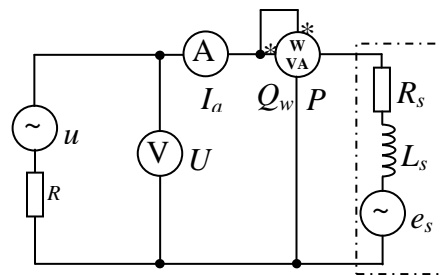
Poszukiwane współczynniki R_s , L_s , E_m , φ_{es} , α reprezentują szereg zjawisk zachodzących w silniku i układzie napędowym. Przykładowo bezwładność wirnika i układu napędowego będzie wpływać na zachowanie się źródła e_s , zaś prędkość kątowna na indukcyjności wzajemne, które są opisane przez L_s . Przy poszukiwaniu parametrów modelu ważny też jest fakt, że współczynników tych nie można wyznaczać przy zatrzymanym silniku. Oznacza to, że R_s nie odzwierciedla rezystancji uzwojeń, a L_s ich indukcyjności. Parametry modelu są określone dla stałego obciążenia na wale maszyny i stałych obrotów. Przy zmianie obciążenia, parametry modelu również ulegną zmianie.

Przedstawiony model jest przybliżeniem wielu zjawisk występujących w rzeczywistym silniku. Pomija on przykładowo oddziaływanie innych faz na analizowaną. Brak jest również odzwierciedlenie zjawisk falowych występujących w uzwojeniu traktowanym jak linia długa. Model taki może być przydatny przy wyznaczaniu zależności prądowo-napięciowych na jednej fazie silnika dla stałej częstotliwości i o małej jej wartości.

Wyznaczanych parametrów nie można w żaden sposób zunifikować dla konkretnej maszyny. Należy je wyznaczyć dla konkretnego zespołu napędowego w konkretnej sytuacji obciążeniowej. Parametry te mogą znacząco się różnić dla tego samego silnika przy różnych obciążeniach mechanicznych wału napędowego silnika.

Model taki nie uwzględnia wpływu dynamicznych zmian obciążenia mechanicznego. Jest on zatem słuszny jedynie w wąskim obszarze zastosowań.

Identyfikacja parametrów modelu polega na poszukiwaniu E_m , φ_{es} , R_s i L_s . Parametry te wyznacza się na odbiorniku [4,5] poprzez dokonanie pomiarów metodą techniczną w stanie ustalonym (praca ze stałym obciążeniem i stałą prędkością) w układzie z rys.2.



Rys.2. Obwód pomiarowy do identyfikacji parametrycznej

Zgodnie z przyjętym modelem wiadomo, że:

$$e_s = E_s \sqrt{2} \sin(\omega t + \varphi_{es}). \quad (5)$$

W dziedzinie liczb zespolonych odpowiada to postaci:

$$\underline{E}_s = E_s e^{j\varphi_{es}}, \quad \underline{U} = U e^{j\varphi_u}. \quad (6)$$

Równanie napięciowe dla oczka jest postaci:

$$\underline{I}_a (R_s + R + jX_s) + \underline{E}_s - \underline{U} = 0, \quad (7)$$

gdzie $X_s = \omega L_s$.

Przekształcając (7), wyznacza się prąd \underline{I}_a :

$$\underline{I}_a = \frac{\underline{U} - \underline{E}_s}{R_s + R + jX_s} \quad (8)$$

Woltomierz V (rys.2) mierzy różnicę napięcia zasilającego i spadku napięcia na rezystancji wewnętrznej źródła R. Zatem, w dziedzinie liczb zespolonych można zapisać:

$$\underline{U}_V = \underline{U} - \underline{I}_a R. \quad (9)$$

Z równań (8) i (9) otrzymać można:

$$\underline{U}_v = \frac{1}{R_s + R + jX_s} [\underline{U}(R_s + jX_s) + \underline{E}_s R]. \quad (10)$$

Wiedząc, że odpowiednio moce i prąd są równe:

$$P_w = \text{Re}(\underline{U}_v \underline{I}_a^*), \quad Q_w = \text{Im}(\underline{U}_v \underline{I}_a^*), \quad \underline{I}_a^* = \frac{\underline{U}^* - \underline{E}_s^*}{R_s + R - jX_s} \quad (11)$$

otrzymuje się następujące równania:

$$\begin{cases} I_a = \left| \frac{\underline{U} - \underline{E}_s}{R_s + R + jX_s} \right| \\ U_v = \left| \frac{1}{R_s + R + jX_s} [\underline{U}(R_s + jX_s) + \underline{E}_s R] \right| \\ P_w = \text{Re} \left\{ \frac{[\underline{U}(R_s + jX_s) + \underline{E}_s R] \cdot [\underline{U}^* - \underline{E}_s^*]}{(R_s + R)^2 + X_s^2} \right\} \\ Q_w = \text{Im} \left\{ \frac{[\underline{U}(R_s + jX_s) + \underline{E}_s R] \cdot [\underline{U}^* - \underline{E}_s^*]}{(R_s + R)^2 + X_s^2} \right\} \end{cases} \quad (12)$$

Równanie (12) jest szczególnym przypadkiem równania (1). Poszukiwane parametry modelu odbiornika, otrzymywane z powyższego układu równań, nie są wyznaczalne dla całej przestrzeni parametrów wejściowych (U_v , I_a , P_w , Q_w). Istnieją obszary, wynikające z niedokładności pomiarowych, w których układ równań (12) nie ma rozwiązań.

Sprawdzono, że współczynników tych nie udaje się również wyznaczyć za pomocą algorytmu interpolacji Newtona. Nie istnieją również funkcje odwrotne do równań (12).

Poszukiwaną stałą czasową przebiegu $1/\alpha$, związaną głównie z bezwładnością wirnika i układu napędzanego, można wyznaczyć doświadczalnie, obserwując przebieg napięcia w funkcji czasu na zaciskach silnika tuż po komutacji.

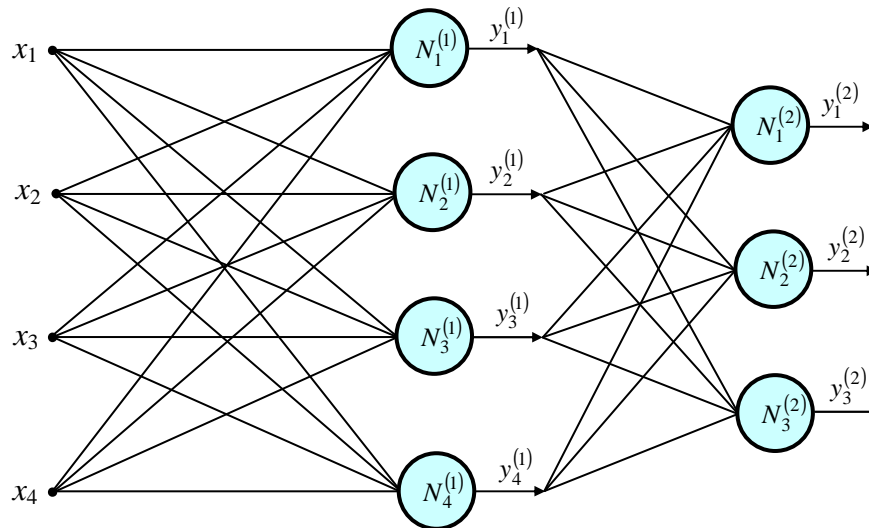
W pracy [5] udowodniono, że amplituda E_s może być równa amplitudzie U oraz częstotliwość E_s jest zbliżona do częstotliwości napięcia sieci. Zauważono też, że przesunięcie fazowe φ_{es} jest równe 0.

W przyjętym modelu zakłada się, że częstotliwości obydwu wymuszeń są identyczne. Jest to założenie nie wpływające zasadniczo na wyniki symulacji.

2. BUDOWA SIECI NEURONOWEJ

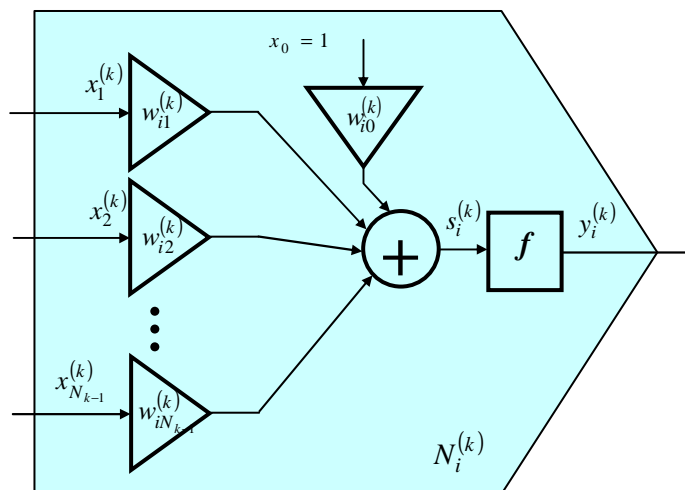
Wyznaczenia współczynników E_m , R_s i L_s z równań (12) można dokonać przy wykorzystaniu sieci neuronowej. Parametry wejściowe sieci: $x_1 = U_v$ [V], $x_2 = I_a$ [A], $x_3 = P_w$ [W], $x_4 = Q_w$ [VAr] zawierają błędy pomiarowe. Ze względu na charakter przyjętej funkcji aktywacji [2], sygnał wyjściowy neuronu warstwy wyjściowej musi zawierać się w przedziale $y \in (0, 1)$. Spełnione jest to dla wartości wyjściowych przedstawionych w odpowiednich skalach: $y_1 = R_s$ [kΩ], $y_2 = L_s$ [H], $y_3 = E_s$ [kV].

Po kilku testach sieci neuronowych zdecydowano się na budowę sieci neuronowej o topologii (rys.3) z jedną warstwą ukrytą. Wagi neuronów wyznaczane są metodą wstecznej propagacji błędów.



Rys.3. Struktura sieci neuronowej

Poszczególne neurony sieci posiadają strukturę zgodną z rys.4.



Rys.4. Schemat neuronu $N_i^{(k)}$

W budowanej sieci przyjęto następujące oznaczenia:

- t – krok iteracji, $t = 1, 2, \dots$
- $y_i^{(k)}(t)$ – i -ty sygnał wyjściowy z neuronu $N_i^{(k)}$
- $y_i^{(L)}(t)$ – i -ty sygnał wyjściowy z sieci
- k – warstwa sieci, $k = 1, \dots, L$
- L – warstwa wyjściowa sieci, liczba warstw sieci
- i – numer neuronu w warstwie, $i = 1, \dots, N_k$
- $x_j^{(k)}(t)$ – sygnał wejściowy w k -tej warstwie
- $x_j(t) = x_j^{(1)}(t)$ – wejścia sieci
- j – numer sygnału wejściowego w k -tej warstwie, $j = 1, \dots, N_{k-1}$
- N_0 – liczba wejść do sieci
- N_k – liczba neuronów w k -tej warstwie
- N_L – liczba neuronów w warstwie ostatniej

- $s_i^{(k)}(t)$ – potencjał membranowy neuronu $N_i^{(k)}$ w k -tej warstwie
- $w_{i,j}^{(k)}(t)$ – waga j -tego wejścia i -tego neuronu $N_i^{(k)}$ w k -tej warstwie
- $d_i^{(L)}(t)$ – i -ty sygnał wzorcowy wyjściowy z wektora uczącego
- $\varepsilon_i^{(L)}(t)$ – błąd i -tego wyjścia sieci $\varepsilon_i^{(L)}(t) = d_i^{(L)}(t) - y_i^{(L)}(t)$
- η – współczynnik uczenia sieci
- $Q(t)$ – błąd na wyjściu sieci dla jednego wzorca
- $Q^*(t)$ – błąd na wyjściu sieci dla całej epoki

Sygnał wyjściowy neuronu $N_i^{(k)}$ (rys.4) w chwili t , jest opisany zależnością:

$$y_i^{(k)}(t) = f(s_i^{(k)}(t)) \quad (13)$$

przy czym potencjał membranowy $s_i^{(k)}$ jest równy:

$$s_i^{(k)}(t) = \sum_{j=0}^{N_{k-1}} w_{i,j}^{(k)}(t) \cdot x_j^{(k)}(t) \quad (14)$$

Sygnał wejściowy neuronu dla warstwy $k=1$ jest równy wejściom sieci $x_j(t) = x_j^{(1)}(t)$. Każda warstwa posiada jedno wejście o stałym potencjale $x_0^{(k)}(t) = 1$. Pozostałe wejścia są wyjściami warstwy poprzedniej.

$$x_j^k(t) = \begin{cases} x_j(t) & \text{for } k=1 \\ y_j^{k-1}(t) & \text{for } k=2, \dots, L \\ 1 & \text{for } j=0, k=1, \dots, L \end{cases} \quad (15)$$

W celu budowy algorytmu uczącego sieć należy zdefiniować miarę błędu. Zdecydowano się na algorytm wstecznej propagacji błędów. Miarą błędu w tym algorytmie jest funkcja, w której jako zmienne występują wszystkie wagi wielowarstwowej sieci. Funkcję tą oznaczono przez $Q(\mathbf{w})$, gdzie \mathbf{w} oznacza wektor wszystkich wag sieci. Celem algorytmu będzie zatem znalezienie minimum funkcji Q względem wektora \mathbf{w} . Dokonać tego można pośrednio, wyznaczając rozwinięcie funkcji Q w szereg Taylora w najbliższym sąsiedztwie znanego aktualnego rozwiązania \mathbf{w} [2] wzdłuż pewnego wektora kierunkowego \mathbf{p} :

$$Q(\mathbf{w} + \mathbf{p}) = Q(\mathbf{w}) + [\mathbf{g}(\mathbf{w})]^T \mathbf{p} + 0,5\mathbf{p}^T \mathbf{H}(\mathbf{w})\mathbf{p} + \dots \quad (16)$$

gdzie: $\mathbf{g}(\mathbf{w}) = \left[\frac{\partial Q(t)}{\partial w_1(t)}, \frac{\partial Q(t)}{\partial w_2(t)}, \frac{\partial Q(t)}{\partial w_3(t)}, \dots, \frac{\partial Q(t)}{\partial w_n(t)} \right]^T$ jest wektorem gradientu,

$$\mathbf{H}(\mathbf{w}) = \begin{bmatrix} \frac{\partial^2 Q}{\partial w_1 \partial w_1} & \dots & \frac{\partial^2 Q}{\partial w_1 \partial w_n} \\ \vdots & & \vdots \\ \frac{\partial^2 Q}{\partial w_n \partial w_1} & \dots & \frac{\partial^2 Q}{\partial w_n \partial w_n} \end{bmatrix} \text{ to macierz drugich pochodnych (hesjan).}$$

Modyfikację wag algorytm dokonuje w następujący sposób:

$$\mathbf{w}(t+1) = \mathbf{w}(t) + \eta(t)\mathbf{p}(t) \quad (17)$$

gdzie $\eta(t)$ to współczynnik uczenia sieci.

Przy tak zdefiniowanej funkcji błędu algorytm będzie wyznaczał taki wektor kierunkowy $\mathbf{p}(t)$, aby Q osiągnęło minimum, lub jej wartość spadła poniżej założonego progu. Oznacza to spełnienie nierówności $Q(\mathbf{w}(t+1)) < Q(\mathbf{w}(t))$.

W celu uproszczenia ogranicza się szereg Taylora (16) do rozwinięcia liniowego:

$$Q(\mathbf{w} + \mathbf{p}) = Q(\mathbf{w}) + [\mathbf{g}(\mathbf{w})]^T \mathbf{p} \quad (18)$$

Funkcja $Q(\mathbf{w})$ zależy od wag wyznaczonych w kroku bieżącym (t), natomiast $Q(\mathbf{w} + \mathbf{p})$ w kroku kolejnym ($t+1$), więc poszukiwany wektor $\mathbf{p}(t)$ musi spełniać nierówność $\mathbf{g}(\mathbf{w}(t))^T \mathbf{p}(t) < 0$. Jest to spełnione, gdy

$$\mathbf{p}(t) = -\mathbf{g}(\mathbf{w}(t)). \quad (19)$$

Po podstawieniu (19) do (17) uzyskuje się algorytm wyznaczający wagi wielowarstwowej sieci zgodny z regułą największego spadku [2]:

$$\mathbf{w}(t+1) = \mathbf{w}(t) - \eta \cdot \mathbf{g}(\mathbf{w}(t)) \quad (20)$$

Dla tak przyjętych reguł uczenia błąd na wyjściu sieci dla jednego wektora uczącego σ wynosi:

$$Q(t) = \sum_{i=1}^{N_L} (\varepsilon_i^L(t))^2 = \sum_{i=1}^{N_L} (d_i^L(t) - y_i^L(t))^2 \quad (21)$$

Z równania (20) dla dowolnej wagi w dowolnej warstwie otrzymuje się:

$$w_{i,j}^{(k)}(t+1) = w_{i,j}^{(k)}(t) - \eta \frac{\partial Q(t)}{\partial w_{i,j}^{(k)}(t)} = w_{i,j}^{(k)}(t) + 2\eta \delta_i^{(k)}(t) x_j^{(k)}(t) \quad (22)$$

Parametr $\delta_i^{(k)}(t)$ jest wyznaczany inaczej dla warstwy wyjściowej i ukrytej:

$$\delta_i^{(k)}(t) = \begin{cases} \varepsilon_i^{(L)}(t) \cdot f'(s_i^{(L)}(t)) & \text{for } k = L \\ f'(s_i^{(k)}(t)) \cdot \sum_{m=1}^{N_{k+1}} \delta_m^{(k+1)}(t) w_{i,m}^{(k+1)}(t) & \text{for } k \neq L \end{cases} \quad (23)$$

gdzie: $\varepsilon_i^{(L)}(t) = d_i^{(L)}(t) - y_i^{(L)}(t)$

Uczenie sieci przeprowadza się metodą przyrostowego uaktualniania wag, czyli każdorazowo po podaniu na wejście kolejnego wektora uczącego wyznacza się odpowiedzi sieci i modyfikuje wagi. Symulacja jest kontynuowana aż do chwili, gdy sumaryczny błąd wyjściowy dla całej epoki $Q^*(t)$ będzie mniejszy od przyjętego Q_{\min} .

$$Q^*(t) = \sum_{m=1}^M Q_m(t) \leq Q_{\min} \quad (24)$$

gdzie: M – liczba wektorów uczących w epoce

Przyjęto funkcję aktywacji neuronu jako ciągłą funkcję unipolarną typu sigmoidalnego:

$$f(s_i(t)) = \frac{1}{1 + e^{-\beta \cdot s_i(t)}} \quad (25)$$

gdzie: β - współczynnik stromości

Przy małej wartości współczynnika β funkcja ma kształt łagodny. Zwiększając β wykres staje się bardziej stromy aż do przebiegu progowego.

Przedstawiona funkcja definiuje wartość wyjściową neuronu zgodnie z zależnością (13). Pochodną funkcji aktywacji jest:

$$f'(s_i(t)) = \frac{\beta \cdot e^{-\beta \cdot s_i(t)}}{(1 + e^{-\beta \cdot s_i(t)})^2} = \beta \cdot f(s_i(t)) \cdot (1 - f(s_i(t))) \quad (26)$$

Funkcja (26) jest częścią algorytmu wyznaczającego korekty wag zgodnie ze wzorami (22) i (23).

PODSUMOWANIE

Przedstawiony algorytm sztucznej sieci neuronowej znalazł szerokie zastosowanie w różnych systemach eksperckich. Wykorzystanie go do wyznaczania parametrów liniowego modelu elektrycznego wydaje się być nadmiarowe. Jednakże w sytuacji, gdy nie istnieją rozwiązania układu równań opisującego ten model, uzasadnienie stosowania algorytmu jest uzasadnione. Dla współczynników równania obciążonych błędami pomiarowymi rozwiązaniem jest pewien obszar. Wybór optymalnego rozwiązania z tego obszaru wymaga zastosowania algorytmu komputerowego. Jedną z metod jest algorytm bisekcji wielowymiarowej zbudowany przez autora i przedstawiony w publikacji [3].

Adaptacja algorytmu sztucznej sieci neuronowej wymaga zastosowania specjalizowanego oprogramowania, lub budowy go w dowolnym języku programowania.

CONSTRUCTION OF THE ARTIFICIAL NEURAL NETWORK TO IDENTIFY THE MODEL PARAMETERS OF THE ELECTRIC RECEIVER

Abstract

The paper presents the method of determining the electrical parameters of the model that contains the serial receiver combination of elements R, L, E, based on artificial neural network. The purpose of this network is determine solutions N equations with N unknowns. This algorithm allows to determine a solution where A_i coefficients in these equations contain errors. For some values of A_i (where $i = 1, \dots, N$) there is no inverse function of input equations. In this case, it is impossible to determine a solution of the equations of classical methods.

BIBLIOGRAFIA

1. Gacek Z.: *Technika wysokich napięć. Izolacja wysokonapięciowa w elektroenergetyce. Przepięcia i ochrona przed przepięciami*. Skrypt uczelniany nr 2137 wyd.III. Wydawnictwo Politechniki Śląskiej, Gliwice 1999
2. Rutkowski L.: *Metody i techniki sztucznej inteligencji. Inteligencja obliczeniowa*. PWN, Warszawa 2005
3. Zajkowski K.: *Algorytm do rozwiązywania układów równań z danymi obciążonymi błędami pomiarowymi.*, Conf.Computer Applications In Electrical Engineering, p.499-502 T.II, IEP Politechniki Poznańskiej 2004
4. Zajkowski K.: *Analysis of overvoltages on inductive system with varistor and capacitor*. Poznań University of Technology, Academic Journals Issue 59 2009, Electrical Engineering, Publ.by Poznań University of Technology 2009, p.87-97
5. Zajkowski K.: *Analiza stanu nieustalonego w obwodach rezystancyjno-indukcyjnych w aspekcie minimalizacji przepieć komutacyjnych*. Wydawnictwo Uczelniane Politechniki Koszalińskiej, Koszalin 2011

Autorzy:

dr inż. Konrad ZAJKOWSKI– Politechnika Koszalińska

dr inż. Stanisław DUER– Politechnika Koszalińska