

POLITECHNIKA WROCŁAWSKA

Wydział Elektryczny

***ADAPTACYJNE SIECI NEURONOWE
W ZASTOSOWANIU DO IDENTYFIKACJI
STANU UKŁADÓW ELEKTRYCZNYCH***

Paweł Kostyła

Rozprawa doktorska

Promotor pracy:

Prof. zw. dr hab. inż. Tadeusz Łobos

WROCŁAW 1998

Dziękuję serdecznie

*Panu prof. dr. hab. inż. Tadeuszowi Łobosowi
za podjęcie się trudu kierowania moją pracą oraz
za wieloletnią opiekę i pomoc w przygotowaniu
rozprawy*

Autor

SPIS TREŚCI

	strona
1. Wprowadzenie	7
1.1. Cel i teza pracy	10
2. Sieci neuronowe	11
2.1. Rys historyczny	11
2.2. Podstawy biologiczne funkcjonowania sieci neuropodobnej	12
2.3. Kierunki badań i zastosowań sieci neuronowych	15
2.4. Model McCullocha-Pittsa	16
2.5. Podstawowe struktury sieci	16
2.6. Podsumowanie	19
3. Optymalizacyjne sieci neuronowe	20
3.1. Wprowadzenie	20
3.2. Sformułowanie i rozwiązanie problemu optymalizacyjnego	20
3.2.1. Podstawowe dyskretne algorytmy spadku gradientu	21
3.2.2. Podstawowy analogowy algorytm spadku gradientu	23
3.3. Określenie funkcji celu (energetycznej)	25
4. Adaptacyjne algorytmy uczenia oparte na kryteriach LS i TLS	27
4.1. Wprowadzenie	27
4.2. Sformułowanie zagadnienia	28
4.3. Kryterium TLS	31
4.4. Adaptacyjny algorytm uczenia dla kryterium LS	33
4.5. Adaptacyjny algorytm uczenia dla kryterium TLS	36
5. Adaptacyjne sieci neuronowe do wyznaczania parametrów zakłóconego sygnału sinusoidalnego	37
5.1. Wprowadzenie	37
5.2. Kryterium LS	38
5.3. Kryterium TLS	39
5.4. Odporne kryterium TLS (RTLS)	40
5.5. Implementacja algorytmów	41
5.6. Wyniki badań	44

6. Sieć neuronowa do detekcji stanu asynchronicznego maszyn synchronicznych	52
6.1. Wprowadzenie.....	52
6.2. Przebiegi prądu uzwojenia stojana.....	53
6.3. Algorytm działania sieci.....	54
6.4. Implementacja algorytmu.....	57
6.5. Wyniki badań.....	59
7. Wykrywanie zwarć łukowych w przesyłowych liniach elektro-energetycznych	75
7.1. Wprowadzenie.....	75
7.2. Prezentacja metody.....	76
7.3. Algorytm działania sieci.....	78
7.4. Implementacja algorytmu.....	79
7.5. Wyniki badań.....	80
Wnioski	89
Literatura	91

Wykaz ważniejszych oznaczeń:

- a_i, b_i - amplitudy składowych harmoniczných sygnału;
- e - błąd estymacji;
- $E(\mathbf{x})$ - funkcja celu (funkcja energetyczna);
- f_s - częstotliwość próbkowania;
- i_s - prąd stojana maszyny synchronicznej;
- I_f, I_b, I_r - amplitudy składowych prądu stojana;
- N - liczba dyskretnych próbek sygnału;
- NT - okno próbkowania;
- s - poślizg wirnika;
- t_c - czas konwergencji trajektorii wyznaczanych parametrów;
- T - okres próbkowania;
- $x(t)$ - model przebiegu sygnału mierzonego;
- $X_{a,...,f}$ - wyznaczone parametry składowych sygnału;
- $X_{1,...,5}$ - obliczane amplitudy składowych sygnału;
- y_m - dyskretna wartość (próbka) sygnału mierzonego;
-
- $\mathbf{b}$ - wektor obserwacji;
- $\mathbf{d}_k$ - dyskretna postać wektora kierunku optymalizacji;
- $\mathbf{r}$ - resztkowy wektor błędu;
- $\mathbf{x}$ - wektor zmiennych optymalizowanych;
- $\mathbf{y}$ - wektor sygnału mierzonego;
- $\|\mathbf{r}\|_p$ - norma L_p wektora $\mathbf{r}$;
- $\|\Delta\|_F$ - norma Frobeniusa wyrażenia Δ ;

- A - macierz współczynników rzeczywistych ze znanych elementów;
- A^T - transpozycja macierzy A ;
- D - macierz wartości sygnału modelowego;
- U - macierz lewych wektorów szczególnych macierzy A ;
- S - macierz wartości osobliwych;
- V - macierz prawych wektorów szczególnych macierzy A ;
- Λ - macierz wartości własnych macierzy A ;
- α - parametr regularyzacji;
- β_k - dyskretna postać współczynnika sprzężenia;
- γ - parametr określający nachylenie funkcji sigmoidalnej;
- $\gamma_f, \gamma_b, \gamma_r$ - fazy składowych prądu stojana;
- η - współczynnik minimalizacji kierunkowej;
- μ - współczynnik uczenia (krok uczenia);
- ρ - funkcja wagowa lub straty;
- $\sigma(e)$ - funkcja błędu;
- τ - stała czasowa integratorów;
- $\Psi(e)$ - nieliniowa funkcja aktywacji;
- ω - pulsacja podstawowej składowej sygnału mierzonego;
- ω_f - prędkość kątowna składowej pola w przód;
- ω_b - prędkość kątowna składowej pola w tył;
- ω_s - prędkość kątowna wirującego pola stojana;
- ω_r - prędkość kątowna wirnika.

1. Wprowadzenie

W wielu zagadnieniach telekomunikacji, elektroniki, automatyki, techniki zabezpieczeń energetycznych, istnieje konieczność identyfikacji stanu pracy, separacji i estymacji parametrów sygnałów elektrycznych, np.: wydzielanie z sygnału podstawowej harmonicznej, filtracja szumu, separacja wyższych harmonicznych (składowe ortogonalne), ocena zawartości wyższych harmonicznych czy estymacja częstotliwości, amplitudy i fazy. Często konieczne jest wyznaczenie tych parametrów i ich śledzenie w czasie rzeczywistym (tzn. bezpośrednio w czasie trwania pomiarów), przy zmieniających się w trakcie pomiarów szumach i zakłóceniach. Większość prezentowanych w literaturze i stosowanych metod estymacji i identyfikacji bazuje na wykonaniu serii pomiarów a następnie przetworzeniu danych pomiarowych przy zastosowaniu algorytmów numerycznych. Algorytmy przetwarzania cyfrowego, chociaż bardzo efektywne, nie są przeznaczone do przetwarzania równoległego i nie zawsze umożliwiają estymację i identyfikację w czasie rzeczywistym (tzw. on-line - tzn. podczas trwania pomiarów). W niniejszej pracy analizowane są rozwiązania tego typu problemów przy użyciu techniki sztucznych sieci neuronowych.

Osiągnięty w ostatnich latach postęp w rozwoju sztucznych sieci neuronowych oraz w zakresie ich technologicznych realizacji, stwarza nowe możliwości konstrukcji wyspecjalizowanych procesorów, umożliwiających estymację i identyfikację parametrów w czasie rzeczywistym.

Przedstawione w rozprawie układy - oparte na teorii optymalizacyjnych sieci neuronowych - zaprojektowano z myślą o ewentualnym zastosowaniu do estymacji parametrów sygnałów prądów i napięć w elektrotechnice oraz identyfikacji wybranych zakłóceń stanu pracy niektórych urządzeń. W układach automatyki i zabezpieczeń, celem rozpoznania stanu chronionego obiektu, konieczne jest określenie wartości niektórych wielkości elektrycznych. Należą do nich przede wszystkim parametry amplitudowe podstawowej składowej prądów i napięć, moc czynna i bierna, impedancja. Dla układów sterowania przekształtników energoelektronicznych zasadnicze znaczenie ma dokładna znajomość parametrów podstawowej składowej napięcia. Z drugiej strony, przekształtniki wywołują znaczne zniekształcenia

przebiegów prądów i napięć. Szereg innych układów automatyki wykorzystuje także parametry podstawowej harmonicznej napięć i prądów. Przedstawione w rozprawie sieci neuronowe dokonują, w czasie rzeczywistym, estymacji amplitudy składowych ortogonalnych sygnału prądu lub napięcia. Znając te wielkości można łatwo obliczyć amplitudę i fazę podstawowej składowej, moce oraz impedancję. Opisane układy mogą być wykorzystane do identyfikacji stanu niektórych urządzeń elektrycznych. Przykładowo przeanalizowano możliwości identyfikacji stanu pracy asynchronicznej maszyny synchronicznej oraz identyfikacji zwarcia łukowego w liniach średnich i wysokich napięć.

Przedstawione w pracy sieci neuronowe są układami analogowymi. Sygnały dostarczane do układu estymującego są próbkowane, podobnie jak w przypadku stosowanych obecnie powszechnie układów cyfrowych. Jednak wartości dyskretne sygnału nie są przekształcane do postaci cyfrowej, lecz bezpośrednio przetwarzane przez sieć neuronową. Sygnał wyjściowy odpowiadający estymowanym parametrom jest ciągły - może być jednak, w razie potrzeby, przetworzony na sygnał dyskretny lub cyfrowy. Należy tu wspomnieć, że istnieje możliwość realizacji opisywanych układów w wersji cyfrowej, po uprzedniej zamianie opisujących je równań różniczkowych na równania różnicowe.

Do chwili obecnej opracowano wiele algorytmów cyfrowych służących do estymacji w czasie rzeczywistym parametrów sygnałów. Bazują one na ogół na technice Fouriera lub teorii filtrów Kalmana. Algorytmy te są dokładne gdy znana jest częstotliwość mierzonego sygnału. Algorytmy Fouriera posiadają najlepsze właściwości filtrujące gdy okno próbkowania jest równe pełnemu okresowi lub całkowitej liczbie okresów. Rekursywne filtry Kalmana charakteryzują się natomiast dłuższymi czasami odpowiedzi. Najpowszechniej stosowane algorytmy cyfrowe opracowane są przy założeniu znanej i niezmiennej podczas pomiaru częstotliwości badanego sygnału. Sygnały występujące w elektrotechnice są niekiedy mocno zniekształcone. Zawierają wyższe harmoniczne i różnego rodzaju składowe przejściowe, np. o przebiegu aperiodycznym. Należy też wziąć pod uwagę błędy wnoszone przez przyrządy pomiarowe takie jak przekładniki pomiarowe, urządzenia próbkujące, przetworniki analogowo-cyfrowe. W tych warunkach dokładność i własności dynamiczne algorytmów cyfrowych ulegają znacznemu pogorszeniu. Dlatego też wydaje się, że w

niektórych przypadkach sieci neuronowe mogą wykazywać pewną przewagę nad algorytmami cyfrowymi, głównie ze względu na równoległy sposób przetwarzania.

W rozdziale 4 przedstawiono podstawy teoretyczne adaptacyjnych algorytmów uczenia dla optymalizacyjnych sieci neuronowych opartych na kryteriach LS (Least Squares) i TLS (Total Least Squares). W rozdziale 5 przedstawiono układy sieci, skonstruowane w oparciu o te algorytmy. Umożliwiają one wyznaczanie amplitud składowych ortogonalnych zakłóconego szumem sygnału sinusoidalnego. Opracowane sieci neuronowe zostały przebadane dla różnego rodzaju zakłóceń sygnału badanego. Opracowano i przebadano zmodyfikowany układ sieci neuronowej, wykorzystujący kryterium RTLS (Robust Total Least Squares). Jest on szczególnie przydatny w sytuacjach, w których w sygnale badanym pojawiają się zakłócenia impulsowe.

W rozdziale 6 zaproponowano układy sieci z przeznaczeniem do wykrywania stanu pracy asynchronicznej maszyny synchronicznej, których algorytm działania oparty jest na standardowym kryterium najmniejszego błędu kwadratowego i metodzie najmniejszego spadku. Identyfikacja stanu asynchronicznego polega na analizie przebiegu prądu stojana. Opracowane sieci wyznaczają amplitudy poszczególnych składowych prądu stojana, których wartość w wyniku utraty synchronizmu zależy od wartości poślizgu wirnika.

W rozdziale 7 zaprezentowano metodę umożliwiającą zapobieganie załączaniu linii elektroenergetycznej na zwarcie trwałe w ramach cyklu SPZ. Metoda opiera się na estymacji w czasie rzeczywistym parametrów sygnałów oraz na analizie błędów estymacji. W celu rozpoznania zwarcia łukowego proponuje się użyć stopień zniekształcenia krzywej napięcia na początku linii oraz rodzaj zmian tego zniekształcenia w miarę upływu czasu palenia się łuku. Zaproponowana rekurencyjna sieć neuronowa wyznacza parametry podstawowej harmonicznej napięcia oraz błędy estymacji tych parametrów. Na ich podstawie określa się następnie stopień zniekształcenia napięcia. Pod wpływem ruchów powietrza łuk ma tendencję do wydłużania się. Wydłużanie się łuku wpływa na wzrost wartości błędów estymacji. Wzrost ten proponuje się zastosować jako kryterium dla rozróżniania zwarć łukowych od zwarć trwałych. Zaproponowana sieć neuronowa umożliwia rozwiązywanie postawionego zagadnienia w czasie rzeczywistym.

1.1. Cel i teza pracy

Celem pracy jest zbadanie możliwości zastosowania sieci neuronowych - których adaptacyjny algorytm działania oparty jest o nowe kryteria TLS i RTLS - do estymacji parametrów sygnałów w czasie rzeczywistym, oraz zbadanie możliwości zastosowania wybranych układów sieci neuronowych do identyfikacji stanu pracy niektórych układów elektrycznych.

T e z a p r a c y :

Adaptacyjne sieci neuronowe umożliwiają wyznaczanie parametrów mocno zniekształconych sygnałów okresowych, także przy nieznacznych wahaniach częstotliwości, jak też dokładniejszą oraz szybszą aniżeli dotychczas stosowane metody identyfikację niektórych stanów pracy układów elektrycznych.

2. Sieci neuronowe

2.1. Rys historyczny

Pierwotnym źródłem badań nad sieciami neuronowymi były prace biologów i neurofizjologów śledzących tajniki naszego mózgu. Odkrycia, jakie poczynili, stanowią fundament dzisiejszego wykorzystania sieci neuronowych do przetwarzania informacji (patrz rozdział 2.3). Wiele z odkryć uhonorowano najwyższym wyrazem naukowego uznania - nagrodą Nobla. Wymienić tu można kilka ważniejszych:

Pawlow	- model odruchu warunkowego (nagroda Nobla w 1904);
Ramon y Cajal	- opis struktury sieci nerwowej (1906);
Sherrington	- model nerwowego sterowania mięśni (1932);
Bekesy	- model percepcji słuchowej (1961);
Hodgkin i Huxley	- model propagacji sygnału w aksonie (1963);
Eccles	- model synapsy (1963);

oraz wiele innych.

Praktycznie można przyjąć, że sama dziedzina sieci neuronowych zaczęła istnieć dopiero wraz z wydaniem historycznej pracy „*A logical calculus of the ideas immanent in nervous activity*” autorstwa W. S. McCullocha i W. Pittsa (1943) [41, 50]. Tam po raz pierwszy pokuszono się o matematyczny opis komórki nerwowej i powiązanie tego opisu z problemem przetwarzania danych (rozdział 2.4). Praca ta miała jednak wymiar wąskiego konkretnego, natomiast znacznie ogólniejszą podstawę teoretyczną prac nad sieciami neuronowymi stworzyli: z punktu widzenia techniki - J. von Neumann w książce „*The Computer and the Brain*” wydanej w 1958 roku, zaś z biologicznego punktu widzenia - W. K. Taylor w pracy „*Computers and the nervous system. Models and analogues in biology*” (1960).

Już w tym pionierskim okresie stwierdzono, iż najcenniejszą właściwością sieci neuronowych jest ich zdolność do **równoległego** przetwarzania informacji. Ta właściwość cechuje pracę sieci neuronowej i w sposób zasadniczy różni od szeregowej pracy tradycyjnego komputera. Bardzo szybko stwierdzono również, iż dodatkowym atutem sieci może być proces ich **uczenia**, zastępujący tradycyjne programowanie.

Pierwszym, szeroko znanym przykładem zbudowanej i działającej sieci neuropodobnej, był **Perceptron** (1957 - F. Rosenblatt i Ch. Wightman). Był to układ

częściowo elektromechaniczny (zmienne wagi synaptyczne realizowano za pomocą potencjometrów, obracanych przez odpowiednio sterowane silniki elektryczne), a częściowo elektroniczny (sumowanie pobudzeń). Przeznaczeniem perceptronu było rozpoznawanie znaków alfanumerycznych z procesem uczenia jako metodą programowania systemu. Oprócz wielu wad i niedociągnięć, niewątpliwą zaletą perceptronu była zdolność do zachowania poprawnego działania nawet po uszkodzeniu pewnej części jego elementów.

Ciekawą konstrukcją neurokomputera, była sieć uczących się elementów elektrochemicznych, zwanych *Adaline* (od **A**daptive **L**inear **E**lement), stworzona w 1960 roku przez B. Widrowa z Uniwersytetu Standforda [34, 50, 59]. Zadania tej sieci, zwanej Madaline (**M**any **A**dalines), związane są z adaptacyjnym przetwarzaniem sygnałów. Jej znaczenie, obok ciekawej zasady działania i efektywnego procesu uczenia się, polegało na tym, że stanowiła podstawę pierwszego neurokomputera oferowanego komercyjnie. Co więcej, układ ten jest wykorzystywany gdzieś do dziś w radarach, sonarach, modemach i liniach telefonicznych.

Trudno wymienić wszystkie implementacje sztucznych sieci neuronowych, jest ich bardzo wiele. Sieci neuronowe bywają także samodzielnym obiektem badań teoretycznych jako całość. Na podstawie ich budowy podejmuje się próby wyjaśnienia zasad funkcjonowania naturalnych fragmentów systemu nerwowego na płaszczyźnie neurofizjologii, neuroanatomii, a także psychologii. Pojawia się nawet nadzieja, że w oparciu o teorię sieci neuronowych uda się zbudować teorię mózgu, chociaż droga do tak zarysowanego celu jest, bez wątpienia, jeszcze bardzo daleka.

2.2. Podstawy biologiczne funkcjonowania sieci neuropodobnej

Pierwowzorem wszelkich sieci neuronowych jest mózg ludzki, którego objętość szacuje się na około 1400 cm^3 , zaś powierzchnię na 2000 cm^2 (należy nadmienić, że kula o tej samej objętości ma zaledwie 600 cm^2). Masa mózgu wynosi około 1.5 kg. Jest to wartość średnia, gdyż są w tym zakresie duże różnice pomiędzy poszczególnymi ludźmi. Nie ma to jednak istotnego znaczenia, gdyż nie istnieje żaden naukowo udowodniony związek pomiędzy masą mózgu a jego intelektualną sprawnością. Na wymienioną masę mózgu w większości składa się woda.

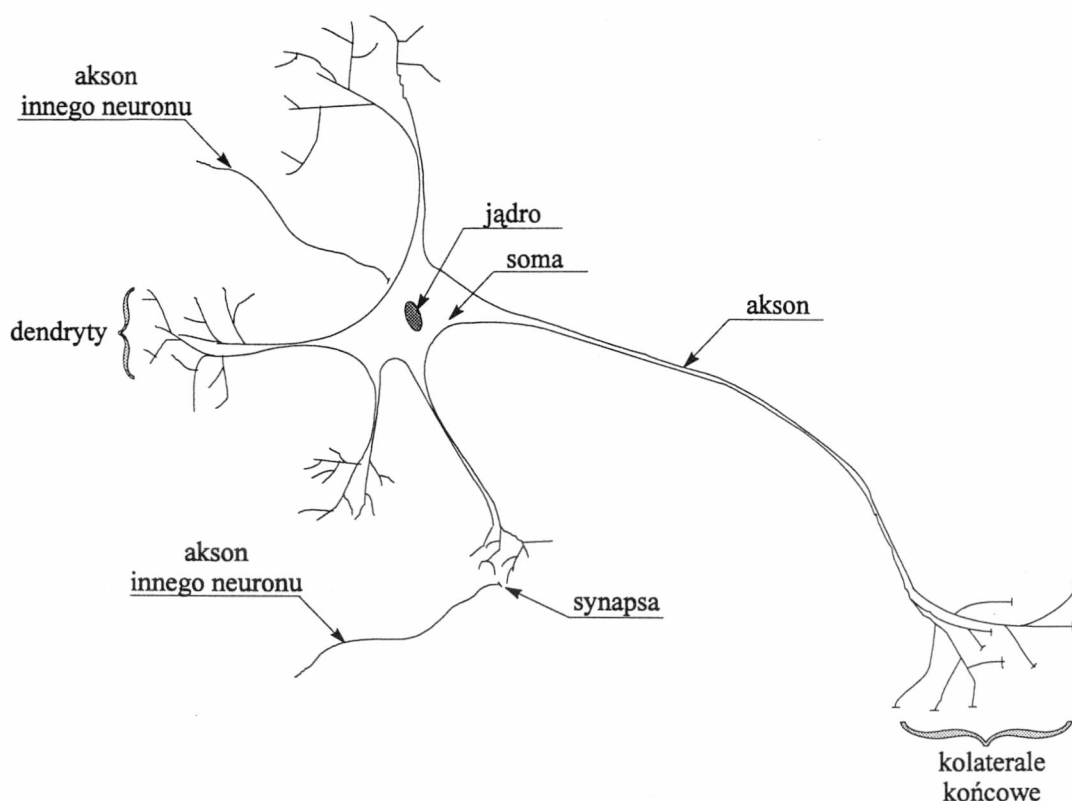
Zasadnicze znaczenie dla intelektualnych funkcji mózgu ma pokrywająca półkule kora o grubości średnio 3 mm, zawierająca około 10^{10} komórek nerwowych (neuronów) i 10^{12} komórek glejowych. Liczbę połączeń międzykomórkowych szacuje się na 10^{15} , przy przeciętnym dystansie od 10 mm do 1 m. Komórki nerwowe wysyłają i przyjmują impulsy o częstotliwości od 1 do 100 Hz, czasie trwania $1 \div 2$ ms, napięciu do ok. 100 mV i szybkości propagacji $1 \div 100$ m/s. Szybkość pracy mózgu można szacować na 10^{18} operacji/s (dla porównania - obecnie najszybsze komputery osiągają „zaledwie” szybkość 10^{10} operacji/s).

Komórka nerwowa, zwana w skrócie **neuronem**, stanowi podstawowy element systemu nerwowego i jak każda inna komórka posiada ciało z elementami wyposażenia cytologicznego, zwane *somą*, wewnątrz którego znajduje się jądro. Z somy neuronu wyrastają liczne wypustki pełniące istotną rolę w połączeniu z innymi komórkami. Zasadniczo można wyróżnić dwa rodzaje wypustek: liczne, cienkie i gęsto rozkrzewione *dendryty* oraz grubszy, rozwidlający się na końcu *akson* (rys. 2.1).

Sygnały wejściowe doprowadzane są do komórki za pośrednictwem *synaps*, zaś sygnał wyjściowy wyprowadzany jest przez akson i jego liczne odgałęzienia, zwane *kolateralami*. Kolaterale docierają do somy i dendrytów innych neuronów, tworząc kolejne synapsy. Dołączając wyjścia innych komórek nerwowych do danej komórki mogą znajdować się one zarówno na dendrytach, jak i bezpośrednio na ciele komórki.

Transmisja sygnałów wewnątrz systemu nerwowego jest skomplikowanym procesem chemiczno-elektrycznym. W dużym uproszczeniu można przyjąć, iż przekazywanie impulsu nerwowego od jednej komórki do drugiej opiera się na wydzielaniu, pod wpływem nadchodzących od synaps bodźców, specjalnych substancji chemicznych (tzw. *neuromediatorów*). Substancje te, oddziałując na błonę komórki, powodują zmianę jej potencjału elektrycznego, jednakże zmiana ta jest tym silniejsza, im więcej neuromediatora pojawi się na błonie. Miarą stopnia pobudzenia neuronu jest stopień polaryzacji błony komórkowej, zależny od sumarycznej ilości neuromediatora wydzielonego we wszystkich synapsach. Poszczególne synapsy różnią się natomiast wielkością oraz możliwością gromadzenia neuromediatora w pobliżu błony synaptycznej. Z tego właśnie powodu taki sam impuls, docierający do neuronu za pośrednictwem określonej synapsy, może powodować silniejsze bądź słabsze jego pobudzenie niż gdy przekaże go inna synapsa.





Rys. 2.1. Komórka neuronowa

Wnioskując, wejściom komórki przypisać można współczynniki liczbowe (wagi), odpowiadające ilości neuromediatora wydzielonego jednorazowo na poszczególnych synapsach. Wagi te są liczbami rzeczywistymi i mogą przyjmować zarówno wartości dodatnie, jak i ujemne, oddziałując pobudzająco bądź hamująco na neuron. Jeśli bilans pobudzeń i hamowań jest ujemny lub jeśli zaburzenie równowagi elektrycznej jest niewielkie, na wyjściu komórki nie da się dostrzec żadnej zmiany. Należy wówczas stwierdzić, iż pobudzenie komórki było mniejsze od progu jej zadziałania. Jeśli natomiast suma pobudzeń i hamowań przekroczyła próg uaktywnienia komórki, wówczas generowany jest sygnał wyjściowy (tzw. impuls nerwowy). Trzeba zauważyć, że sygnał ten jest niezależny od stopnia pobudzenia neuronu, czyli komórka działa na zasadzie wszystko albo nic. Po spełnieniu swej roli neuromediator jest usuwany (tzn. wchłaniany przez komórkę bądź rozkładany).

2.3. Kierunki badań i zastosowań sieci neuronowych

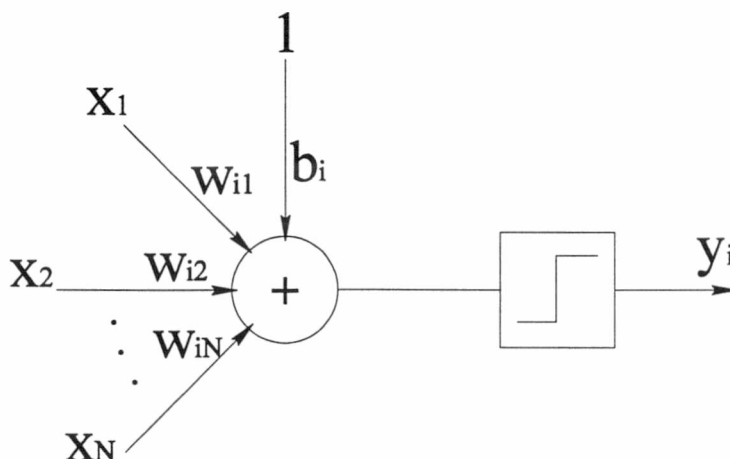
Jak już wcześniej wspomniano, bardzo trudno jest wymienić wszystkie aktualnie spotykane zastosowania sztucznych sieci neuronowych. Idąc za [50] można wskazać między innymi następujące:

- diagnostyka układów elektronicznych;
- badania psychiatryczne;
- prognozy giełdowe;
- prognozowanie sprzedaży;
- poszukiwania ropy naftowej;
- interpretacja badań biologicznych;
- prognozy cen;
- analiza badań medycznych;
- planowanie remontów maszyn;
- prognozowanie postępów w nauce;
- typowania na wyścigach konnych;
- analiza problemów produkcyjnych;
- optymalizacja działalności handlowej;
- analiza spektralna;
- optymalizacja utylizacji odpadów;
- dobór surowców;
- selekcja celów śledztwa w kryminalistyce;
- sterowanie procesów przemysłowych.

Najczęstszym obszarem ogólniejszych badań, angażujących sieci neuronowe, są prace nad algorytmami automatycznego uczenia się maszyn i ich zdolności do uogólniania nabywanych doświadczeń. Sieci neuronowe, będąc modelami naturalnych fragmentów systemu nerwowego człowieka, używane są często jako narzędzie do modelowania innych systemów i zjawisk. Rozważane są również systemy podejmowania decyzji i wspomagające wnioskowanie na przykład w sieciach energetycznych.

2.4. Model McCullocha - Pittsa

Jak już wspomniano, opierając się na zasadzie funkcjonowania rzeczywistego neuronu, stworzono szereg modeli matematycznych uwzględniających w stopniu mniejszym lub większym własności komórki nerwowej. Jeden z pierwszych modeli został sformułowany przez McCullocha i Pittsa w 1943 r. Model ten zakłada, iż neuron jest jednostką binarną. Schemat modelu przedstawia rys. 2.2 [41]:



Rys. 2.2. Model komórki nerwowej według McCullocha i Pittsa

Sygnały wejściowe x_j sumowane są z odpowiednimi wagami w_{ij} w sumatorze i porównywane z progiem b_i . Sygnał wyjściowy y_i wyraża się wzorem:

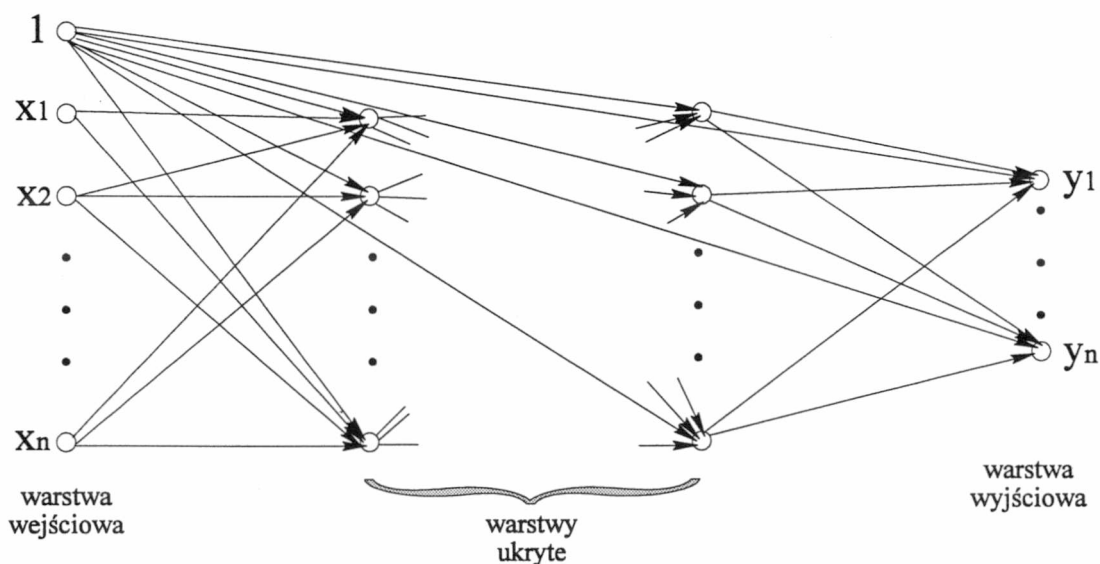
$$y_i = f\left(\sum_{j=1}^N w_{ij}x_j + b_i\right). \quad (2.1)$$

Funkcja f jest tak zwaną *funkcją aktywacji* - w rozpatrywanym modelu jest to funkcja skokowa. Należy tu nadmienić, iż istnieje szereg innych funkcji aktywacji, np. funkcja liniowa (sieć Madaline), czy funkcja sigmoidalna (np. $f(a) = \tanh(a)$) w dokładniejszy sposób odwzorowująca charakterystyki przejścia rzeczywistego - biologicznego neuronu.

2.5. Podstawowe struktury sieci

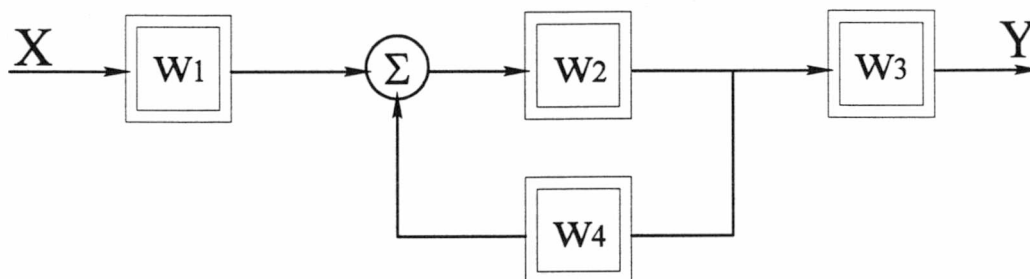
Pojedyncze neurony powiązane wzajemnie ze sobą tworzą sieć. W zależności od sposobu połączeń neuronów między sobą można wyróżnić trzy podstawowe rodzaje sieci [41, 50].

Najbardziej popularnym i jednocześnie najwcześniejszym typem sieci jest *sieć jednokierunkowa* (feedforward), zwana też perceptronem wielowarstwowym. Tworzą ją neurony, ułożone w warstwach o jednokierunkowym przepływie informacji (od wejścia do wyjścia). Z reguły sieć taka posiada warstwę wejściową, kilka (czasami tylko jedną) warstw ukrytych i warstwę wyjściową (rys. 2.3).



Rys. 2.3. Sieć neuronowa jednokierunkowa

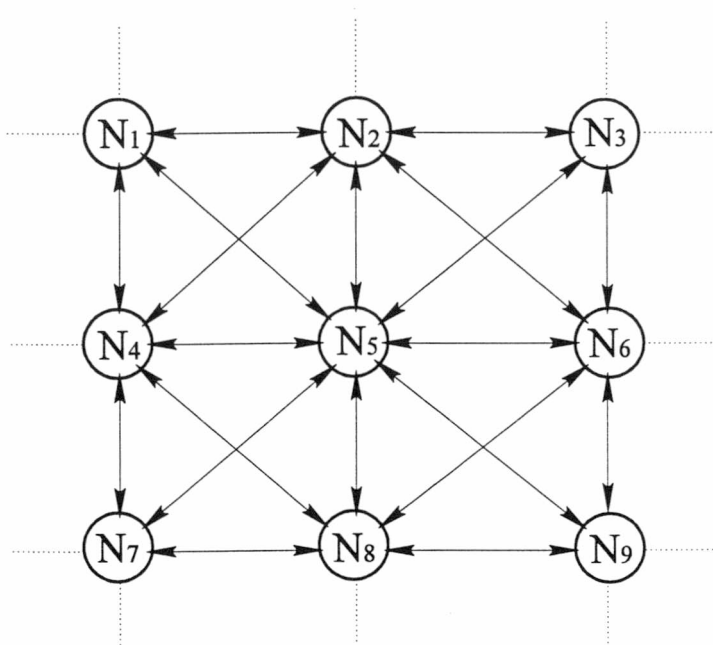
Drugą grupę stanowią *sieci rekurencyjne* (rys. 2.4), charakteryzujące się głównie tym, iż ustalanie wartości sygnałów wyjściowych przebiega nie w sposób bezpośredni, jak to miało miejsce w sieci jednokierunkowej, lecz rekurencyjnie. W sieci takiej przepływ sygnałów jest dwukierunkowy, zaś powiązanie pomiędzy neuronami w zasadzie dowolne.



Rys. 2.4. Struktura ogólna sieci rekurencyjnej

Bloki w_1 i w_3 oznaczają odpowiednio warstwę wejściową i wyjściową neuronów, natomiast w_2 i w_4 stanowią warstwy ukryte sieci realizujące sprzężenie zwrotne. Najbardziej znanym przedstawicielem tej kategorii sieci jest *sieć Hopfielda*, zdefiniowana w 1982 roku [41]. Typowym przedstawicielem sieci rekurencyjnych jest *sieć neuronowa optymalizująca*, w której sprzężenie dozwolone jest od dowolnego punktu sieci i nie musi posiadać struktury uporządkowanej jak na rys. 2.4. Sieci neuronowe optymalizujące, będące przedmiotem niniejszej pracy, zostaną opisane w jej dalszej części.

Wreszcie trzecią grupę stanowią *sieci komórkowe* (rys. 2.5) [27], w których poszczególne neurony powiązane są dwukierunkowo jedynie ze swymi sąsiadami. Sieci takie mają regularną budowę geometryczną, co ułatwia ich implementację. Najczęściej, każdy neuron wewnętrzny sieci powiązany jest jedynie z ośmioma swoimi najbliższymi sąsiadami. Podstawową dziedziną, w której tego typu sieci znalazły zastosowanie, jest przetwarzanie obrazów. Dotychczasowe badania wskazują na analogię między zasadą działania tej sieci a mechanizmem postrzegania zachodzącym w mózgu.



Rys. 2.5. Sieć neuronowa komórkowa

2.6. Podsumowanie

Podsumowując, *sieć neuropodobna* jest bardzo uproszczonym *modelem mózgu*. Składa się z dużej liczby (od kilkuset do kilkudziesięciu tysięcy) elementów przetwarzających informację - neuronów. Jednakże funkcje takiego neuronu w stosunku do rzeczywistych komórek nerwowych są bardzo uproszczone. Neurony są powiązane w sieć za pomocą połączeń o pewnych parametrach (tzw. wagach), modyfikowanych w trakcie procesu uczenia. Topologia połączeń oraz ich parametry stanowią program działania takiej sieci, zaś sygnały pojawiające się na jej wyjściach, w odpowiedzi na określone pobudzenia, są rozwiązaniami stawianych jej zadań.

Jedną z ważniejszych cech sieci jest ich zdolność do *adaptacji* i *samoorganizacji*. Jest to atut wykorzystywany niemal w większości z zastosowań. Sieci neuronowe cechuje także zmniejszona wrażliwość na uszkodzenia elementów. Systemami technicznymi, z którymi najczęściej się sieci neuronowe porównuje, są systemy *współbieżne*. Jednak stopień współbieżności obliczeń jest w sieciach neuronowych setki razy większy, niż w najnowocześniejszych systemach wieloprocesorowych. W pełni uzasadnionym wydaje się więc być fakt dużego zainteresowania, jakie budzi zagadnienie sprawności przetwarzania informacji w sieciach neuronowych. Dodatkowym atutem sieci jest wygoda ich programowania poprzez *uczenie*. Miał projektować algorytm wymaganego przetwarzania informacji, następnie dzielić go na moduły nadające się do współbieżnego wykonania, stawia się sieci przykładowe zadania, a sieć automatycznie, zgodnie z założoną strategią uczenia modyfikuje połączenia poszczególnych elementów i ich współczynniki wagowe.

3. Optymalizacyjne sieci neuronowe

3.1. Wprowadzenie

Sieci neuronowe, zarówno jednokierunkowe jak i rekurencyjne, w swoim działaniu dokonują minimalizacji pewnej *funkcji błędu* zwanej też *funkcją celu* lub *energetyczną*. Na ogół funkcja ta jest definiowana na etapie projektowania sieci i pełni w niej rolę pomocniczą. Istnieją wszakże sieci, dla których minimalizacja pewnej funkcji energetycznej jest głównym zadaniem [41, 46]. Sieci tego typu noszą nazwę *optymalizacyjnych sieci neuronowych*. Sieci zastosowane do badań będących tematem tej rozprawy należą do tej kategorii sztucznych sieci neuronowych.

Struktura tego typu sieci jest projektowana na podstawie układu równań różniczkowych (analogowa implementacja), bądź różnicowych (dyskretna implementacja), do których sprowadza się problem minimalizacji określonej funkcji energetycznej.

Opisane dalej układy będą stanowić implementację analogową algorytmu minimalizacyjnego, projektowanego na podstawie układu równań różniczkowych. Użycie równań różniczkowych zamiast różnicowych przynosi pewne korzyści, szczególnie zaś umożliwia rozwiązywanie zagadnień optymalizacyjnych w czasie rzeczywistym (przetwarzanie równoległe). Dużą szybkość działania układów analogowych uzyskuje się dzięki zastosowaniu elementów charakteryzujących się niewielkimi stałymi czasowymi, bez groźby utraty stabilności układu, jak to ma miejsce w przypadku układów dyskretnych.

3.2. Sformułowanie i rozwiązanie problemu optymalizacyjnego

Głównym zadaniem realizowanym przez optymalizacyjne sieci neuronowe jest znalezienie wektora $\mathbf{x}$, dla którego funkcja energetyczna $E(\mathbf{x})$ osiąga minimum. Jest to problem optymalizacji bez ograniczeń, który można przedstawić w sposób następujący:

$$\min E(\mathbf{x}), \quad (3.1)$$

gdzie:

$\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ jest wektorem zmiennych optymalizowanych.

Bardzo szeroką i ważną klasą metod w zagadnieniach optymalizacji są procedury, które bazują na algorytmach spadku gradientowego (gradient descent methods) często zwanych metodami gradientowymi [11]. Wszystkie one wywodzą się z technik standardowych znanych w literaturze jako: metody stromego spadku, podstawowe dynamiczne systemy gradientowe lub metody Newtona. W rozprawie będą omówione sieci optymalizacyjne, których algorytm minimalizacyjny oparty jest na tych metodach.

Ze względu na sposób realizacji powyższego zagadnienia, można wyróżnić następujące algorytmy:

- iteracyjne (dyskretne);
- ciągłe (analogowe).

3.2.1. Podstawowe dyskretne algorytmy spadku gradientu [11]

Podstawowym zadaniem metod iteracyjnych jest minimalizacja funkcji $E(\mathbf{x})$, gdzie $\mathbf{x} \in \mathbf{R}^n$. Ponadto zakłada się, że istnieje gradient i hesjan tej funkcji. Większość metod iteracyjnych rozwiązuje to zadanie generując sekwencje badanych (poszukiwanych) punktów $\mathbf{x}^{(k)}$ według poniższej procedury:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \eta^{(k)} \mathbf{d}_k, \quad \mathbf{x}(0) = \mathbf{x}^{(0)} \quad (k = 0, 1, 2, \dots), \quad (3.2)$$

gdzie $\mathbf{d}_k$ jest kierunkiem (wektorem) optymalizacji obliczonym dla $\mathbf{x} = \mathbf{x}^{(k)}$, zaś $\eta^{(k)}$ jest współczynnikiem minimalizacji kierunkowej, zwanej też krokiem uczenia (może to być wielkość stała lub dobierana adaptacyjnie podczas obliczeń).

W zależności od sposobu obliczania (definiowania) wektora optymalizacji $\mathbf{d}_k$ można wyróżnić następujące podstawowe metody:

1. Metoda największego spadku

Stosuje się tutaj liniowe przybliżenie funkcji $E(\mathbf{x})$, zaś kierunek optymalizacji określa wzór:

$$\mathbf{d}_k = -\nabla E(\mathbf{x}^{(k)}). \quad (3.3)$$

2. Metoda Newtona

Korzysta się w niej z aproksymacji kwadratowej funkcji $E(\mathbf{x})$. Wektor $\mathbf{d}_k$ wyznaczany jest z zależności:

$$\mathbf{d}_k = -[\nabla^2 E(\mathbf{x}^{(k)})]^{-1} \nabla E(\mathbf{x}^{(k)}). \quad (3.4)$$

3. Metoda gradientów sprzężonych

Oblicza ona aktualny poszukiwany kierunek optymalizacji jako liniową kombinację bieżącego wektora gradientu oraz poprzedniego kierunku optymalizacyjnego, zgodnie ze wzorem:

$$\mathbf{d}_k = -\nabla E(\mathbf{x}^{(k)}) + \beta_k \mathbf{d}_{k-1}, \quad (3.5)$$

gdzie:

$$\mathbf{d}_0 = -\nabla E(\mathbf{x}^{(0)}),$$

β_k jest współczynnikiem sprzężenia, którego wartość może być obliczana na podstawie różnych strategii.

Zaletą metody największego spadku jest przede wszystkim jej prostota i łatwość implementacji. Do jej wad należy zaliczyć stosunkowo wolną zbieżność, szczególnie tam, gdzie gradient przyjmuje małe wartości - czyli w okolicy poszukiwanego minimum.

Metoda Newtona, szybciej zbieżna od poprzedniej, jest wrażliwa na wybór punktu startowego, co czasem powoduje jej niezbieżność. Ponadto jest ona dość złożona i trudna w implementacji, głównie ze względu na konieczność obliczania pochodnych drugiego rzędu.

Metoda gradientów sprzężonych, podobnie jak metoda Newtona, jest znacznie szybsza w porównaniu z algorytmem największego spadku, jest jednak znacznie bardziej skomplikowana, wymaga obliczania współczynnika minimalizacji kierunkowej η oraz współczynnika sprzężenia β_k w każdym kroku iteracji.

W rozprawie opisane są układy stanowiące analogową implementację algorytmów optymalizacyjnych. Układy te są oparte na algorytmie największego spadku. Łatwość implementacji tego algorytmu pozwala zakładać, że możliwa

i opłacalna byłaby realizacja opisywanych sieci neuronowych z zastosowaniem dostępnych układów VLSI.

3.2.2. Podstawowy analogowy algorytm spadku gradientu

W przypadku implementacji analogowej, zadaniem algorytmu spadku gradientu jest przekształcenie problemu minimalizacyjnego w układ równań różniczkowych:

$$\frac{dx_j}{dt} = -\sum_{i=1}^n \mu_{ji} \frac{\partial E}{\partial x_i}, \quad (3.6)$$

z warunkiem początkowym $x_j(0) = x_j^{(0)}$,

które mogą być zapisane w postaci macierzowej:

$$\frac{d\mathbf{x}}{dt} = -\mu(\mathbf{x}, t) \nabla_{\mathbf{x}} E(\mathbf{x}), \quad (3.7)$$

gdzie:

$$\frac{d\mathbf{x}}{dt} = \left[\frac{dx_1}{dt}, \dots, \frac{dx_n}{dt} \right]^T,$$

$$\mathbf{x} = [x_1, x_2, \dots, x_n]^T$$

oraz $\mu(\mathbf{x}, t)$ jest symetryczną, dodatnio określoną macierzą uczącą z wejściami zależnymi od czasu t i wektora $\mathbf{x}(t)$.

W celu znalezienia pożądanego wektora $\mathbf{x}^*$, który minimalizuje funkcję $E(\mathbf{x})$, należy rozwiązać układ równań różniczkowych z warunkami początkowymi. Oznacza to, że minima funkcji energetycznej są zdeterminowane rozwiązaniem równania $\mathbf{x}^* = \lim_{t \rightarrow \infty} \mathbf{x}(t)$. Należy zwrócić uwagę na to, że głównym zadaniem jest znalezienie granicy, nie zaś określenie szczegółowego przebiegu i kształtu trajektorii $\mathbf{x}(t)$.

W celu pokazania, że powyższy układ równań różniczkowych jest stabilny, należy określić pochodną funkcji energetycznej (Lyapunova) po czasie [11]:

$$\frac{dE}{dt} = \sum_{j=1}^n \frac{\partial E}{\partial x_i} \frac{dx_j}{dt} = -[\nabla_{\mathbf{x}} E(\mathbf{x})]^T \mu(\mathbf{x}, t) \nabla_{\mathbf{x}} E(\mathbf{x}) \leq 0, \quad (3.8)$$

pod warunkiem, że macierz μ jest symetryczna i dodatnio określona. To powyższe określenie gwarantuje, że funkcja energetyczna $E(\mathbf{x})$ maleje w czasie i ustala się w stabilnym minimum lokalnym przy $t \rightarrow \infty$. Szybkość ustalenia się minimum jest zależna od wyboru wejść macierzy μ . Różne sposoby wyboru tej macierzy doprowadziły do powstania różnych metod gradientowych znanych w teorii optymalizacji. I tak, w najpopularniejszej metodzie największego spadku (steepest-descent method), macierz $\mu(\mathbf{x}, t)$ zredukowana jest do macierzy jednostkowej pomnożonej przez stały dodatni parametr. Wówczas, równanie opisujące algorytm największego spadku przyjmuje postać:

$$\frac{d\mathbf{x}}{dt} = -\mu(t)\nabla E(\mathbf{x}), \quad (3.9)$$

gdzie:

$\mu(t) > 0$ - współczynnik uczenia.

O ile w przypadku implementacji dyskretniej, algorytm opisany równaniem (3.2) generuje ciąg punktów $\{\mathbf{x}^k\}$, to w przypadku implementacji analogowej otrzymuje się ciągłą trajektorię $\mathbf{x}(t)$, dążącą od punktu początkowego $\mathbf{x}(0)$ do punktu stacjonarnego $\mathbf{x}^*$. W punkcie $\mathbf{x} = \mathbf{x}^*$, w którym $\nabla E(\mathbf{x}) = 0$ pochodna $d\mathbf{x}/dt$ zeruje się. Punkt $\mathbf{x}^*$ może być poszukiwanym minimum lub punktem siodłowym. W układach rzeczywistych, ze względu na wpływ nieuniknionych zakłóceń, pozostawanie układu w punkcie siodłowym jest praktycznie mało prawdopodobne. Można więc napisać:

$$\lim_{t \rightarrow \infty} \mathbf{x} = \mathbf{x}^*, \quad (3.10)$$

gdzie $\mathbf{x}^*$ jest poszukiwanym rozwiązaniem problemu optymalizacyjnego.

Współczynnik uczenia $\mu(t)$ ma istotny wpływ na przebieg trajektorii $\mathbf{x}(t)$. Przede wszystkim, jak wykazały badania, od niego zależy szybkość zbieżności implementowanego algorytmu. Dokładne określenie kształtu trajektorii nie ma raczej znaczenia. Optymalny dobór $\mu(t)$ jest zagadnieniem dość złożonym. W opisywanych układach (rozdział 6 i 7) przyjęto $\mu(t) = \mu = 1/\tau = \text{const}$ (gdzie τ - stała czasowa integratorów), co znacznie upraszcza projektowany układ. W zależności od zastosowanej technologii realizacji układu (np. CMOS) stała czasowa może

przyjmować wartości rzędu mikrosekund; oznacza to dużą wartość współczynnika μ i dużą szybkość zbieżności układu.

3.3. Określenie funkcji celu (energetycznej)

Odpowiedni wybór funkcji energetycznej (funkcji celu) $E(\mathbf{x})$ ma decydujący wpływ na strukturę i właściwości projektowanej sieci neuronowej. Omawiane układy są przeznaczone do estymowania parametrów sygnałów na podstawie dostarczonych do układu próbek ich wartości chwilowych. Z każdą próbką sygnału y_m związany jest błąd e_m , będący różnicą między otrzymaną z pomiaru wartością próbki y_m a wartością estymowaną przez układ x_m :

$$e_m = y_m - x_m. \quad (3.11)$$

Wartości funkcji celu $E(\mathbf{x})$ są wyznaczane na podstawie wartości błędów e_m . Opisuje to ogólne równanie:

$$E(\mathbf{x}) = \sum_{m=1}^N \sigma_m[e_m(\mathbf{x})], \quad (3.12)$$

gdzie:

$\mathbf{x}$ - jest wektorem zmiennych optymalizowanych, którego elementami są estymowane parametry sygnału,

N - jest liczbą próbek sygnału,

$\sigma_m[e_m(\mathbf{x})]$ - jest odpowiednio dobraną funkcją błędu.

Postać funkcji $\sigma_m[e_m(\mathbf{x})]$ zależy od wybranego kryterium estymacji.

Podstawowe znaczenie ma kilka następujących przypadków:

1. Przyjmując $\sigma_m[e_m] = e_m^2 / 2$ otrzymuje się *kryterium najmniejszego błędu kwadratowego* (kryterium minimum normy L_2).

Funkcja energetyczna ma postać:

$$E_2(\mathbf{x}) = \frac{1}{2} \|\mathbf{e}\|_2^2 = \frac{1}{2} \mathbf{e}^T \mathbf{e} = \frac{1}{2} \sum_{m=1}^N e_m^2(\mathbf{x}); \quad (3.13)$$

2. Przyjmując $\sigma_m[e_m] = \frac{1}{\gamma} \ln\{\cosh(\gamma e_m)\}$ otrzymuje się *odporne kryterium najmniejszego błędu kwadratowego*.

Funkcja energetyczna ma postać:

$$E_{12}(\mathbf{x}) = \frac{1}{\gamma} \sum_{m=1}^N \ln\{\cosh[\gamma e_m(\mathbf{x})]\}, \quad (3.14)$$

gdzie $\gamma > 0$;

3. Przyjmując $\sigma_m[e_m] = |e_m|$ otrzymuje się *kryterium najmniejszego błędu bezwzględnego* (kryterium minimum normy L_1).

Funkcja energetyczna ma postać:

$$E_1(\mathbf{x}) = \sum_{m=1}^N |e_m(\mathbf{x})|; \quad (3.15)$$

4. Przyjmując $\sigma_m[e_m] = \begin{cases} |e_m| & \text{gdy } |e_m| = \max_{1 \leq i \leq N} \{|e_i|\} \\ 0 & \text{w pozostałych przypadkach} \end{cases}$ otrzymuje się

kryterium minimax (kryterium Czebyszewa, kryterium minimum normy L_∞).

Funkcja energetyczna ma postać:

$$E_\infty(\mathbf{x}) = \max_{1 \leq m \leq N} \{|e_m(\mathbf{x})|\}. \quad (3.16)$$

Wybór kryterium estymacji jest pierwszym etapem projektowania sieci neuronowej do wyznaczania parametrów sygnału. Zależy ono przede wszystkim od przewidywanego rodzaju zakłóceń w sygnale. Kryterium najmniejszego błędu kwadratowego jest optymalne, jeżeli sygnał zakłócony jest szumem o rozkładzie normalnym (Gaussowskim). Kryterium to jest powszechnie stosowane w zagadnieniach optymalizacji. Jeżeli jednak rozkład zakłóceń jest inny, lepsze wyniki daje zastosowanie innych kryteriów, chociaż są one zwykle bardziej złożone i trudniejsze w implementacji [56].

W rozprawie będą badane sieci, których funkcja energetyczna określona jest na podstawie kryterium najmniejszego błędu kwadratowego.

4. Adaptacyjne algorytmy uczenia oparte na kryteriach LS i TLS

4.1. Wprowadzenie

Wiele problemów w nauce i technice sprowadza się do rozwiązywania poniższego układu równań liniowych [14]:

$$\mathbf{Ax} \equiv \mathbf{b}, \quad (4.1a)$$

gdzie: $\mathbf{A}$ - jest to $m \times n$ wymiarowa macierz danych,
 $\mathbf{b}$ - jest to m wymiarowy wektor obserwacji.

Zazwyczaj $\mathbf{A}$ i $\mathbf{b}$ są zniekształconymi lub zakłóconymi wersjami dokładnych, lecz nie obserwowalnych danych $\mathbf{A}_0$ i $\mathbf{b}_0$, zaś ścisła relacja pomiędzy nimi opisana jest przez równania macierzowe:

$$\mathbf{A}_0 \mathbf{x} = \mathbf{b}_0. \quad (4.1b)$$

Taki opis, można znaleźć w wielu dziedzinach naukowych i problematyce inżynierskiej. Macierz $\mathbf{A}$ występuje pod różną nazwą, w zależności od dziedziny nauki, w której jest rozpatrywana. Każde zagadnienie estymacji liniowej parametrów, występujące w: sterowaniu i kontroli automatycznej, systemach identyfikacji i przetwarzania sygnałów, fizyce, medycynie, biologii, prowadzi do nadokreślonych układów równań liniowych (4.1a) [4, 11, 12, 13, 22, 23, 54, 55]. W pewnych przypadkach ich ilość może być ograniczana z powodu braku pełnej informacji. Generalnie, można powiedzieć, że w zastosowaniach przetwarzania sygnałów (signal processing applications) przypadki nadokreślone ($m > n$) opisują np.: filtrację, estymację parametrów, zagadnienia identyfikacji, podczas gdy przypadki podokreślone ($m < n$) - opisują zagadnienia odwrotności oraz ekstrapolacji [4, 11].

W wielu zastosowaniach (np. w robotyce, przetwarzaniu sygnałów) pożądane jest rozwiązanie zagadnień w czasie rzeczywistym. Dopuszczalne czasy przetwarzania bywają niekiedy bardzo krótkie (rzędu setek nanosekund), tak że zastosowanie komputerów może okazać się niewystarczające lub bardzo kosztowne. Jedną z możliwości rozwiązania takich kłopotów jest zastosowanie analogowych sieci neuronowych. Jednak znane metody oraz architektury sieci żądają użycia co najmniej n (n jest liczbą nieznanymi zmiennymi) sztucznych neuronów (części procesorowych)

[11, 12, 13, 51]. W wielu zastosowaniach inżynierskich (np. rekonstrukcji obrazu) wymagane jest rozwiązanie w bardzo krótkim czasie dużego układu algebraicznych równań liniowych. Dla takich zaś zagadnień, znane struktury sieci neuronowych wymagają z kolei ekstremalnie dużej liczby neuronów, co w praktycznej implementacji hardwarowej sieci neuronowej może być trudne, kosztowne, a nawet niemożliwe.

W tym rozdziale, przedstawione będą adaptacyjne algorytmy uczenia, opracowane z zastosowaniem kryterium *Least Squares* (LS) oraz *Total Least Squares* (TLS). W rozdziale następnym (rozdział 5) będą omówione, skonstruowane na ich podstawie, adaptacyjne sztuczne sieci neuronowe do wyznaczania parametrów podstawowej składowej sygnału, zawierające tylko jeden uproszczony neuron (single adaptive processing unit) z implementowanym adaptacyjnym algorytmem uczenia.

Wielu badaczy z różnych dyscyplin naukowych wynalazło klasę skutecznych algorytmów adaptacyjnych do rozwiązania układów liniowych równań algebraicznych [4, 11, 12, 13, 22, 23, 54, 55, 58]. Najwcześniej, bo w 1937 roku, Kaczmarz [28] wynalazł bardzo prosty i skuteczny algorytm o nazwie RAP (**R**ow-**A**ction **P**rojection). Ponieważ układy równań liniowych są stosowane w różnych dyscyplinach i zagadnieniach naukowych, algorytm RAP i jego modyfikacje szybko się rozpowszechniły. Obecnie występują pod różnymi nazwami w różnych zastosowaniach, np.: w medycynie (tomografia komputerowa) algorytm ten można spotkać pod nazwą ART (an Algebraic **R**econstructed **T**echnique), zaś w literaturze filtrów adaptacyjnych - jako algorytm α - LMS lub reguła delty Widrowa-Hoffa.

4.2. Sformułowanie zagadnienia

Zakładamy, że chcemy rozwiązać układ równań liniowych w postaci skalarnej [11, 14]:

$$\sum_{j=1}^n a_{ij}x_j = b_i \quad (i=1, 2, \dots, m), \quad (4.2a)$$

lub w postaci macierzowej:

$$\mathbf{Ax} = \mathbf{b}, \quad (4.2b)$$

gdzie: $\mathbf{x}$ - jest to n wymiarowy nieznaną wektor,

$\mathbf{b}$ - jest to m wymiarowy wektor źródłowy (obserwacji lub pomiarów),

$\mathbf{A} = [a_{ij}]$ - jest to $m \times n$ wymiarowa macierz współczynników rzeczywistych ze znanych elementów.

Liczebność równań nie musi być dokładnie równa n , może być mniejsza, równa lub większa od ilości zmiennych, tj. składników wektora $\mathbf{x}$.

Jeśli $m < n$ wówczas układ równań jest zwany jako podokreślony, jeśli zaś $m > n$ - nadokreślony. Układ taki może mieć jedno rozwiązanie $\mathbf{x}^*$, kilka rozwiązań lub może ich nie mieć. W problematyce estymacji liniowej powyższy układ równań nadokreślonych (overdetermined), można zapisać w ogólnej postaci jako:

$$\mathbf{Ax} = \mathbf{b} + \mathbf{n}_e = \mathbf{b}_t, \quad (4.2c)$$

gdzie:

$$\begin{aligned} \mathbf{n}_e &\in \mathbf{R}^m && - \text{nieznany wektor szumów lub błędów pomiarowych,} \\ \mathbf{b}_t &\in \mathbf{R}^n && - \text{wektor prawdziwych, lecz nieznanymi wartościami.} \end{aligned}$$

Zadaniem jest znalezienie takiego wektora $\mathbf{x} \in \mathbf{R}^n$, który minimalizuje funkcję celu (energetyczną) [11]:

$$E(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_p = \|\mathbf{r}(\mathbf{x})\|_p, \quad p \geq 1, \quad (4.3)$$

gdzie wektor błędu (residual error vector), dla aktualnych wartości wektora $\mathbf{x}$ wynosi:

$$\mathbf{r}(\mathbf{x}) = [r_1(\mathbf{x}), r_2(\mathbf{x}), \dots, r_m(\mathbf{x})]^T,$$

i składa się z następujących części:

$$r_i(\mathbf{x}) = \mathbf{a}_i^T \mathbf{x} - b_i = \sum_{j=1}^n a_{ij} x_j - b_i \quad (i=1, 2, \dots, m) \quad (4.4)$$

a $\|\mathbf{r}\|_p$ jest normą L_p wektora $\mathbf{r}$.

Rozwiązanie $\mathbf{x}^*$ według normy L_p spełnia równanie:

$$\mathbf{Ax}^* = \mathbf{b} + \mathbf{r}(\mathbf{x}^*), \quad (4.5)$$

tak więc, wektor $\mathbf{x}^*$ minimalizuje według normy L_p wektor błędu $\mathbf{r}(\mathbf{x})$, spełniając relację:

$$\|\mathbf{Ax}^* - \mathbf{b}\|_p \leq \|\mathbf{Ax} - \mathbf{b}\|_p \quad \forall \mathbf{x} \in \mathbf{R}^n. \quad (4.6)$$

W praktyce, w zależności od przyjętej wartości p , można wyróżnić kilka przypadków, które zostały omówione w rozdziale 3.3.

W standardowym przybliżeniu najmniejszych kwadratów LS ($p = 2$) żąda się, aby elementy macierzy $\mathbf{A}$ były **wolne** od błędów, a wszystkie błędy - zawarte w wektorze obserwacji $\mathbf{b}$ [11, 12].

Przy takim założeniu, definiuje się odpowiednią funkcję energetyczną $E(\mathbf{x})$ jako:

$$\begin{aligned} E(\mathbf{x}) &= \frac{1}{2} \|\mathbf{Ax} - \mathbf{b}\|_2^2 + \frac{1}{2} \alpha \|\mathbf{x}\|_2^2 = \\ &= \frac{1}{2} (\mathbf{Ax} - \mathbf{b})^T (\mathbf{Ax} - \mathbf{b}) + \frac{1}{2} \alpha \mathbf{x}^T \mathbf{x} = \quad , \quad (4.7) \\ &= \frac{1}{2} \sum_{i=1}^m r_i^2(\mathbf{x}) + \frac{1}{2} \alpha \sum_{j=1}^n x_j^2 \end{aligned}$$

gdzie błąd $r_i(\mathbf{x})$ jest definiowany wzorem (4.4), zaś $\alpha \geq 0$ jest parametrem regularyzacji. Pierwszy człon równania (4.7) określa się terminem *standardowy najmniejszy błąd kwadratowy*, który wymusza minimalizację sumy kwadratu błędów, drugi - *regularyzacja*, której celem jest wymuszenie gładkiego ograniczenia estymowanego rozwiązania $\mathbf{x}^*$ dla słabo-warunkowanych zagadnień.

Używając standardowej metody największego spadku do minimalizacji funkcji celu, powyższe zagadnienie może być przekształcone w układ liniowych równań różniczkowych [11, 12]:

$$\frac{d\mathbf{x}}{dt} = -\mu [\mathbf{A}^T (\mathbf{Ax} - \mathbf{b}) + \alpha \mathbf{x}] \quad (4.8)$$

z warunkiem początkowym $\mathbf{x}(0) = \mathbf{x}^{(0)}$ (przeważnie $\mathbf{x}(0) = \mathbf{0}$), gdzie:

$$\mu = \text{diag}(\mu_1, \mu_2, \dots, \mu_n), \quad \mu_j > 0 \quad \forall j.$$

Dokładna implementacja tego układu, wymaga użycia $m + n$ liniowych części procesorowych [12], co dla sieci neuronowych, rozwiązujących postawione zadania według kryterium LS, wymusza konieczność zastosowania nadmiernej liczby wbudowanych bloków (np. członów mnożących, sumatorów). W następnym rozdziale, zostanie zaproponowane nowe rozwiązanie - skuteczniejsze i bardziej ekonomiczne w realizacji.

Oprócz uwag wymienionych wyżej, należy w tym miejscu zaznaczyć, że wprowadzenie parametru α znacznie utrudnia implementację układu sieci. Intensywne badania symulacyjne sieci neuronowych omówione w pracach [5, 6, 7, 8, 56] nie stwierdziły

konieczności stosowania tego parametru dla klasy zagadnień poruszanych w niniejszej pracy.

4.3. Kryterium TLS

Jak już wspomniano, podstawowe kryterium LS jest optymalne tylko wówczas, gdy wszystkie błędy są zawarte w wektorze obserwacji $\mathbf{b}$ oraz są one o rozkładzie Gaussa. Elementy macierzy $\mathbf{A}$ przyjmuje się jako wolne od błędów. W praktyce elementy macierzy $\mathbf{A}$ także mogą być obciążone błędami [54].

Kryterium TLS, w przeciwieństwie do metody LS, jest metodą stosowaną do rozwiązywania nadokreślonego układu równań liniowych (4.1a), kiedy elementy zarówno wektora $\mathbf{b}$ jak i macierzy $\mathbf{A}$ nie są wolne od błędów [4, 22, 23, 54, 55, 58]. Podczas gdy kryterium LS umieszcza wszystkie błędy w wektorze obserwacji $\mathbf{b}$, kryterium TLS przyjmuje zakłócenia danych macierzy $\mathbf{A}$ przez macierz $\Delta\mathbf{A}$ oraz wektora obserwacji $\mathbf{b}$ przez wektor $\Delta\mathbf{b}$, a następnie próbuje je skompensować.

Zagadnienie TLS może być traktowane jako metoda optymalizacyjna stosowana w celu znalezienia wektora $\mathbf{x}_{\text{TLS}}^*$, który minimalizuje wyrażenie:

$$\|\Delta\mathbf{A}\|_F^2 + \|\Delta\mathbf{b}\|_F^2, \quad (4.9)$$

pod warunkiem zachowania wymaganej równości:

$$(\mathbf{A} - \Delta\mathbf{A})\mathbf{x}_{\text{TLS}}^* = (\mathbf{b} - \Delta\mathbf{b}),$$

gdzie wyrażenie $\|\Delta\|_F$ oznacza normę Frobeniusa wyrażenia Δ .

Innymi słowy, zagadnienie TLS dąży do:

$$\min_{[\mathbf{A}_0; \mathbf{b}_0] \in \mathbf{R}^{m \times (n+1)}} \|\mathbf{A}; \mathbf{b} - [\mathbf{A}_0; \mathbf{b}_0]\|_F \quad (4.10)$$

pod warunkiem, że zakres $(\mathbf{b}_0) \subset \text{zakres}(\mathbf{A}_0)$.

Głównym narzędziem rozwiązania tego problemu jest rozkład według wartości osobliwych SVD (Singular Value Decomposition) [45, 53, 54, 55]. SVD jest bardzo ważnym narzędziem matematycznym, stosowanym szeroko np. w przetwarzaniu i analizie sygnałów, analizie danych, w szczególności gdy macierz jest słabo warunkowana.

Niech $\mathbf{A}$ będzie $m \times n$ macierzą rzeczywistą. Zakładamy dla uproszczenia, że $m \geq n$, wówczas macierz $\mathbf{A} \in \mathbf{R}^{m \times n}$ ma rozkład SVD w postaci:

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T = \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \quad (4.11)$$

gdzie wektory $\mathbf{u}_i$ oraz $\mathbf{v}_i$ macierzy $\mathbf{A}$ są ortonormalnymi kolumnami macierzy $\mathbf{U} \in \mathbf{R}^{m \times n}$ oraz $\mathbf{V} \in \mathbf{R}^{n \times n}$, natomiast wartości osobliwe $\sigma_i \geq 0$ są diagonalnymi elementami macierzy $\mathbf{S}$, tzn.:

$$\mathbf{S} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n) \in \mathbf{R}^{n \times n}.$$

Generalnie spełniają one relację: $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$, i jeśli macierz $\mathbf{A}$ jest nieosobliwa to $\sigma_n > 0$.

Szczególnym, szeroko znanym przypadkiem SVD jest sytuacja, kiedy macierz $\mathbf{A} \in \mathbf{R}^{n \times n}$ jest symetryczna i dodatnio określona. Wówczas, można zapisać że:

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T, \quad (4.12)$$

gdzie $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ oraz $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ są to wartości własne macierzy $\mathbf{A}$ odpowiadające wektorom własnym $\mathbf{v}_i$ ($i = 1, 2, \dots, n$) macierzy $\mathbf{V}$.

Niech $\mathbf{v}_{n+1} = [\mathbf{v}_{1,n+1}, \mathbf{v}_{2,n+1}, \dots, \mathbf{v}_{n+1,n+1}]^T$ będzie osobliwym wektorem, odpowiadającym najmniejszej wartości osobliwej σ_{n+1} rozszerzonej macierzy $[\mathbf{A}, \mathbf{b}]$.

Rozwiązanie (według kryterium TLS) $\mathbf{x}_{\text{TLS}}^*$ otrzymuje się jako [22]:

$$\mathbf{x}_{\text{TLS}}^* = - \frac{1}{\mathbf{v}_{n+1,n+1}} \cdot [\mathbf{v}_{1,n+1}, \mathbf{v}_{2,n+1}, \dots, \mathbf{v}_{n,n+1}]^T. \quad (4.13)$$

Jeśli wartość osobliwa σ_{n+1} dąży do zera, to rozwiązania zarówno LS jak i TLS zbliżają się do siebie [54]. Należy w tym miejscu zwrócić uwagę na to, że zagadnienie LS minimalizuje sumę kwadratów reszt (4.4), czyli funkcję energetyczną (3.13).

Zagadnienie TLS minimalizuje natomiast sumę kwadratów ważonych reszt [54]:

$$\mathbf{r}_{\text{TLS}}(\mathbf{x}) = \frac{\mathbf{Ax} - \mathbf{b}}{\left[1 + \mathbf{x}^T \mathbf{x}\right]^{\frac{1}{2}}},$$

co jest równoważne z minimalizacją ze względu na wektor $\mathbf{x}$ funkcji energetycznej:

$$E_{\text{TLS}}(\mathbf{x}) = \sum_{i=1}^m \left(\frac{r_i(\mathbf{x})}{\sqrt{1 + \mathbf{x}^T \mathbf{x}}} \right)^2 = \sum_{i=1}^m \frac{\left(\sum_{j=1}^n a_{ij} x_j - b_i \right)^2}{1 + \mathbf{x}^T \mathbf{x}}. \quad (4.14)$$

W porównaniu ze standardowym kryterium LS, otrzymanie rozwiązania TLS jest trudne i czasochłonne.

4.4. Adaptacyjny algorytm uczenia dla kryterium LS

W planowaniu algorytmu dla sieci neuronowych, kluczowym krokiem jest skonstruowanie odpowiedniej funkcji energetycznej $E(\mathbf{x})$ tak, że najmniejszy jej stan energetyczny będzie odpowiadał pożądanemu rozwiązaniu $\mathbf{x}^*$ [11].

Określenie tej funkcji umożliwi przekształcenie zagadnienia minimalizacyjnego w system równań różniczkowych, na bazie których może zostać zaprojektowana odpowiednia sieć neuronowa z towarzyszącym jej algorytmem uczenia.

W tym celu, skonstruowana została funkcja błędu:

$$e(t) = \mathbf{S}^T (\mathbf{A}\mathbf{x} - \mathbf{b}) = \mathbf{S}^T \mathbf{r}(\mathbf{x}) = \sum_{i=1}^m S_i(t) r_i(\mathbf{x}), \quad (4.15)$$

gdzie:

$r_i(\mathbf{x})$ - zgodnie z (4.4),

$\mathbf{S} = [S_1(t), S_2(t), \dots, S_m(t)]^T$ - jest to niezależne źródło wzbudzonych sygnałów o średnio-zerowym rozkładzie.

Funkcję tę można zapisać:

$$\begin{aligned} e(t) &= \sum_{i=1}^m S_i(t) r_i[\mathbf{x}(t)] = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j(t) - b_i \right) S_i(t) = \\ &= \sum_{j=1}^n \left(\sum_{i=1}^m a_{ij} S_i(t) \right) x_j(t) - \sum_{i=1}^m b_i S_i(t) = \\ &= \sum_{j=1}^n \tilde{a}_j(t) x_j(t) - \tilde{b}(t) \end{aligned} \quad (4.16)$$

$$\text{gdzie: } \tilde{a}_j(t) = \sum_{i=1}^m a_{ij} S_i(t),$$

$$\tilde{b}(t) = \sum_{i=1}^m b_i S_i(t).$$

Dla tak sformułowanego błędu $e(t)$ można skonstruować zmienną w czasie funkcję energetyczną w postaci:

$$E(x(t)) = \rho[e(t)], \quad (4.17)$$

gdzie ρ jest odpowiednio wypukłą funkcją, zwaną funkcją wagową lub funkcją straty a jej pochodna jest zwana funkcją aktywacji [11, 40, 24].

Minimalizacja tej funkcji prowadzi do układu równań różniczkowych:

$$\frac{dx_j(t)}{dt} = -\mu_j \frac{\partial E(t)}{\partial x_j} = -\mu_j \frac{\partial E}{\partial e} \frac{\partial e}{\partial x_j} = -\mu_j \Psi[e(t)] \tilde{a}_j(t), \quad (4.18a)$$

gdzie:

$$\mu_j > 0,$$

$$\Psi[e(t)] = \frac{\partial \rho[e(t)]}{\partial e(t)},$$

$$e(t) = \sum_{j=1}^n \tilde{a}_j(t) x_j(t) - \tilde{b}(t) \quad (j = 1, 2, \dots, n);$$

lub w postaci macierzowej:

$$\frac{dx(t)}{dt} = -\mu \Psi[e(t)] \tilde{a}(t), \quad (4.18b)$$

gdzie:

$$\mu = \text{diag}(\mu_1, \mu_2, \dots, \mu_n), \quad \mu_j > 0 \quad \forall j,$$

$$\tilde{a}(t) = [\tilde{a}_1(t), \tilde{a}_2(t), \dots, \tilde{a}_n(t)]^T.$$

Powyższy układ (4.18b) stanowi podstawowy adaptacyjny algorytm uczenia pojedynczego sztucznego neuronu.

Istnieją różne postaci funkcji strat $\rho(e)$, które mogą być zastosowane jako funkcje kosztów do minimalizacji funkcji energetycznej [40, 24], np. podstawiając $\rho(e) = e^2 / 2$ otrzymujemy kryterium najmniejszego błędu kwadratowego. W celu redukcji wpływu zakłóceń, w statystykach proponowanych jest kilka różnych postaci funkcji $\rho(e)$.

Można tutaj przedstawić trzy poniższe:

1) wartość bezwzględna funkcji:

$$\rho_A(e) = |e|;$$

2) funkcja Huber'a:

$$\rho_H(e) = \begin{cases} \frac{1}{2}e^2 & \text{dla } |e| \leq \beta \\ \beta|e| - \beta^2 / 2 & \text{dla } |e| > \beta \end{cases} ;$$

3) funkcja Talvar'a:

$$\rho_T(e) = \begin{cases} \frac{1}{2}e^2 & \text{dla } |e| \leq \beta \\ \beta^2 / 2 & \text{dla } |e| > \beta \end{cases} ,$$

gdzie $\beta > 0$ jest zadaniowo zależnym parametrem zwanym parametrem odcinającym (cut-off parameter).

Powyższe funkcje należą do klasy mocnych estymatorów, a ich celem jest, podczas trwania procesu estymacji parametrów, zredukowanie do minimum wpływu na ten proces pojawiających się nieoczekiwane błędów o dużej wartości lub tzw. „dzikich” zakłóceń impulsowych w sygnale. Celem zastosowania takich funkcji jest skondensowanie dużych aktualnych błędów, przez niedopuszczenie aby ich bezwzględna wartość była większa niż zalecana przez wartość parametru odcinającego.

Standardowe kryterium LS definiowane jest (4.7) następująco:

znaleźć wektor $\mathbf{x}_{LS}^$ który minimalizuje funkcję energetyczną:*

$$E_{LS}(\mathbf{x}) = \frac{1}{2}e^2(t) + \frac{\alpha}{2}\|\mathbf{x}\|_2^2, \quad (4.19)$$

gdzie $\alpha \geq 0$ oraz błąd $e(t)$ jest definiowany wzorem (4.16). Minimalizacja tego błędu, według metody największego spadku (metoda gradientowa), prowadzi do powstania poniższego algorytmu uczącego:

$$\frac{d\mathbf{x}(t)}{dt} = -\mu(t)[e(t)\tilde{\mathbf{a}}(t) + \alpha\mathbf{x}(t)], \quad (4.20)$$

a przy założeniu $\alpha = 0$ (p. 4.2):

$$\frac{d\mathbf{x}(t)}{dt} = -\mu(t)e(t)\tilde{\mathbf{a}}(t). \quad (4.20a)$$

4.5. Adaptacyjny algorytm uczenia dla kryterium TLS

Dla kryterium TLS sformułowanego w rozdziale 4.3 można skonstruować funkcję energetyczną w postaci (4.14):

$$E_{\text{TLS}}(\mathbf{x}) = \frac{1}{2} \frac{e^2(t)}{\mathbf{x}^T \mathbf{x} + 1}, \quad (4.21)$$

gdzie:

$$\mathbf{x} = [x_1(t), x_2(t), \dots, x_n(t)]^T,$$

$e(t)$ - zgodnie ze wzorem (4.16).

Stosując standardowy algorytm spadku gradientowego, otrzymujemy układ równań różniczkowych w postaci:

$$\begin{aligned} \frac{dx_j(t)}{dt} &= -\mu(t) \frac{\partial E_{\text{TLS}}(\mathbf{x})}{\partial x_j} = \\ &= -\mu(t) e(t) \frac{\tilde{a}_j(t)(\mathbf{x}^T \mathbf{x} + 1) - e(t)x_j(t)}{(\mathbf{x}^T \mathbf{x} + 1)^2}, \end{aligned} \quad (4.22)$$

gdzie $\mu(t) > 0$.

Po linearyzacji można go zapisać w uproszczonej postaci jako:

$$\frac{d\mathbf{x}(t)}{dt} = -\mu(t) e(t) [\tilde{\mathbf{a}}(t) + \tilde{\mathbf{b}}(t)\mathbf{x}(t)]. \quad (4.23)$$

Układ ten jest podstawowym równoległym algorytmem uczenia do rozwiązywania nadokreślonych równań liniowych (4.1a), zgodnie z przyjętym kryterium TLS.

Dla słabo uwarunkowanych zadań, funkcja energetyczna może być sformułowana jako:

$$E_{\text{TLS}}(\mathbf{x}, \alpha) = \frac{1}{2} \frac{e^2(t)}{\mathbf{x}^T \mathbf{x} + 1} + \frac{\alpha}{2} \|\mathbf{x}\|_2^2, \quad (4.24)$$

gdzie $\alpha \geq 0$. Uwzględniając (4.24), algorytm uczenia może zostać zmodyfikowany do postaci:

$$\frac{d\mathbf{x}(t)}{dt} = -\mu(t) \left\{ e(t) [\tilde{\mathbf{a}}(t) + \tilde{\mathbf{b}}(t)\mathbf{x}(t)] + \alpha \mathbf{x}(t) \right\}, \quad (4.25)$$

gdzie $\mu(t) > 0$.

5. Adaptacyjna sieć neuronowa do wyznaczania parametrów zakłóconego sygnału sinusoidalnego

5.1. Wprowadzenie

W wielu zastosowaniach praktycznych, pożądanym jest wyznaczenie parametrów sygnałów harmoniczych, zniekształconych przez szumy. Mierzone sygnały mają często zmienną w czasie amplitudę. Większość znanych algorytmów cyfrowych nie działa w pełni równolegle, tak więc prędkość ich przetwarzania jest ograniczona w czasie. W rozdziale tym będą zaproponowane nowe, w pełni równoległe algorytmy, realizowane jako adaptacyjne układy analogowe ze sprzężeniem zwrotnym, których działanie jest oparte na zasadach pracy sztucznych neuronów. Problem estymacji jest formułowany jako zagadnienie optymalizacyjne, rozwiązywane przy użyciu metod największego spadku (rozdział 3.2). Ponadto będą porównywane algorytmy bazujące na kryterium *Least Squares* (LS), *Total Least Squares* (TLS) oraz *Robust Total Least Squares* (RTLS).

Wyznaczenie parametrów (amplitud) składowych harmoniczych sygnału ma istotne znaczenie w systemach elektroenergetycznych i energoelektronice. Obecność wyższych harmoniczych jest spowodowana wzrostem zastosowań nieliniowych obciążeń dynamicznych. Występujące wyższe harmoniczne mają zazwyczaj zmienną w czasie amplitudę, wynikającą z dynamicznej natury obciążeń nieliniowych. Szybka w czasie estymacja parametrów ma podstawowe znaczenie w urządzeniach automatyki i zabezpieczeń systemów elektroenergetycznych. Jest ponadto bardzo przydatna w modelowaniu, miernictwie i kompensacji wyższych harmoniczych. Dotychczas były proponowane różnorodne cyfrowe i analogowe (sieci neuronowe) algorytmy do estymacji parametrów sygnałów harmoniczych. Można tu wymienić: standardowe metody najmniejszego błędu kwadratowego (LS) [44, 9], kryteria najmniejszej wartości bezwzględnej (LAV) [52, 9], normy L_p [10], minimax (Chebyshev norm) [52, 9, 10, 6], metody bazujące na rozkładzie według wartości osobliwej (SVD) [45], DFT [38, 19], filtry Kalmana [38].

Rozpatrywany jest zakłócony szumem sygnał:

$$y(t) = \sum_{i=1}^n a_i \sin(i\omega t) + \sum_{i=1}^n b_i \cos(i\omega t) + r(t), \quad (5.1)$$

gdzie:

- $\omega = 2\pi f$ - jest znaną lub przybliżenie oszacowaną pulsacją,
- $r(t)$ - jest nieznanym szumem lub błędem,
- a_i, b_i - są nieznanymi amplitudami sygnałów harmoniczných.

Na podstawie wartości sygnału $y(t)$ należy wyznaczyć w czasie rzeczywistym amplitudy a_i, b_i . Zakładając, że sygnał analogowy $y(t)$ jest próbkowany z okresem próbkowania T , zaś otrzymane próbki zapamiętane, powyższe zagadnienie może być matematycznie sformułowane jako zadanie do rozwiązania nadokreślonego układu równań liniowych w postaci:

$$\mathbf{D}\mathbf{x} = \mathbf{y}, \quad (5.2)$$

gdzie:

$$\mathbf{x} = [a_1, b_1, a_2, b_2, \dots, a_n, b_n]^T \in \mathbf{R}^{2n},$$

$$\mathbf{y} = [y(T), y(2T), \dots, y(mT)] \in \mathbf{R}^m,$$

$$\mathbf{D} = \begin{bmatrix} \sin(\omega T) & \cos(\omega T) & \cdots & \sin(n\omega T) & \cos(n\omega T) \\ \sin(2\omega T) & \cos(2\omega T) & \cdots & \sin(2n\omega T) & \cos(2n\omega T) \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \sin(m\omega T) & \cos(m\omega T) & \cdots & \sin(mn\omega T) & \cos(mn\omega T) \end{bmatrix} =$$

$$= [\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \mathbf{d}_4, \dots, \mathbf{d}_{2n-1}, \mathbf{d}_{2n}].$$

5.2. Kryterium LS

Powyższy układ może być rozwiązany przy użyciu standardowego zagadnienia LS. Zgodnie z nim, korzystając z zależności (3.13), minimalizujemy funkcję energetyczną zdefiniowaną jako:

$$E_2(\mathbf{x}) = \frac{1}{2} \|\mathbf{e}\|_2^2 = \frac{1}{2} \mathbf{e}^T \mathbf{e}, \quad (5.3)$$

gdzie:

$$\mathbf{e} = \mathbf{D}\mathbf{x} - \mathbf{y} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_{2n}] \mathbf{x} - \mathbf{y}.$$

W tym przypadku, szukany wektor $\mathbf{x}$ można zapisać jako:

$$\mathbf{x} = [\mathbf{D}^T \mathbf{D}]^{-1} \mathbf{D}^T \mathbf{y}. \quad (5.4)$$

Równanie (5.4) wymaga obliczenia macierzy odwrotnej, a to jest bardzo czasochłonne. W celu znalezienia rozwiązania w czasie rzeczywistym zgodnie z kryterium LS, można zastosować sieci neuronowe typu Hopfielda [3, 12]. Zgodnie ze standardowym algorytmem spadku gradientu dla kryterium LS, omówionym w rozdziale 3.2.2 oraz 3.3, zadanie sprowadza się do rozwiązania układu równań różniczkowych:

$$\frac{d\mathbf{x}}{dt} = -\mu \mathbf{D}^T (\mathbf{D}\mathbf{x} - \mathbf{y}) = -\mu \mathbf{D}^T \mathbf{e}, \quad (5.5)$$

gdzie:

$\mu > 0$ - jest krokiem uczenia.

Macierz równań różniczkowych (5.5) może być zapisana w postaci skalarnej jako:

$$\frac{da_i}{dt} = -\mu \sum_{k=1}^m e_k \sin(ik\omega T), \quad (5.6)$$

$$\frac{db_i}{dt} = -\mu \sum_{k=1}^m e_k \cos(ik\omega T), \quad (5.7)$$

gdzie:

$$e_k = \sum_{i=1}^n [a_i \sin(ik\omega t) + b_i \cos(ik\omega t)] - y(kT),$$

$$i = 1, 2, \dots, n; \quad k = 1, 2, \dots, m.$$

Stosowanie kryterium LS jest względnie proste. Otrzymane rozwiązanie jest jednakże optymalne **tylko** jeśli macierz $\mathbf{D}$ jest dokładnie znana a wektor $\mathbf{y}$ jest zakłócony szumem o rozkładzie normalnym (Gaussowskim) (rozdział 4.2). Metoda jest także wrażliwa na szumy o charakterze impulsowym.

5.3. Kryterium TLS

W praktyce, pulsacja ω nie jest dokładnie znana i może lekko oscylować podczas pomiarów, a wahania te nie są znane. Okres próbkowania może czasami nieznacznie się zmieniać (tj. próbkowanie sygnału nie jest idealnie regularne). Z tego powodu, w celu otrzymania dokładniejszych i pewniejszych rozwiązań, zastosowano kryterium *Total Least Squares* (TLS), znane w statystyce jako *orthogonal regression* lub *errors-in-variable regression*. Kryterium TLS zakłada występowanie błędów zarówno w macierzy $\mathbf{D}$ jak i w wektorze $\mathbf{y}$. Dlatego, podczas gdy metoda LS

minimalizuje przewidywany błąd; zagadnienie TLS minimalizuje błąd, który jest odległością (normalną) danego punktu obliczeniowego do liniowego wykresu przewidywanego rozwiązania [10, 53].

Zgodnie z kryterium TLS otrzymujemy układ równań liniowych w zapisie macierzowym:

$$(\hat{\mathbf{D}} + \mathbf{R})\mathbf{x}_{\text{TLS}} = \hat{\mathbf{y}} + \mathbf{r}, \quad (5.8)$$

gdzie:

$$\mathbf{D} = \hat{\mathbf{D}} + \mathbf{R} \text{ oraz } \mathbf{y} = \hat{\mathbf{y}} + \mathbf{r};$$

$$\hat{\mathbf{D}} \in \mathbb{R}^{m \times n}, \hat{\mathbf{y}} \in \mathbb{R}^m - \text{są dokładnymi lecz nieznanymi macierzami,}$$

$$\mathbf{R} \in \mathbb{R}^{m \times n}, \mathbf{r} \in \mathbb{R}^m - \text{są odpowiednimi błędami.}$$

Innymi słowy, zagadnienie TLS może być rozpatrywane jako problem optymalizacyjny:

znaleźć wektor $\mathbf{x}_{\text{TLS}}^$, który minimalizuje wyrażenie:*

$$\|\mathbf{R}\|_{\text{F}}^2 + \|\mathbf{r}\|_{\text{F}}^2 \quad (5.9)$$

zgodnie z równaniem (5.8), gdzie $\|\mathbf{R}\|_{\text{F}}$ oznacza normę Frobeniusa wyrażenia $\mathbf{R}$.

Stosując tryb postępowania zgodnie z wzorami: (4.11), (4.13)÷(4.16), (4.21)÷(4.23) otrzymamy adaptacyjny algorytm uczenia sieci do rozwiązywania równań liniowych (5.2), który można zapisać w postaci:

$$\frac{dx_j}{dt} = -\mu(t)e(t)[\tilde{d}_j(t) + \tilde{y}(t)x_j(t)]. \quad (5.10)$$

5.4. Odporne kryterium TLS (RTLS)

W praktyce, rzeczywiste rozkłady błędów występujące w układach, w mniejszym lub w większym stopniu, odbiegają od zakładanych. Algorytm TLS jest raczej czuły na szумы oraz na rodzaj rozkładu błędów, w szczególności, na obecność tzw. „dzikich” szumów. Ta niepożądana czułość algorytmu TLS, zmusza do jego modyfikacji, w celu eliminacji, jak dalece jest to możliwe, jego czułości na występowanie tego rodzaju zakłóceń oraz szumów o charakterze impulsowym.

Jest to główną motywacją do powstania i przebadania nowego uogólnionego algorytmu o nazwie *Robust Total Least Squares (RTLS)*.

Algorytm uczenia według równania (5.10) może być rozszerzony do postaci:

$$\frac{dx_j(t)}{dt} = -\mu(t)\Psi[e(t)][\tilde{d}_j(t) + \delta\tilde{y}(t)x_j(t)], \quad (5.11)$$

gdzie: $\Psi(e)$ - jest nieliniową funkcją aktywacji dająca możliwość tłumienia dużych błędów, np.: $\Psi(e) = \tanh(\gamma e)$ lub $\Psi(e) = \begin{cases} e & \text{dla } |e| \leq \beta \\ 0 & \text{pozostałe} \end{cases}$,

a $\delta \geq 0$ jest dodatnim współczynnikiem.

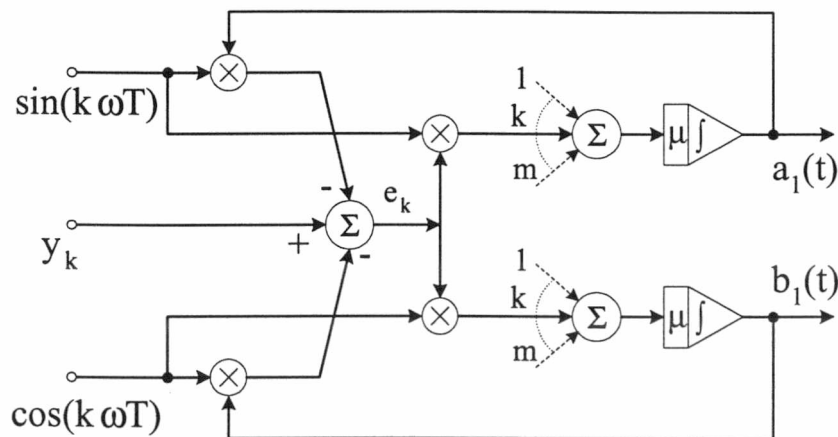
W szczególnym przypadku, gdy $\Psi(e) \equiv e$ oraz $\delta = 0$ otrzymamy standardowy algorytm LS, zaś w przypadku gdy $\Psi(e) \equiv e$ oraz $\delta = 1$ - standardowy algorytm TLS. Należy w tym miejscu zwrócić uwagę na to, że poprzez zmianę wartości parametru δ w większym lub mniejszym stopniu podkreślamy wzajemny udział występowania błędów w macierzy $\mathbf{D}$ oraz w wektorze $\mathbf{y}$. Jeśli $\delta = 0$, przyjmujemy, że błąd jest zawarty tylko w wektorze $\mathbf{y}$. W drugim przypadku, dla dużych wartości parametru δ (np. 100) można założyć, że wektor $\mathbf{y}$ jest wolny od błędów, zaś one zawarte są w macierzy danych $\mathbf{D}$. Taka ekstremalna sytuacja odnosi się do zagadnienia zwanego *Data Least Squares* (DLS).

5.5. Implementacja algorytmów

Przedstawione powyżej układy równań: (5.6)-(5.7), (5.10), (5.11) są podstawą do skonstruowania odpowiednich sieci neuronowych. Sieci te są analogową implementacją tych układów.

Na rys. 5.1 jest przedstawiony fragment sztucznej sieci neuronowej do wyznaczania amplitudy składowej podstawowej sygnału, stanowiącej rozszerzenie sieci typu Hopfielda. Implementuje ona układ równań (5.6)-(5.7). Sieć składa się z typowych elementów elektronicznych: sumatorów, układów mnożących, generatorów znaku, integratorów, generatorów funkcji trygonometrycznych, wzmacniaczy. W strukturze sieci możemy wyróżnić dwa integratory oraz m równoległych kanałów obliczeniowych. Na ich wejście dostarcza się próbki sygnału oraz estymowane w danej chwili przez układ wartości parametrów sygnału. Mamy tu do czynienia ze sprzężeniem zwrotnym, dlatego też sieci takie zalicza się do typu rekurencyjnego. Wszystkie m próbek sygnału doprowadza się do układu jednocześnie, co wymaga zastosowania odpowiedniego

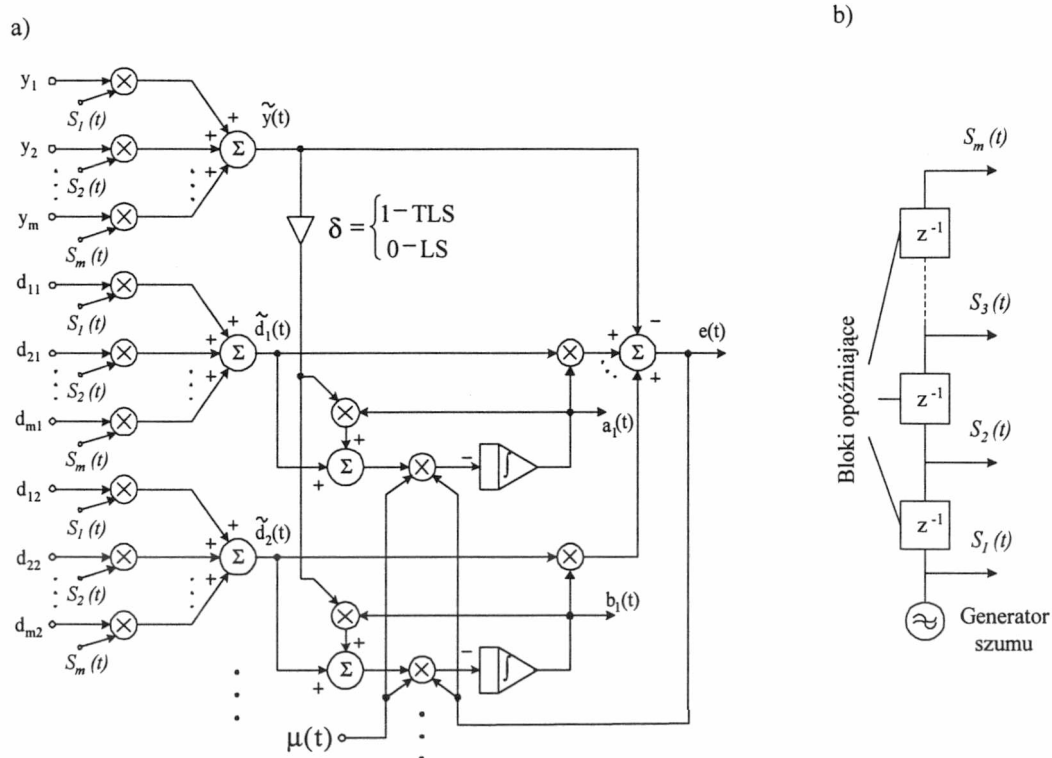
układu próbkująco-pamiętającego. Na wyjściu każdego z kanałów pojawia się sygnał błędu e_k . Sieć wyznacza tylko amplitudy a_1 oraz b_1 . Tak otrzymane wyniki będą stanowić podstawę do porównania z wynikami uzyskanymi przez sieć przedstawioną na rys. 5.2, skonstruowaną na nowym algorytmie (5.10).



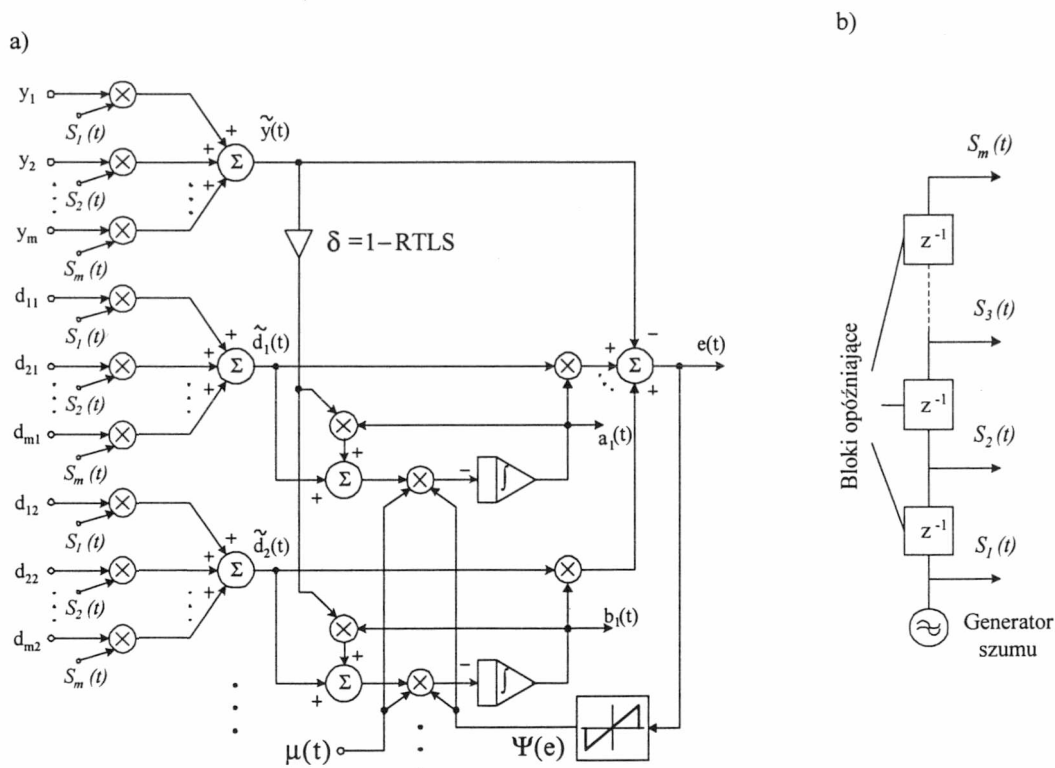
Rys. 5.1. Fragment sztucznej sieci neuronowej do wyznaczania amplitudy podstawowej składowej sygnału, opartej na standardowym kryterium LS

Rys. 5.2 przedstawia fragment schematu funkcjonalnego sieci implementującej algorytm opisany równaniem (5.10), którego celem jest wyznaczenie składowych podstawowych sygnału a_1 oraz b_1 . Schemat ten powinien być rozważany jako pojedynczy neuron x_j ($j = 1, 2$), uczony według równania (5.10). Sieć składa się z integratorów, sumatorów, układów mnożących. Na wejścia odpowiednich sumatorów podawane są **równocześnie** próbki sygnału mierzonego y_k oraz sygnału modelowego d_{kj} ($k = 1, 2, \dots, m$; $j = 1, 2$), pomnożone przez wartości sygnałów $S_k(t)$ pochodzących z niezależnego źródła (rys. 5.2b). Tak uzyskane na wyjściu sumatorów sygnały dostarcza się następnie na wejścia poszczególnych elementów układu: sumatorów, członów mnożących oraz integratorów. W końcowej fazie obliczeń, po scałkowaniu, na wyjściu integratorów otrzymuje się poszukiwane wartości składowych podstawowych sygnału a_1 i b_1 . W zależności od wartości parametru δ , sieć realizuje obliczenia według algorytmu opartego na kryterium LS (dla $\delta = 0$) lub - kryterium TLS (dla $\delta = 1$).

Rys. 5.3 przedstawia fragment schematu funkcjonalnego implementującego algorytm opisany równaniem (5.11) - kryterium RTLS.



Rys. 5.2. a) Fragment sieci neuronowej do wyznaczania parametrów sygnału opartej na algorytmie uczenia według równania (5.10),
b) Przykładowy sposób generowania sygnału $S_k(t)$



Rys. 5.3. a) Fragment sieci neuronowej do wyznaczania parametrów sygnału opartej na algorytmie uczenia według równania (5.11),
b) Przykładowy sposób generowania sygnału $S_k(t)$

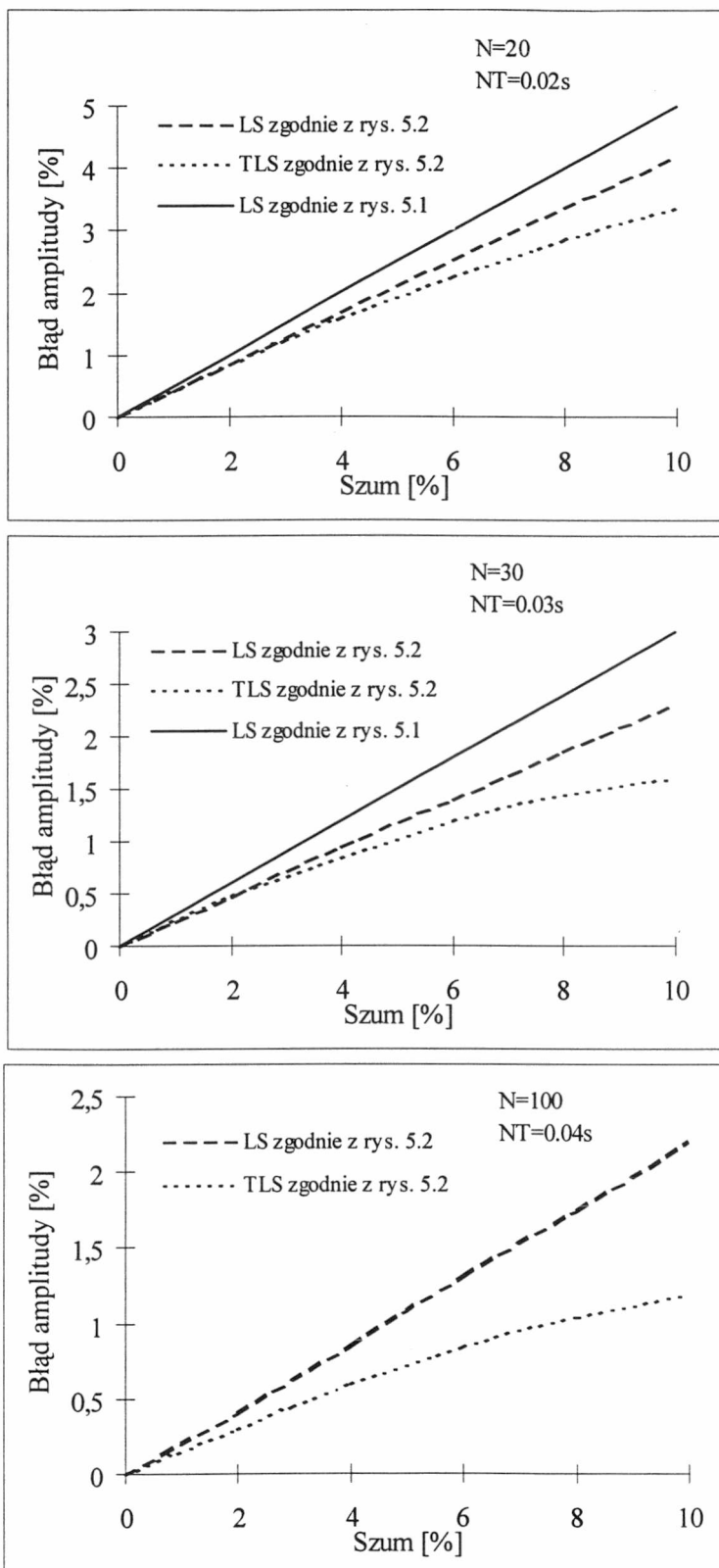
Do badań wybrano przykładowo sygnały: $y(t) = \sin(\omega t) + r(t)$ oraz $y(t) = 100\sin(\omega t) + 100\cos(\omega t) + r(t)$. Częstotliwość sygnału modelowego wynosiła 50 Hz. W pierwszej kolejności zbadano, jaki wpływ na dokładność pomiaru amplitudy, ma poziom zakłóceń w sygnale, długość okna pomiarowego NT oraz liczba próbek N doprowadzonych do sieci.

Rys. 5.6 przedstawia wykres zależności błędów estymacji amplitudy od poziomu zakłóceń szumem o rozkładzie normalnym dla sieci o oknie próbkowania NT: 20, 30 i 40 ms oraz liczbie próbek N odpowiednio: 20, 30, 100. Porównane zostały ze sobą układy skonstruowane zgodnie z rys. 5.1 (kryterium LS) oraz - z rys. 5.2 (nowe kryterium LS i TLS). Analizując wykresy można zauważyć, że we wszystkich przypadkach lepsze wyniki (większą odporność) na poziom zakłóceń szumami wykazała sieć skonstruowana na nowym algorytmie, według równania (5.10) - w szczególności kryterium TLS. Ponadto można zauważyć, że dokładność obliczeń w tej metodzie wzrasta wraz ze wzrostem częstotliwości próbkowania oraz długości okna próbkowania. Najlepsze wyniki pomiarów uzyskano dla częstotliwości próbkowania sygnału $f_s = 2500$ Hz oraz przy oknie próbkowania NT = 40 ms.

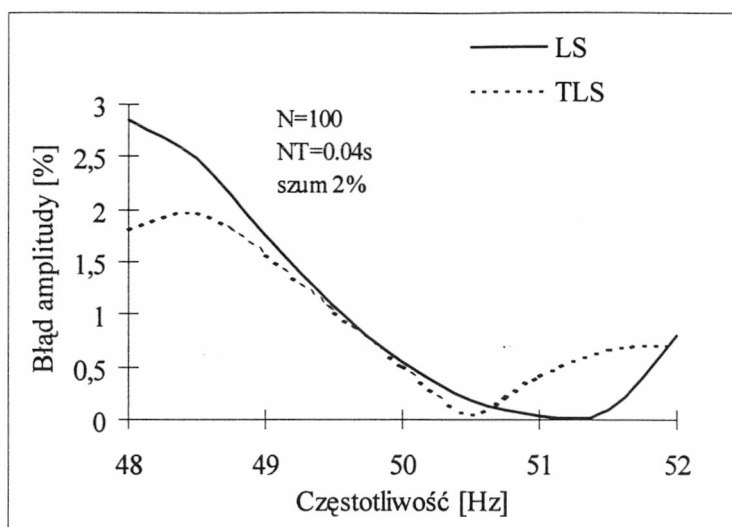
W kolejnym etapie pracy poddano dokładniejszemu badaniu sieci, działające według nowego algorytmu obliczeniowego, skonstruowanego na podstawie równania (5.10) - kryterium TLS. Przebadano układy o oknie próbkowania NT: 20, 30, 40, 60 i 80 ms oraz liczbie próbek $N = 20 + 100$. W każdym analizowanym przypadku, poziom szumów pomocniczych $S_k(t)$ wynosił 1% wartości amplitudy badanego sygnału. Sprawdzano jaki wpływ na dokładność pomiaru amplitudy ma: wahania częstotliwości mierzonego sygnału (rys. 5.7), występowanie w nim wyższych harmonicznym (rys. 5.8), zakłóceń impulsowych (rys. 5.10) oraz wartość współczynnika uczenia sieci (stała lub zmienna) (rys. 5.11). Podobnie jak w przypadku opisanym wyżej, najlepsze wyniki dostarcza sieć przy najdłuższym oknie próbkowania NT = 40 ms i częstotliwości próbkowania $f_s = 2500$ Hz.

Rys. 5.7 przedstawia błędy estymacji amplitudy w zależności od wahań częstotliwości, dla sygnału zakłóconego 2% szumem o rozkładzie normalnym. Badano sieć opracowaną dla podstawowej częstotliwości sygnału równej 50 Hz. Do sieci doprowadzano próbki sygnału o częstotliwości podanej na osi odciętych. Dokładniejsze wyniki, w przypadku sygnałów o częstotliwości podstawowej mniejszej niż 50.5 Hz, dostarcza kryterium TLS. Błędy wyznaczonej amplitudy, przy częstotliwościach z

przedziału $49.5 \text{ Hz} \div 52 \text{ Hz}$, zarówno w przypadku kryterium LS jak i TLS, nie przekraczały wartości 1%.

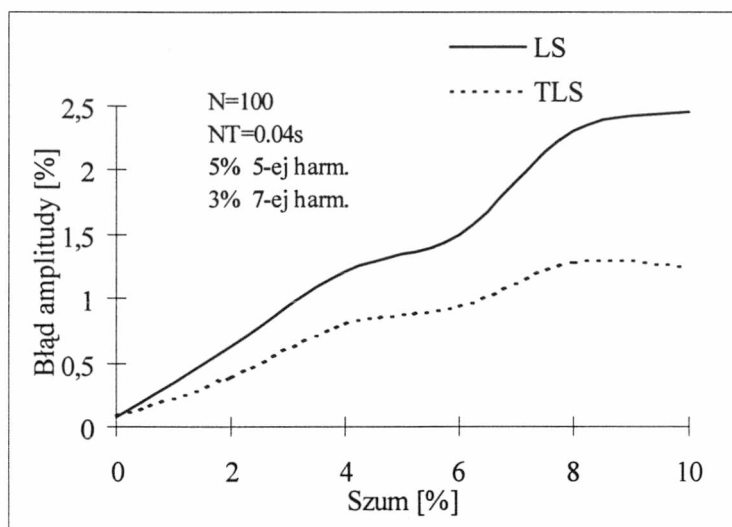


Rys. 5.6. Błędy estymacji amplitudy przy zastosowaniu sieci o algorytmie opartym na:
a) kryterium LS zgodnie z rys. 5.1,
b) kryterium LS i TLS zgodnie z rys. 5.2



Rys. 5.7. Błędy estymacji amplitudy sygnału $y(t) = \sin(\omega t) + r(t)$ zakłóconego 2% szumem o rozkładzie normalnym

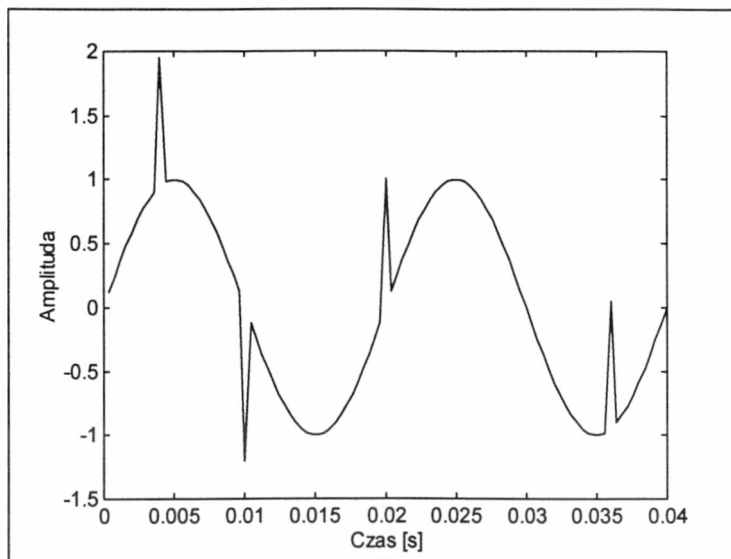
Przeprowadzone badania symulacyjne pozwoliły ponadto stwierdzić, że jeśli w badanym sygnale występują wyższe harmoniczne, to kryterium TLS jest odporniejsze na szumy w porównaniu z kryterium LS. Przykładowo, rys. 5.8 przedstawia wyniki dla przypadku, w którym analizie poddano sygnał zawierający, oprócz składowej podstawowej, składowe harmoniczne: 5-ą (5% wartości amplitudy podstawowej składowej) oraz 7-ą (3%).



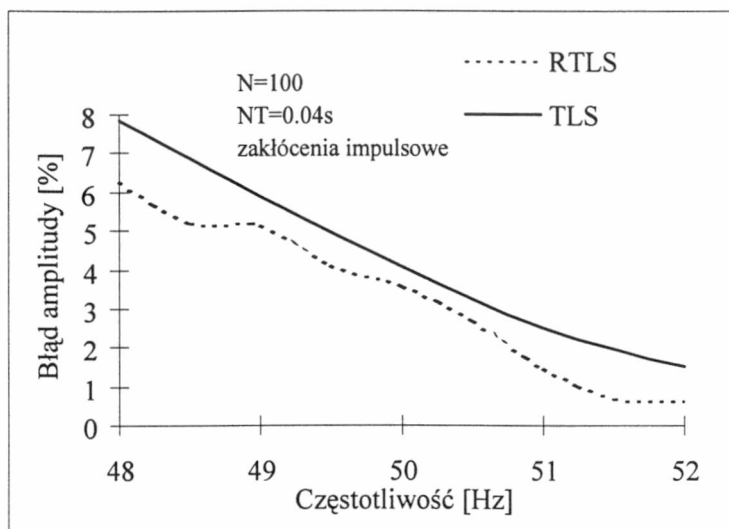
Rys. 5.8. Błędy estymacji amplitudy sygnału $y(t) = 100 \sin(\omega t) + 100 \cos(\omega t) + r(t)$ zawierającego 5-ą i 7-ą harmoniczną

Rys. 5.10 przedstawia wyniki pomiarów dla sytuacji, w której badany sygnał był zakłócony impulsowo (rys.5.9). Analiza wyników pozwala stwierdzić, że odporniejsze jest kryterium RTLS w porównaniu do kryterium TLS, co potwierdziło cel i założenia

teoretyczne omówione w rozdziale 5.4. Takiej zależności nie zaobserwowano w sytuacji, kiedy sygnał był zakłócony szumem o rozkładzie równomiernym.

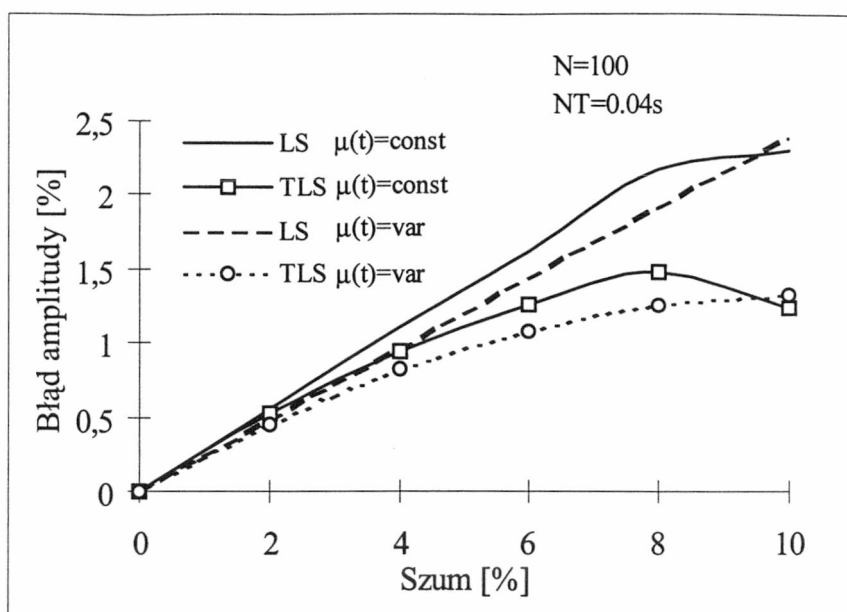


Rys. 5.9. Przykładowy przebieg sygnału zakłóconego impulsowo

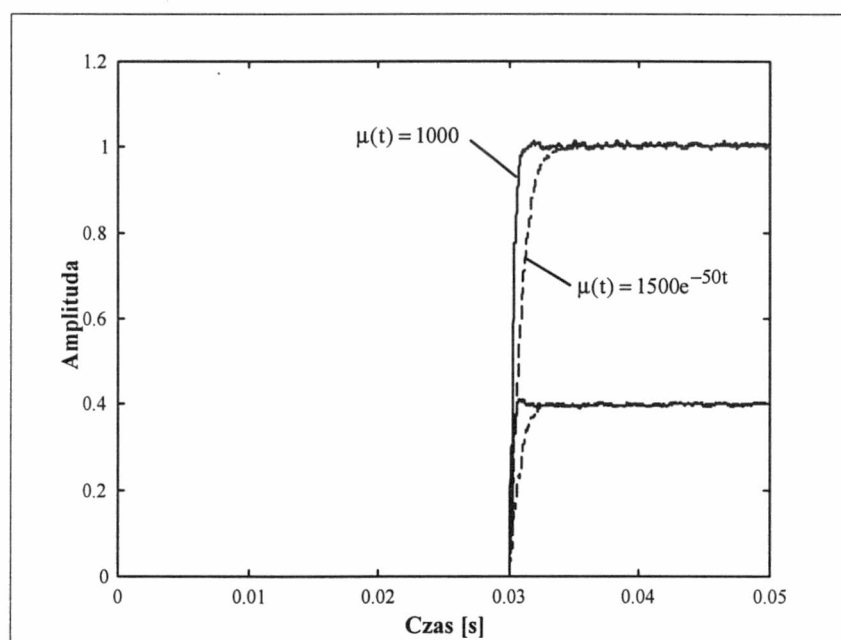


Rys. 5.10. Błędy estymacji amplitudy sygnału zawierającego zakłócenia impulsowe, wyznaczone przez sieć działającą według kryterium TLS oraz RTLS

Ważnym czynnikiem, wpływającym na dokładność estymacji parametrów sygnału, jest odpowiedni dobór współczynnika uczenia sieci μ . W badaniach brano pod uwagę różne jego wartości: stałe lub zmienne (malejące asymptotycznie) w czasie. Dobre wyniki estymacji, w porównaniu do stałej wartości współczynnika uczenia $\mu(t) = 1500 = \text{const}$, uzyskano dla współczynnika: $\mu(t) = 1500 \exp(-50t)$. Przykładowo ilustruje to rys. 5.11. Należy jednak podkreślić, że sposób jego wyboru wpływa również na czas ustalania się oraz kształt trajektorii estymowanego sygnału. W przypadku $\mu(t) = \text{const}$ czas ten jest krótszy (rys. 5.12).

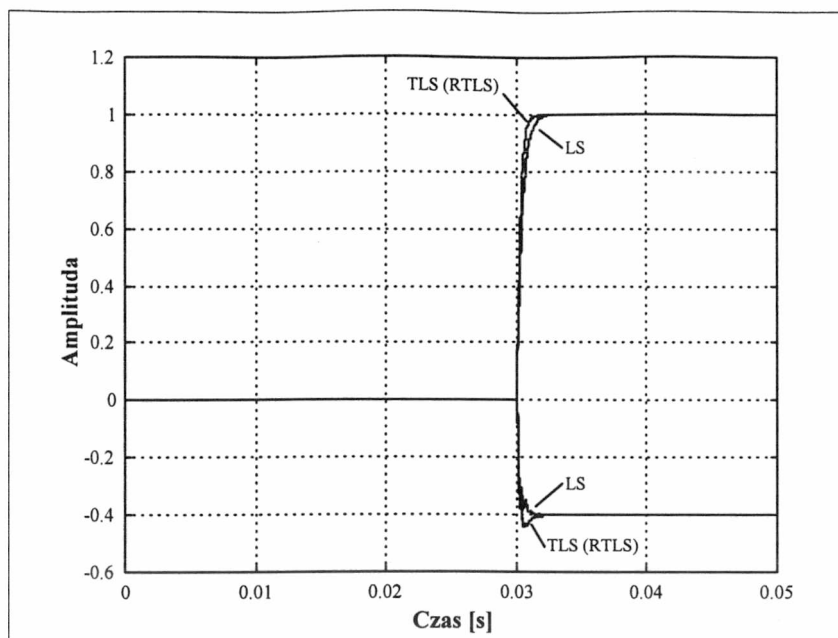


Rys. 5.11. Błędy estymacji amplitudy sygnału w zależności od doboru kroku uczenia sieci



Rys. 5.12. Przykładowy przebieg trajektorii wyznaczanych parametrów sygnału $y(t) = \sin \omega t + 0.4 \cos \omega t$ dla różnych postaci współczynnika uczenia sieci; okno próbkowania $NT = 60$ ms; $t = 0.03$ s - czas rozpoczęcia przez sieć estymacji parametrów sygnału

Zauważono ponadto, że w przypadku stosowania kryterium TLS i RTLS czas ustalania się trajektorii wyznaczanych parametrów sygnału był krótszy niż w przypadku stosowania kryterium LS (rys. 5.13).



Rys. 5.13. Przykładowy przebieg trajektorii wyznaczanych parametrów sygnału $y(t) = \sin \omega t - 0.4 \cos \omega t$ dla różnych kryteriów estymacji; pozostałe parametry sieci zgodnie z rys. 5.12

Podsumowując wyniki pomiarów można zauważyć, że:

1. Sieć skonstruowana na podstawie standardowego algorytmu (LS) (rys. 5.1):
 - a) wykazała najmniejszą odporność na poziom zakłóceń sygnału szumami o rozkładzie normalnym, w porównaniu do sieci przedstawionych w p-ch 2, 3 oraz 4;
2. Sieć skonstruowana na podstawie nowego algorytmu (rys. 5.2 - kryterium LS) wykazała:
 - a) większą odporność na poziom zakłóceń szumami o rozkładzie normalnym, aniżeli sieć omówiona wyżej;
 - b) mniejszą dokładność pomiarów dla wszystkich rodzajów badań oraz struktur sieci opisanych w p-ch 3 oraz 4;
3. Sieć skonstruowana na podstawie nowego algorytmu (rys. 5.2 - kryterium TLS):
 - a) posiada największą odporność na poziom zakłóceń szumami o rozkładzie normalnym, w porównaniu z sieciami omówionymi wyżej;
 - b) wykazuje największą odporność na poziom zakłóceń szumem w sygnale zawierającym wyższe harmoniczne;
 - c) dostarcza dokładniejszych wyników estymacji amplitudy w przypadku wystąpienia wahań częstotliwości w badanym sygnale - błąd nie przekracza wartości 1 %;
 - d) jest wrażliwa na zakłócenia impulsowe;
 - e) ma czas konwergencji trajektorii wyznaczanych parametrów krótszy, niż w przypadku stosowania kryterium LS;

4. Sieć skonstruowana na podstawie nowego algorytmu (rys. 5.3 - kryterium **RTLS**) charakteryzuje się:
- a) identycznymi uwagami jak w p-ch 3a), b), c) oraz e);
 - b) lepszą odpornością na zakłócenia impulsowe w porównaniu do kryterium TLS.

Dla wszystkich omówionych wyżej sieci można stwierdzić, że:

- a) dokładność estymacji amplitudy zależy od częstotliwości próbkowania oraz od długości okna próbkowania - wzrasta wraz ze wzrostem tych parametrów,
- b) spośród badanych sieci najlepsze wyniki estymacji amplitudy dostarcza sieć o najszerszym oknie próbkowania $NT = 40$ ms (częstotliwość próbkowania $f_s = 2500$ Hz),
- c) wpływ na dokładność estymacji parametrów ma odpowiedni dobór współczynnika uczenia sieci μ ; lepsze wyniki estymacji uzyskano dla $\mu(t) = 1500 \exp(-50t)$, aniżeli dla $\mu(t) = 1500 = \text{const.}$

6. Sieć neuronowa do detekcji stanu asynchronicznego maszyn synchronicznych

6.1. Wprowadzenie

Identyfikacja zjawiska utraty synchronizmu w maszynach synchronicznych ma podstawowe znaczenie dla ich ochrony oraz kontroli stanu pracy. Podczas zwarcia w pobliżu generatora, na jego zaciskach pojawia się nadwyżka mocy. Jest to spowodowane: po pierwsze - przekazywaniem mocy wyjściowej do obwodu o charakterze głównie indukcyjnym, po drugie - brakiem dostatecznie szybkiej zmiany mocy wejściowej dostarczanej przez turbinę. Nadmiar mocy powoduje wzrost prędkości wirnika. Jeśli awaria trwa dostatecznie długo, kąt mocy wzrasta do wartości, przy której następuje utrata synchronizmu.

Pracę asynchroniczną maszyny synchronicznej można traktować jako jeden z przypadków pracy nieustalonej, podczas której wszystkie wielkości, zarówno elektryczne jak i mechaniczne, ulegają zmianie na skutek istnienia momentów pulsujących. Niekiedy jednak, w analizie pracy asynchronicznej zakłada się stałość prędkości kątowej [47], tak więc pracę asynchroniczną maszyny synchronicznej, traktuje się jako pracę quasi stacjonarną.

Po utracie synchronizmu przez maszynę synchroniczną w uzwojeniach wirnika indukują się siły elektromotoryczne o pulsacji proporcjonalnej do poślizgu [47]. Prądy w uzwojeniach wzbudzenia płynące pod wpływem tych sił, wywołują pulsujący strumień magnetyczny, który może być interpretowany jako superpozycja dwóch strumieni wirujących w przeciwnych kierunkach. Indukują one w uzwojeniach stojana odpowiednie składowe prądów (rozdział 6.2). Trzecią składową prądu indukuje prąd wzbudzenia. Te zjawiska występują również przy utracie synchronizmu przez silnik synchroniczny.

Identyfikację stanu asynchronicznego maszyny synchronicznej można przeprowadzić wyznaczając amplitudy poszczególnych składowych prądu stojana, przy założeniu stałej i znanej wartości poślizgu. Zakłada się przy tym, że przyjęta wartość zostanie, po utracie synchronizmu, osiągnięta i zapewne przekroczona. W rozdziale omówione zostaną sztuczne sieci neuronowe realizujące powyższe zadanie, których

algorytm działania oparty jest na kryterium najmniejszego błędu kwadratowego (rozdz. 3.3) oraz metodzie największego spadku (rozdz. 3.2.2).

6.2. Przebiegi prądu uzwojenia stojana

Jak już wspomniano wcześniej, jeżeli maszyna pracująca równolegle z innymi „wypadnie” z warunków stanu synchronicznego, to strumienie uzwojenia wirnika i stojana wirują z różnymi prędkościami. Strumień uzwojenia stojana generuje siły elektromotoryczne w uzwojeniu wirnika, których prędkość kątowna zależy od wartości jego poślizgu. Prądy w uzwojeniach wirnika, wywołane tymi siłami elektromotorycznymi, indukują pulsujący strumień magnetyczny, który może być rozdzielony na dwie składowe wirujące w przeciwnych kierunkach. Prędkości kątowne strumieni są zależne od poślizgu wirnika:

$$\omega_f = s \cdot \omega_s; \quad \omega_b = -s \cdot \omega_s, \quad (6.1)$$

gdzie:

- ω_f - prędkość kątowna składowej pola magnetycznego, wirującej zgodnie z polem stojana,
- ω_b - prędkość kątowna składowej pola, wirującej w kierunku przeciwnym,
- ω_s - prędkość kątowna wirującego pola magnetycznego stojana,
- s - poślizg wirnika.

Prędkość kątowna wirnika opisana jest zależnością:

$$\omega_r = (1 - s)\omega_s. \quad (6.2)$$

Składowe pola przecinają uzwojenia stojana z prędkościami, które zależą od prędkości tych składowych i prędkości wirnika. W uzwojeniach stojana są indukowane siły elektromotoryczne, wywołujące przepływ składowych prądu o pulsacjach odpowiednio: ω_s i $(1 - 2s)\omega_s$.

Prąd stały w uzwojeniach wirnika jest źródłem siły magnetomotorycznej, która również wytwarza strumienie magnetyczne w szczelinie. Wirują one z prędkością ω_r , przecinają uzwojenia stojana z prędkością poślizgu, indukują siły elektromotoryczne oraz powodują pojawienie się drugiej składowej prądu stojana.

Przyjmując stan ustalony pracy asynchronicznej maszyny synchronicznej ($s = \text{const}$), prąd uzwojenia stojana może być wyrażony jako:

$$i_s = I_f \cos(\omega_s t + \gamma_f) + I_b \cos[(1 - 2s)\omega_s t + \gamma_b] + I_r \cos[(1 - s)\omega_s t + \gamma_r], \quad (6.3)$$

gdzie:

$I_f, I_b, I_r, \gamma_f, \gamma_b, \gamma_r$ amplitudy oraz fazy składowych prądu stojana.

Przykładowy, symulacyjny przebieg prądu w uzwojeniach stojana oraz jego składowych, przedstawiono na rysunku 6.1.

6.3. Algorytm działania sieci

Sygnał opisany równaniem (6.3), można przedstawić w następujący sposób:

$$x(t) = X_a \sin(\omega t) + X_b \cos(\omega t) + X_c \sin[(1 - s)\omega t] + X_d \cos[(1 - s)\omega t] + X_e \sin[(1 - 2s)\omega t] + X_f \cos[(1 - 2s)\omega t], \quad (6.4)$$

gdzie:

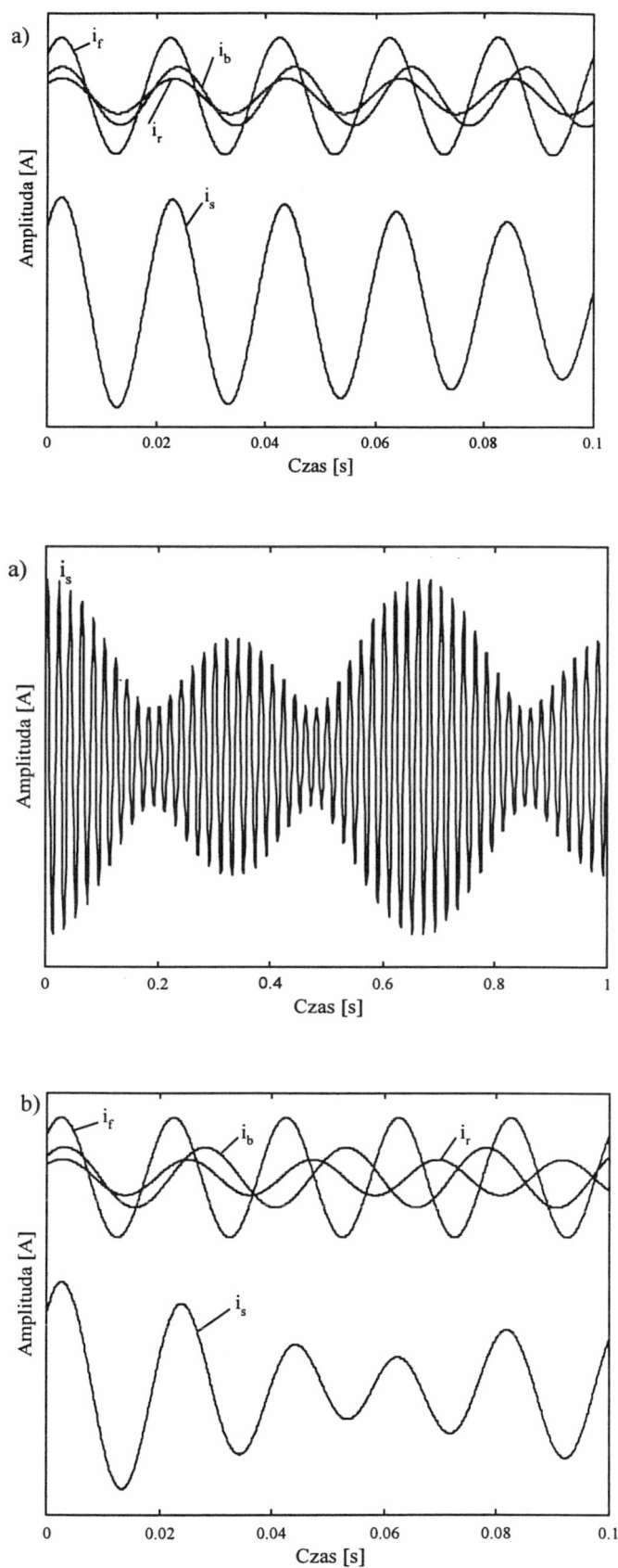
$X_{a,...,f}$ - amplitudy składowych sygnału,

ω - znana pulsacja podstawowej składowej sygnału (50 Hz).

W wyniku utraty przez maszynę stanu pracy synchronicznej, następuje stosunkowo wolna zmiana poślizgu wirnika, z uwagi na znaczne wartości elektromechanicznych stałych czasowych. Wraz z upływem czasu następuje wzrost wartości poślizgu. W celu identyfikacji stanu asynchronicznego maszyny, przyjęto stałą wartość poślizgu, która zostanie osiągnięta, w wyniku stopniowego wzrostu jego wartości.

Z praktycznego punktu widzenia założono, że prąd stojana $y(t)$ jest mierzony w skończonym przedziale czasu, w wyniku czego otrzymano N próbek mierzonego sygnału $y_m = y(mT)$ (T - okres próbkowania). Liczba próbek przekracza liczbę składowych mierzonego prądu. Zagadnienie estymacji może być sformułowane w następujący sposób:

wyznaczyć wektor $\mathbf{X} = [X_a, X_b, X_c, X_d, X_e, X_f]$, który minimalizuje odpowiednio dobraną funkcję energetyczną $E(\mathbf{X})$.



Rys. 6.1. Składowe i_f , i_b , i_r oraz przebieg całkowitego prądu stojana i_s maszyny synchronicznej w stanie asynchronicznym; poślizg: a) 0.03, b) 0.1 (symulacja)

W celu rozwiązania tak postawionego zadania, zastosowano standardowe kryterium najmniejszego błędu kwadratowego (norma L_2) (rozdz. 3.3) opisane równaniem (3.13). Stosując algorytm optymalizacyjny największego spadku (rozdz. 3.2.2) otrzymujemy następujący układ nieliniowych równań różniczkowych [44, 8, 9, 10, 60]:

$$\frac{dX_a}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \sin(m\omega T), \quad \frac{dX_b}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \cos(m\omega T), \quad (6.5a, b)$$

$$\frac{dX_c}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \sin(m\omega_1 T), \quad \frac{dX_d}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \cos(m\omega_1 T), \quad (6.5c, d)$$

$$\frac{dX_e}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \sin(m\omega_2 T), \quad \frac{dX_f}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \cos(m\omega_2 T), \quad (6.5e, f)$$

gdzie:

$$\omega_1 = (1-s)\omega,$$

$$\omega_2 = (1-2s)\omega,$$

τ - stała czasowa integratora.

Ze względu na zmieniający się w funkcji czasu poślizg, powyższy algorytm opisany równaniami (6.5a, ..., f) można rozszerzyć o dalsze równania. Umożliwiłyby one wyznaczanie amplitud $X_g, \dots, X_j$ składowych, pojawiających się przy wyższych wartościach poślizgu maszyny. Podwyższyłoby to niezawodność identyfikacji stanu asynchronicznego, szczególnie w przypadkach dłuższego czasu trwania takiego stanu. Powiększając układ równań różniczkowych o dalsze cztery równania otrzymuje się:

$$\frac{dX_a}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \sin(m\omega T), \quad \frac{dX_b}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \cos(m\omega T), \quad (6.6a, b)$$

$$\frac{dX_c}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \sin(m\omega_1 T), \quad \frac{dX_d}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \cos(m\omega_1 T), \quad (6.6c, d)$$

$$\frac{dX_e}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \sin(m\omega_2 T), \quad \frac{dX_f}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \cos(m\omega_2 T), \quad (6.6e, f)$$

$$\frac{dX_g}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \sin(m\omega_3 T), \quad \frac{dX_h}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \cos(m\omega_3 T), \quad (6.6g, h)$$

$$\frac{dX_i}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \sin(m\omega_4 T), \quad \frac{dX_j}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \cos(m\omega_4 T), \quad (6.6i, j)$$

gdzie:

ω_1, ω_2, τ - zgodnie ze wzorem (6.5),

$\omega_3 = (1 - 3s)\omega$,

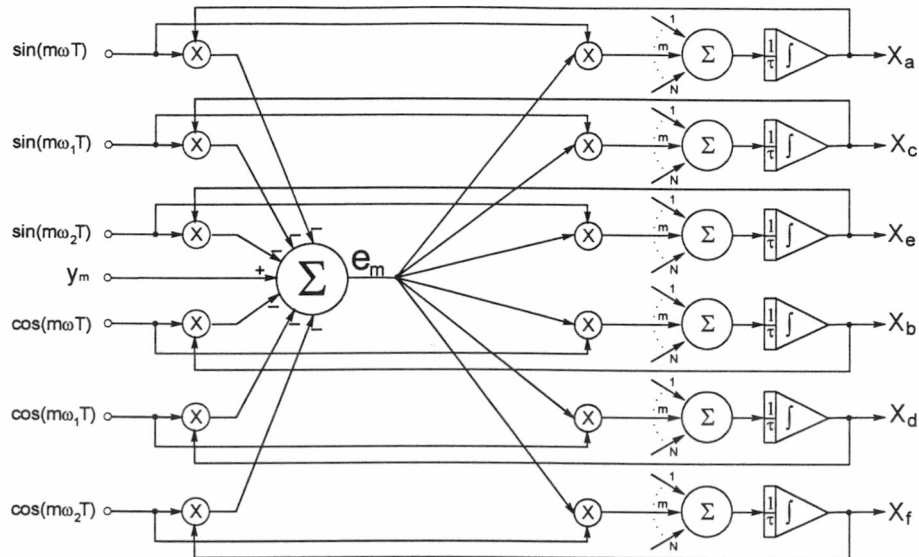
$\omega_4 = (1 - 4s)\omega$.

6.4. Implementacja algorytmu

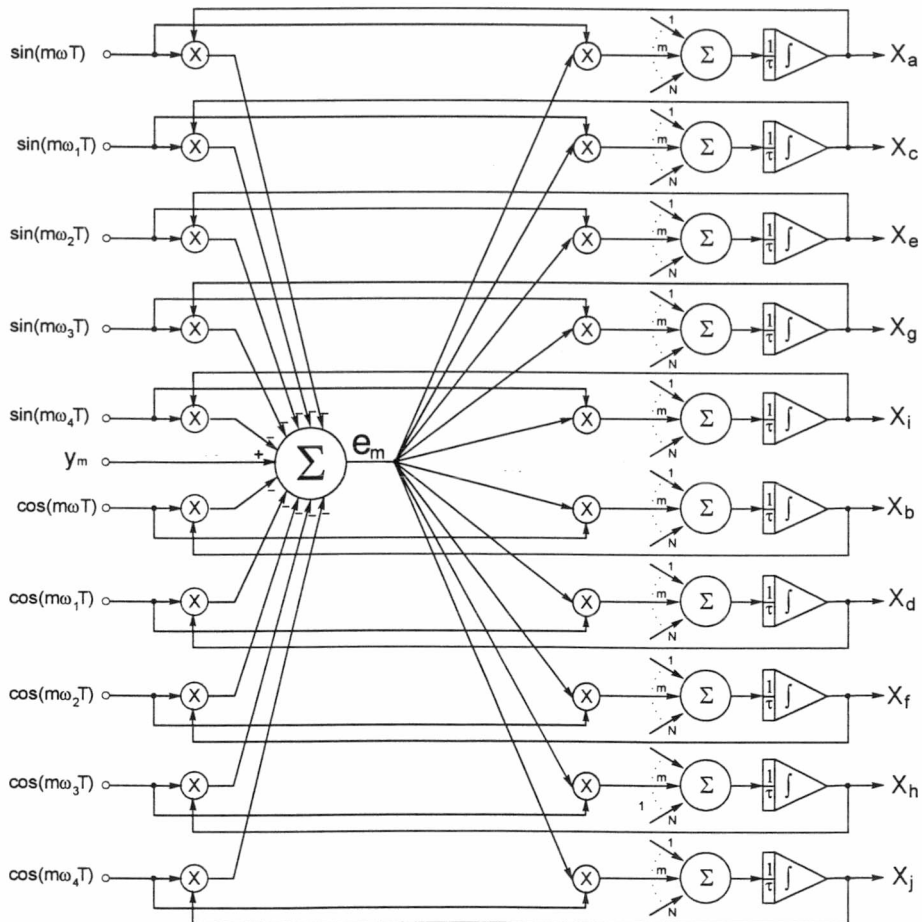
Przedstawione powyżej układy równań: (6.5) oraz (6.6), stanowią podstawę do skonstruowania odpowiednich sieci neuronowych, będących ich analogową implementacją.

Na rysunkach 6.2 oraz 6.3 przedstawione są fragmenty sztucznych, adaptacyjnych sieci neuronowych do wyznaczania wartości amplitud poszczególnych składowych badanego sygnału. Implementują one odpowiednio: układ równań (6.5) - sieć zgodna z rys. 6.2; układ równań (6.6) - rys. 6.3. W obu przypadkach sieci składają się z typowych elementów elektronicznych: sumatorów, układów mnożących, generatorów znaku, integratorów. Wartości funkcji dla różnych pulsacji mogą być przechowywane w pamięci. Wyznaczane wartości amplitud X oraz chwilowe wartości błędów e_m są mnożone przez te przechowywane w pamięci wartości funkcji *sinus* i *cosinus*. Tego rodzaju operacje mnożenia przez wartość stałą są łatwe w realizacji. W strukturach sieci możemy wyróżnić: 6 (rys. 6.2) lub 10 integratorów (rys. 6.3) oraz N równoległych kanałów obliczeniowych. Na ich wejścia dostarcza się próbki sygnału oraz estymowane w danej chwili przez układy wartości parametrów sygnału. Występuje tu również sprzężenie zwrotne, dlatego też sieci takie zalicza się do typu rekurencyjnego. Wszystkie N próbek sygnału doprowadza się do układu jednocześnie, co wymaga zastosowania odpowiedniego układu próbkująco-pamiętającego. Na wyjściu

każdego z kanałów pojawia się sygnał błędu e_m . Sieci wyznaczają amplitudy odpowiednio: $X_{a,...,f}$ - dla sieci z rys. 6.2 oraz $X_{a,...,j}$ - rys. 6.3.



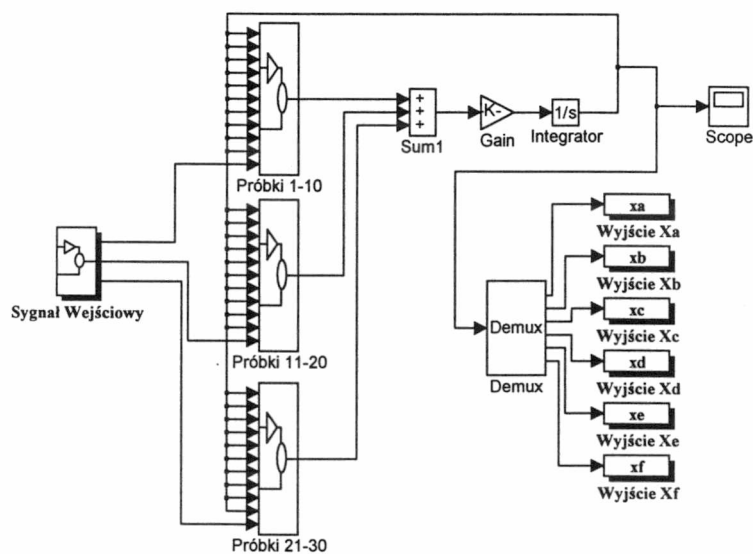
Rys. 6.2. Fragment sztucznej sieci neuronowej do detekcji stanu asynchronicznego, realizującej algorytm według układu równań (6.5)



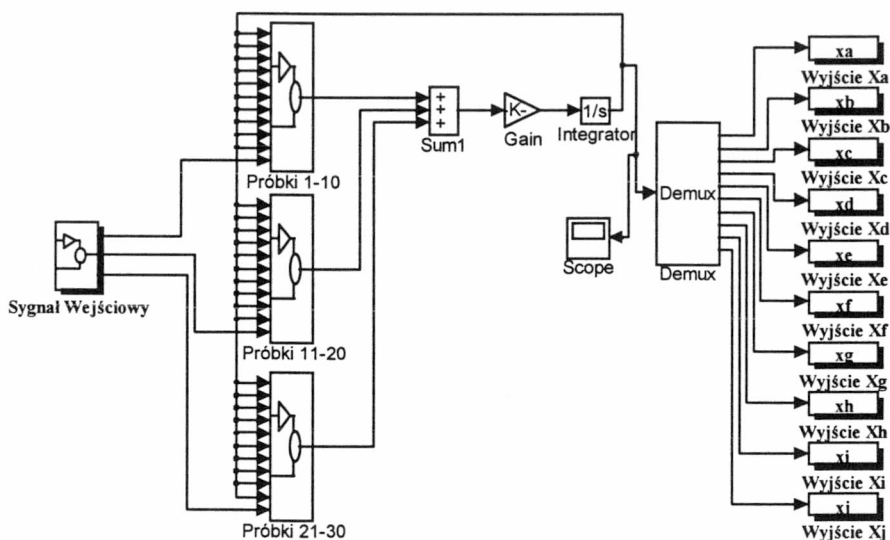
Rys. 6.3. Fragment sztucznej sieci neuronowej do detekcji stanu asynchronicznego, realizującej algorytm według układu równań (6.6)

6.5. Wyniki badań

Opisane i omówione w poprzednim rozdziale schematy sztucznych sieci neuronowych zostały zaimplementowane, a następnie przebadane na komputerze z wykorzystaniem programu SIMULINK. Skonstruowane w tym programie ich schematy blokowe, przedstawiono na rys. 6.4 oraz rys. 6.5.



Rys. 6.4. Implementacja układu (rys. 6.2) w programie SIMULINK



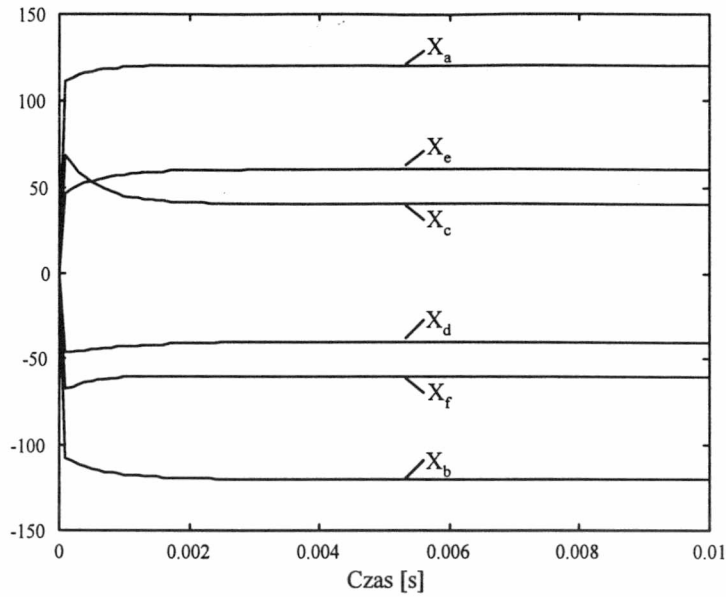
Rys. 6.5. Implementacja układu (rys. 6.3) w programie SIMULINK

Pierwszym etapem pracy był wybór optymalnej struktury sieci. W tym celu do analizy wybrano sygnał opisany równaniem (6.4), którego wartości amplitud wynosiły: $X_a = X_b = 100$, $X_c = X_d = 30$, $X_e = X_f = 50$, zaś wartość poślizgu dobierana była z przedziału $s = 0.01 \div 0.1$ [35]. Przebadano układy sieci o oknie próbkowania NT: 30, 60 i 90 ms i liczbie próbek N: 20, 30, 60. Najlepsze wyniki uzyskała sieć o oknie próbkowania 60 i 90 ms, liczbie próbek N = 30 oraz wartości poślizgu $s = \geq 0.03$. Będzie ona głównym narzędziem analizy sygnałów w dalszych etapach rozprawy. Przykładowe wyniki badań zawarto w tabeli 1, natomiast trajektorie estymowanych parametrów sygnału dla poślizgu o wartości $s = 0.05$ przedstawia rys. 6.6.

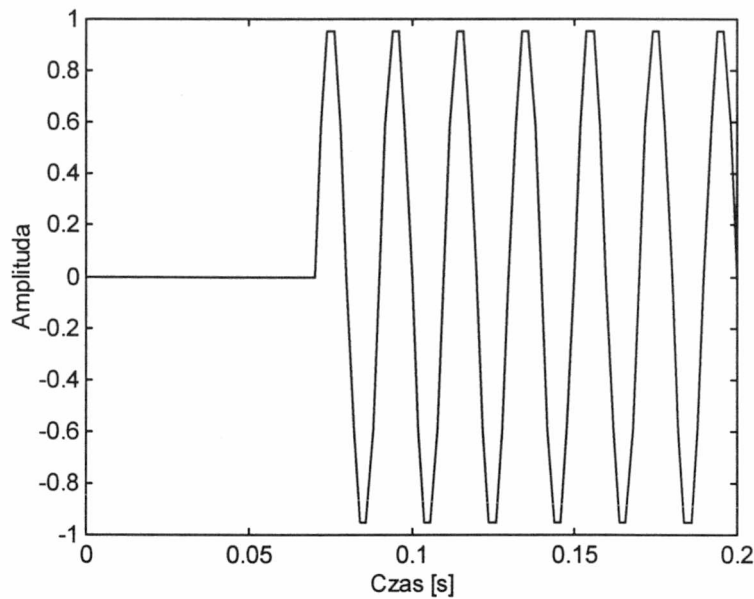
Tabela 1. Wartości wyznaczonych amplitud sygnału (6.4); liczba próbek N=30; okna próbkowania: 60 i 90 ms; t_c - czas konwergencji; X_s - założona wartość wyznaczanej amplitudy

NT [ms]	60		90				X_s
s	0.05	0.1	0.01	0.03	0.05	0.1	
X_a	92.13	98.33	88.74	100.79	99.54	99.51	100
X_b	105.05	100.71	85.89	99.74	100.13	100.01	100
X_c	38.65	30.53	35.07	28.30	30.30	30.33	30
X_d	15.33	27.51	40.71	31.63	30.05	29.91	30
X_e	49.38	50.98	55.97	50.40	49.78	49.96	50
X_f	58.93	51.32	54.17	49.97	50.36	50.44	50
t_c [μs]	-	14	400	300	60	6	-

Następnie przebadano sygnał otrzymany w wyniku symulacji komputerowej, którego fragment przedstawiono na rys. 6.7. Celem tej operacji było znalezienie właściwego punktu odniesienia do przeprowadzenia poprawnej analizy wyników estymacji parametrów sygnałów omówionych w dalszej części pracy. W szczególności chodziło o wyznaczenie stanu nieustalonego sieci, występującego w przypadku doprowadzenia na jej wejście próbek niezakłóconego sygnału sinusoidalnego. Do przeprowadzenia badań zastosowano sieć neuronową zgodnie z rys. 6.2 oraz 6.4. Do sieci dostarczono N = 30 próbek sygnału badanego, przy częstotliwości próbkowania $f_s = 500$ Hz, z których przy pierwszym obliczeniu tylko jedna próbka dotyczy sygnału badanego. Po zakończeniu obliczeń, przesuwano okno próbkowania (NT = 60 ms) o jeden okres próbkowania i całą procedurę powtarzano od nowa. Druga estymacja odbywała się zatem w sytuacji, gdy w oknie próbkowania znajdowały się dwie próbki badanego sygnału. Zastosowana procedura symuluje proces załączania sygnału.



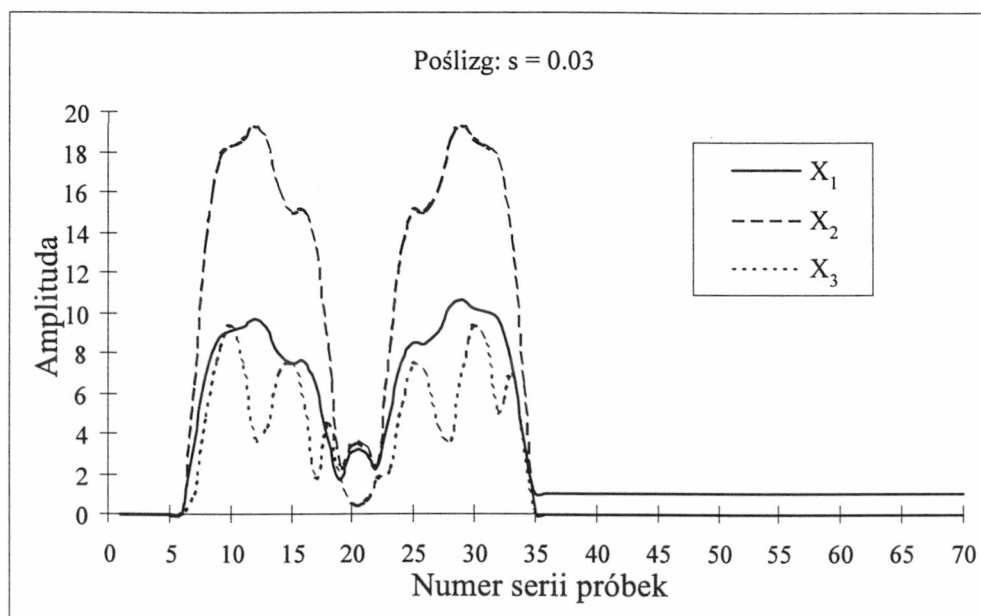
Rys. 6.6. Trajektorie wyznaczanych parametrów sygnału: $x(t) = 120 \sin \omega t - 120 \cos \omega t + 40 \sin(1-s)\omega t - 40 \cos(1-s)\omega t + 60 \sin(1-2s)\omega t - 60 \cos(1-2s)\omega t$, $N = 30$, $NT = 90 \text{ ms}$, $s = 0.05$



Rys. 6.7. Fragment przebiegu czasowego sygnału $y = \sin \omega t$ włączonego w chwili $t = 0.07 \text{ s}$

W każdym kroku omówionej wyżej procedury, sieć wyznaczała wartości poszczególnych parametrów składowych sygnału, zgodnie z algorytmem opisanym układem równań (6.5). Na ich podstawie, wykorzystując arkusz kalkulacyjny MS EXCEL (ver. 7.0a) [63], obliczono amplitudy składowych sygnału: $X_1 = \sqrt{X_a^2 + X_b^2}$, $X_2 = \sqrt{X_c^2 + X_d^2}$, $X_3 = \sqrt{X_e^2 + X_f^2}$ dla poszczególnych serii próbek (kroków

obliczeniowych). Przez pojęcie „seria próbek” należy rozumieć 30 próbek, które w następnej serii przesuwają się o jedną do przodu. Badania wykonano dla przyjętej w równaniu (6.4) wartości poślizgu s : 0.03. Przebiegi amplitud składowych sygnału w funkcji serii próbek przedstawia rysunek 6.8.



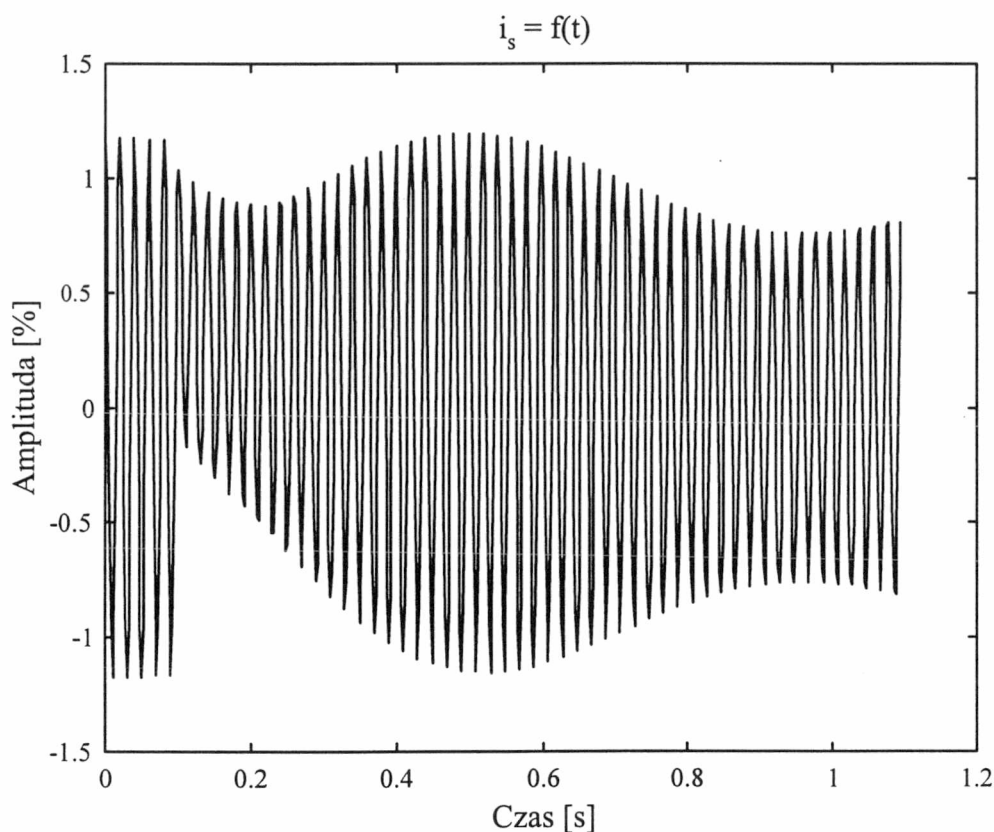
Rys. 6.8. Przebiegi wyznaczanych przez sieć amplitud składowych sygnału (rys. 6.7) dla założonej wartości poślizgu: 0.03;
 X_1 - amplituda składowej o pulsacji ω ;
 X_2 - amplituda składowej o pulsacji $(1-s)\omega$;
 X_3 - amplituda składowej o pulsacji $(1-2s)\omega$

Rysunek 6.8 przedstawia przebiegi amplitud poszczególnych składowych sygnału (rys. 6.7) wyznaczonych przez sieć dla założonej wartości poślizgu. Analizując wykresy można wyróżnić trzy zasadnicze etapy:

1. etap I - obejmujący zakres 0÷6 serii próbek co w odniesieniu do rys. 6.7, odpowiada sytuacji przed pojawieniem się sygnału na wejściu sieci;
2. etap II - obejmujący zakres 7÷36 serii próbek co odpowiada procesowi stopniowego nachodzenia okna próbkowania sieci NT na sygnał badany - obszar ten charakteryzuje się dużymi wartościami błędów estymacji amplitudy spowodowany pojawieniem się stanu nieustalonego wynikającego z sytuacji, że w kolejnych seriach próbek część z nich jest ze stanu poprzedniego a część ze stanu następnego;
3. etap III - obejmujący zakres powyżej 36 serii próbek odpowiada sytuacji gdy całe okno próbkowania obejmuje sygnał badany - obszar ten charakteryzuje się prawidłową estymacją amplitud składowych sygnału, które wynoszą odpowiednio: $X_1 = 1$, $X_2 = 0$, $X_3 = 0$.

Rozpatrując powyższe wyniki można wnioskować, że w celu stwierdzenia wystąpienia stanu asynchronicznego maszyny należy brać tę część uzyskanych wyników estymacji dla przypadków opisanych niżej, które odpowiadają stanowi ustalonymu procesowi obliczeniowemu.

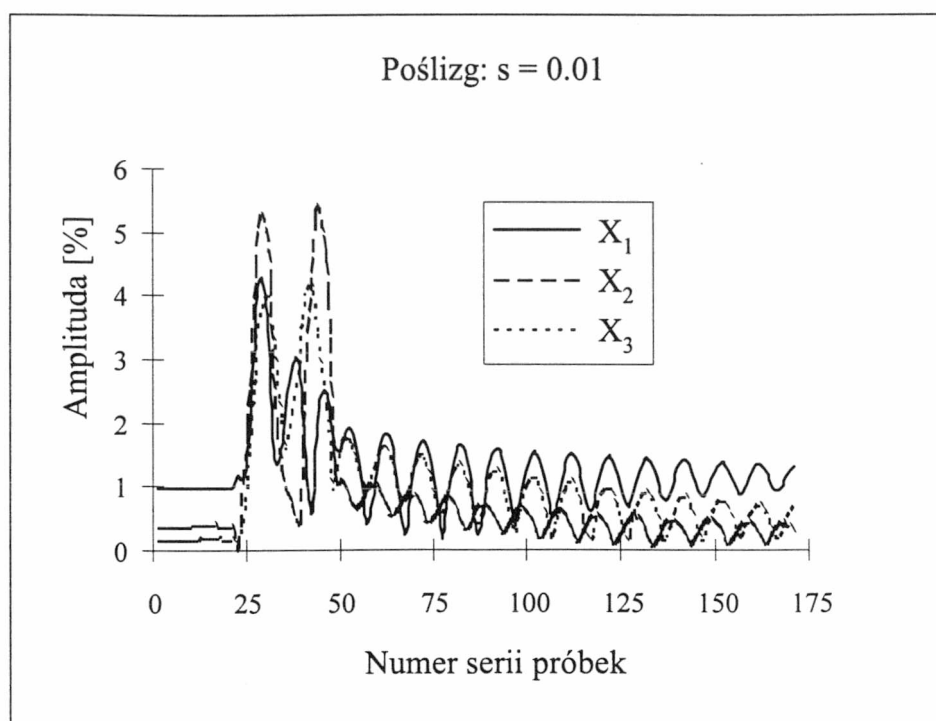
W kolejnym etapie badań wykorzystano wyniki prac teoretycznych, przeprowadzonych na modelu matematycznym hydrobloku elektrowni wodnej Żarnowiec-HG1 (podstawowe parametry: moc pozorna $S_n = 209 \text{ MV} \cdot \text{A}$; napięcie $U_n = 15.75 \text{ kV}$; prąd $I_n = 7.66 \text{ kA}$). Wzięto w szczególności pod uwagę te wyniki symulacji cyfrowej wybranych zakłóceń hydrogeneratora, które opisują stan jego pracy asynchronicznej [36, 37]. Poddano analizie zarejestrowany przebieg czasowy prądu stojana hydrogeneratora (w jednostkach względnych) podczas jego samowzbudzenia asynchronicznego, którego fragment przedstawiono na rys. 6.9. Badania wykonano dla przyjmowanych kolejno w równaniu (6.4) wartości poślizgu s : 0.01, 0.02 i 0.03. Przebiegi amplitud składowych sygnału w funkcji serii próbek dla różnych wartości poślizgu przedstawiono na rysunkach 6.10 i 6.11.



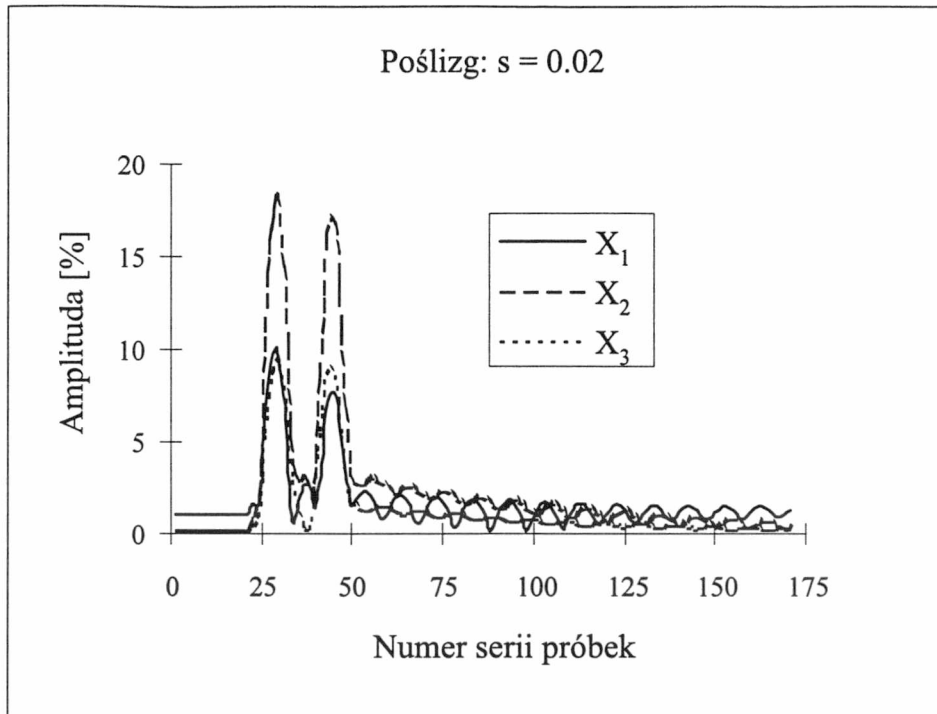
Rys. 6.9. Fragment przebiegu czasowego prądu uzwojenia stojana hydrogeneratora

Rysunki 6.10a÷6.10c przedstawiają przebiegi amplitud poszczególnych składowych prądu stojana wyznaczonych przez sieć dla założonych wartości poślizgów. Analizując wykresy można zauważyć, że dla każdej składowej prądu stojana występuje znaczny wzrost jej wartości amplitudy w zakresie 25 ÷ 50 serii próbek, co w porównaniu z rys. 6.9 odpowiada stanowi nieustalonemu. Biorąc pod uwagę wnioski wysunięte po analizie rys. 6.8, można stwierdzić, że przekroczenie 50-tej serii próbek odpowiada momentowi przejścia generatora z synchronicznego stanu pracy w stan asynchroniczny. Sytuacja ta wystąpiła dla każdej założonej dla sieci wartości poślizgu, co można zinterpretować osiągnięciem tych wartości przez wirnik maszyny oraz gwałtowną zmianą wartości prądu. Ponadto można przyjąć, że podczas trwania awarii wartości te zostały przekroczone. Na rys. 6.10 i 6.11 odpowiada to przekroczeniu 50-tej serii próbek, gdzie występuje i utrzymuje się wyraźny spadek wartości amplitudy składowych prądu.

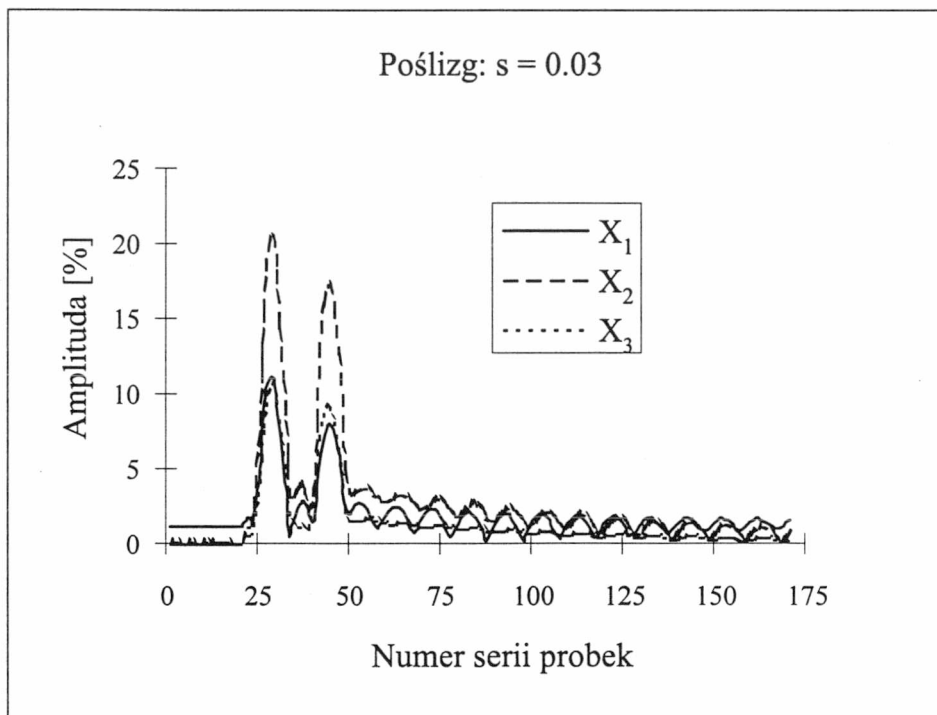
Na rysunku 6.11 zestawiono przebiegi estymowanych przez sieci amplitud składowych prądu stojana dla wszystkich przyjętych przy implementacji sieci według równania (6.4) wartości poślizgów. Pojawienie się składowych X_2 oraz X_3 oznacza stan asynchroniczny maszyny.



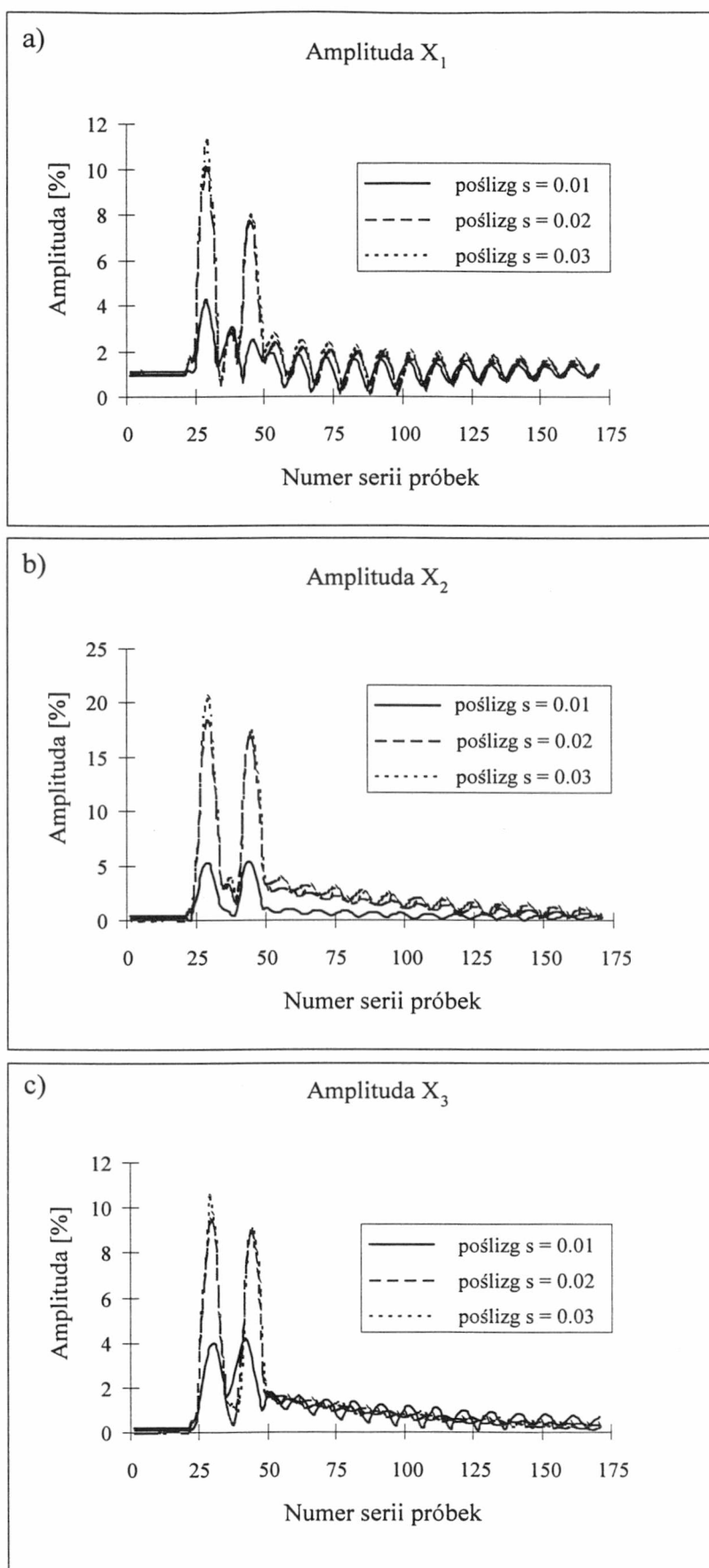
Rys. 6.10a. Przebiegi wyznaczanych przez sieci amplitud składowych prądu stojana (rys. 6.9) dla założonej wartości poślizgu: 0.01; amplitudy X_1 , X_2 , X_3 opisane zgodnie z rys. 6.8



Rys. 6.10b. Przebiegi wyznaczanych przez sieci amplitud składowych prądu stojana dla założonej wartości poślizgu: 0.02; amplitudy X_1 , X_2 , X_3 opisane zgodnie z rys. 6.8



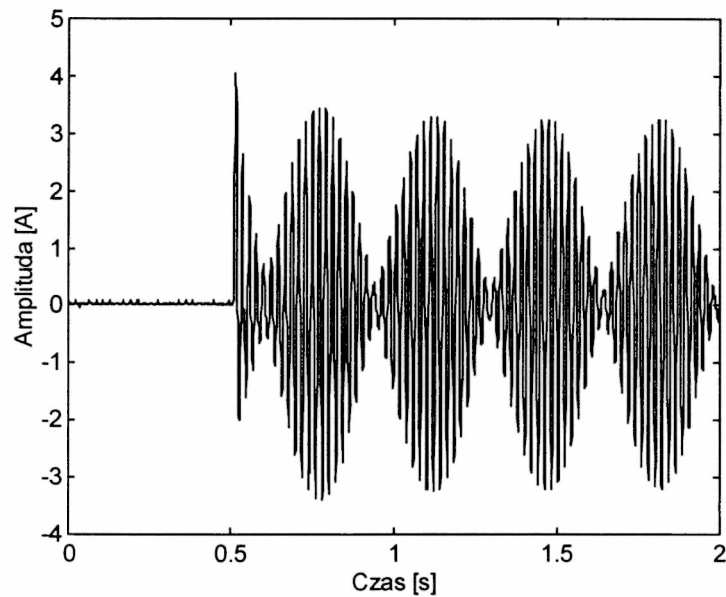
Rys. 6.10c. Przebiegi wyznaczanych przez sieci amplitud składowych prądu stojana dla założonej wartości poślizgu: 0.03; amplitudy X_1 , X_2 , X_3 opisane zgodnie z rys. 6.8



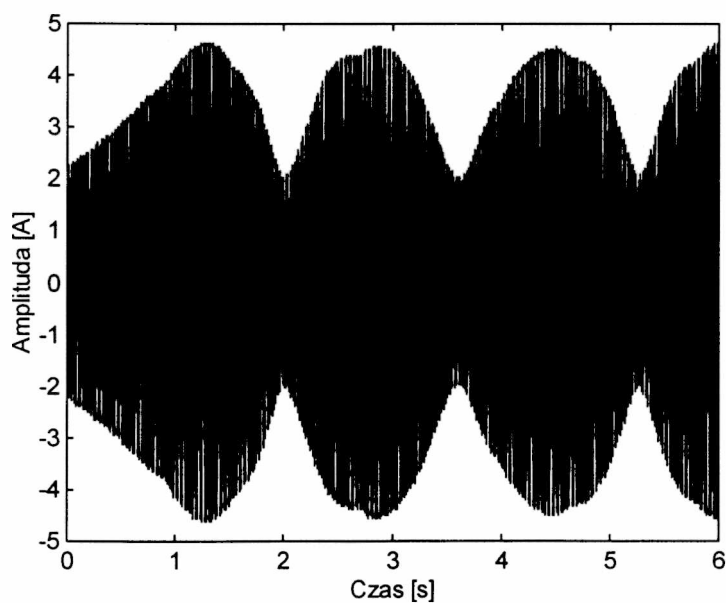
Rys. 6.11. Zestawienie przebiegów wartości amplitud składowych: a) X_1 , b) X_2 , c) X_3 , dla przyjętych w implementacji sieci wartości poślizgów: 0.01, 0.02, 0.03

W ostatnim etapie pracy, badaniom zostały poddane rzeczywiste przebiegi prądu stojana maszyny synchronicznej (moc modelu 1 kW co odpowiada wartości 1 MW mocy rzeczywistej), zarejestrowane podczas prób laboratoryjnych na Uniwersytecie w Saarbruecken (Niemcy). Badano 3 rodzaje stanów asynchronicznych:

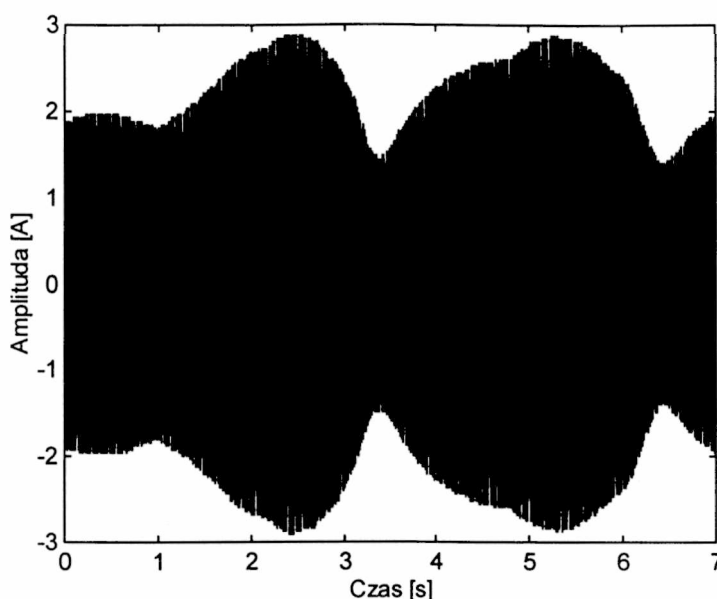
1. załączanie generatora (rys. 6.12),
2. praca generatora do granicy stabilności i odłączenie prądu wzbudzenia (przypadek I) (rys. 6.13),
3. praca generatora do granicy stabilności i odłączenie prądu wzbudzenia (przypadek II) (rys. 6.14).



Rys. 6.12. Załączanie generatora



Rys. 6.13. Praca generatora do granicy stabilności i odłączenie prądu wzbudzenia (przypadek I)

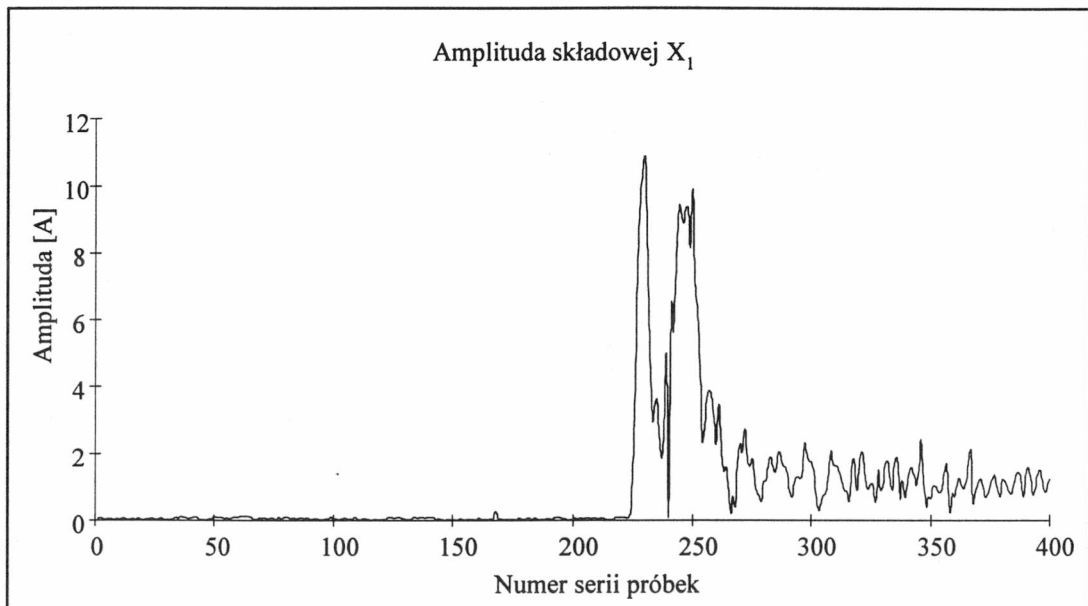


Rys. 6.14. Praca generatora do granicy stabilności i odłączenie prądu wzbudzenia (przypadek II)

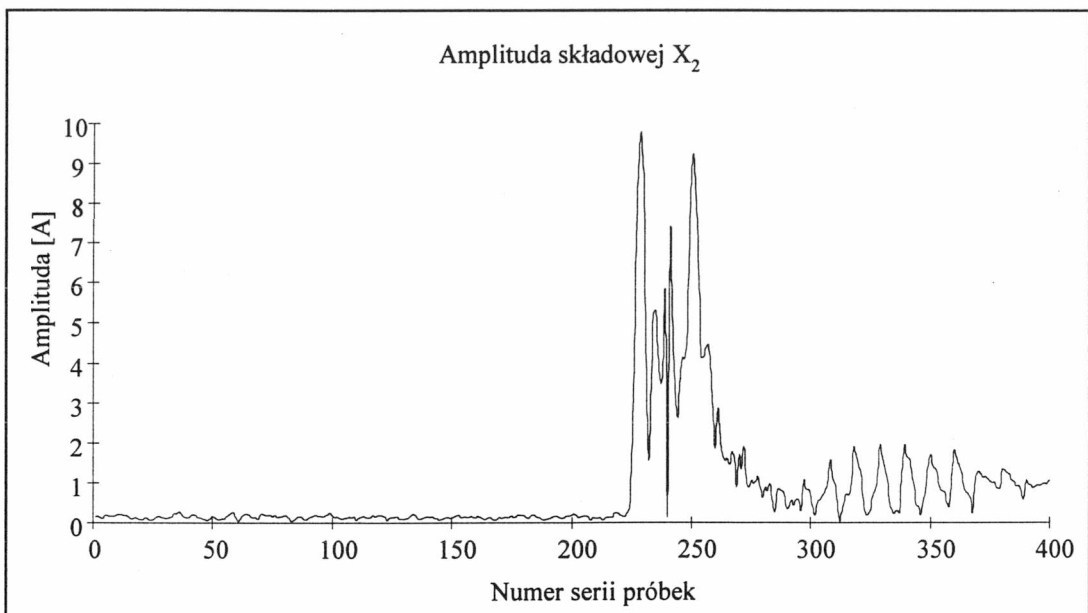
Do przeprowadzenia badań zastosowano sieć neuronową zgodnie z rys. 6.5. Podobnie jak to omówiono wyżej, do sieci dostarczono $N = 30$ próbek sygnału badanego, na podstawie których sieć wyznaczała parametry jego składowych. Po zakończeniu jednego cyklu obliczeń, przesuwano okno próbkowania sieci o jedną próbkę i całą procedurę powtarzano od nowa, aż do momentu wyczerpania się próbek sygnału badanego. W każdym kroku omówionej wyżej procedury, sieć wyznaczała wartości poszczególnych parametrów składowych sygnału, zgodnie z algorytmem opisanym równaniem (6.6). Na ich podstawie, wykorzystując arkusz kalkulacyjny MS EXCEL, obliczono amplitudy składowych sygnału: $X_1 = \sqrt{X_a^2 + X_b^2}$, $X_2 = \sqrt{X_c^2 + X_d^2}$, $X_3 = \sqrt{X_e^2 + X_f^2}$, $X_4 = \sqrt{X_g^2 + X_h^2}$, $X_5 = \sqrt{X_i^2 + X_j^2}$. Badania przeprowadzono dla przyjętej wartości poślizgu s : 0.03. Na kolejnych stronach rozprawy przedstawione są wykresy amplitud poszczególnych składowych dla analizowanych przebiegów.

Rysunki 6.15a÷6.15e ilustrują wyniki dotyczące przebiegu z rys. 6.12 (załączanie generatora), rysunki: 6.16a÷6.16e wyniki dotyczące przebiegu z rys. 6.13 (praca generatora do granicy stabilności i odłączenie prądu wzbudzenia - przypadek I), natomiast rysunki: 6.17a÷6.17e wyniki dotyczące przebiegu z rys. 6.14 (praca generatora do granicy stabilności i odłączenie prądu wzbudzenia - przypadek II). We

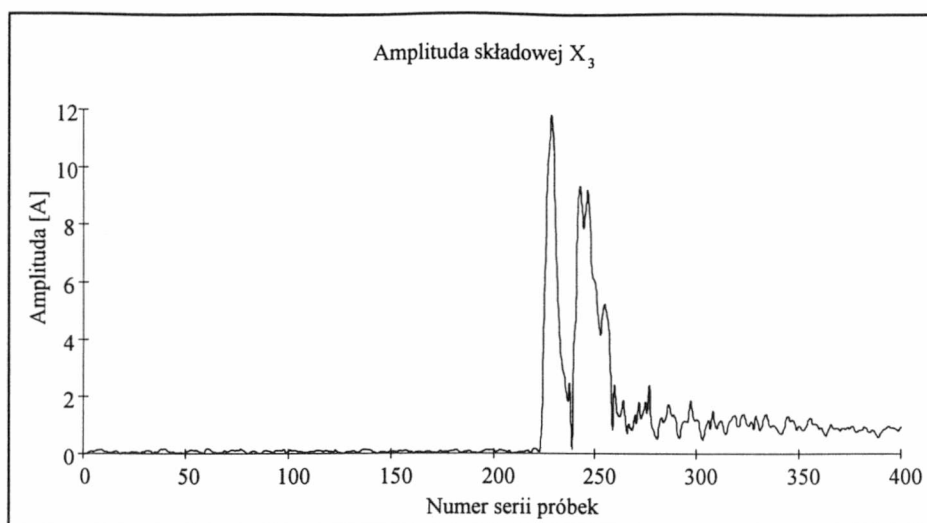
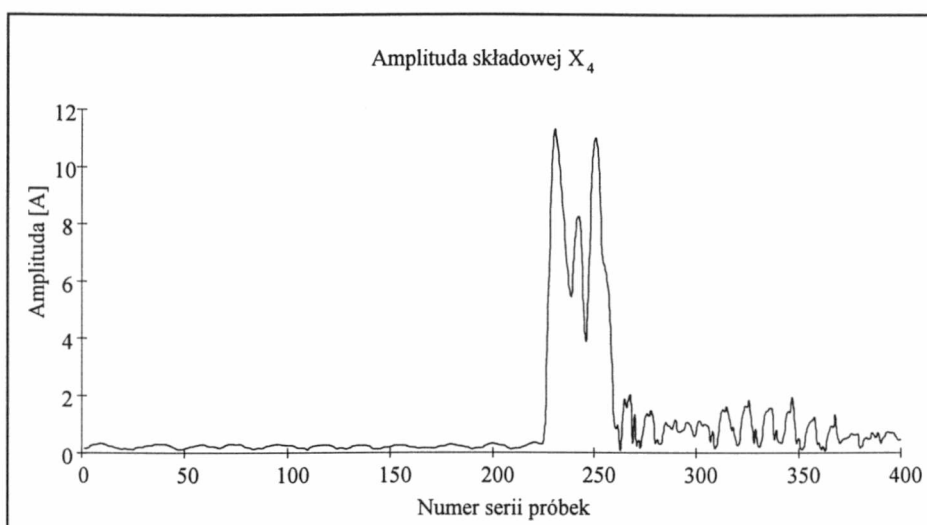
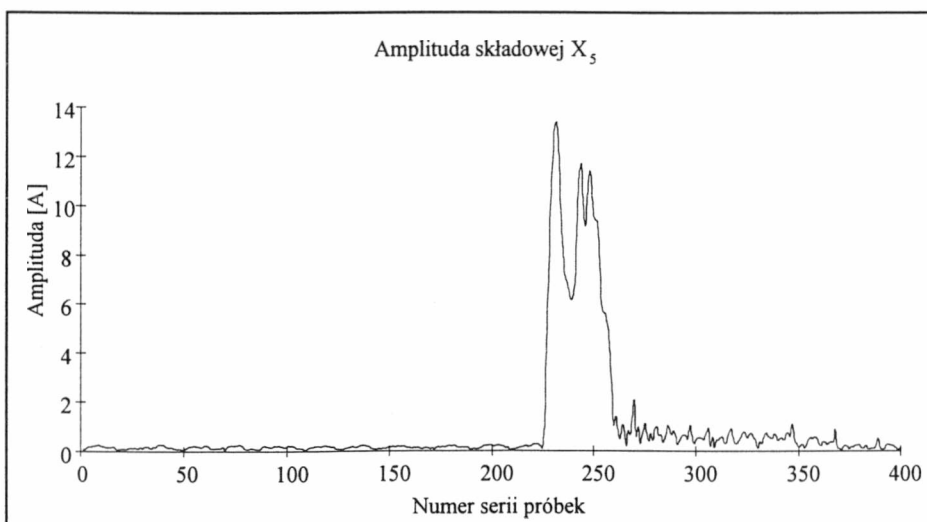
wszystkich przypadkach stanu asynchronicznego obserwuje się pojawienie się dodatkowych składowych prądu stojana. Obecność tych składowych jest wykrywana przez sieć neuronową, co może być wykorzystane do identyfikacji stanu asynchronicznego.

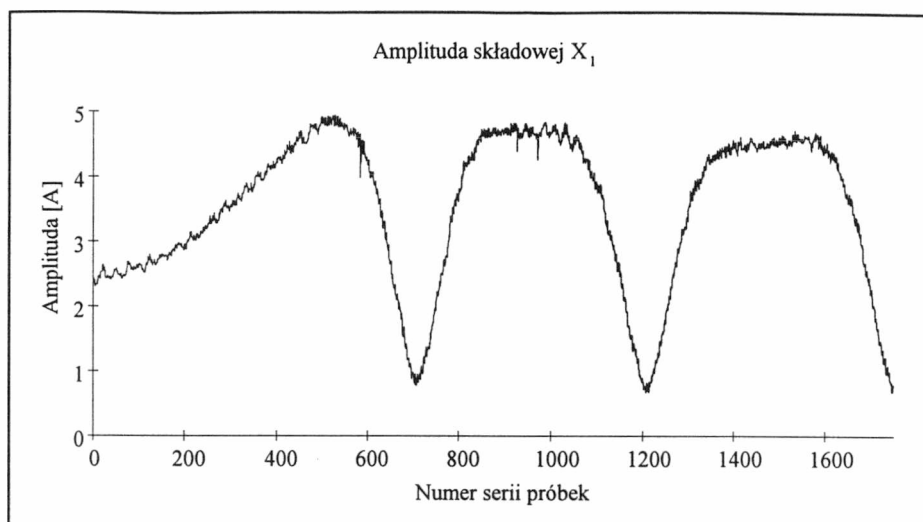
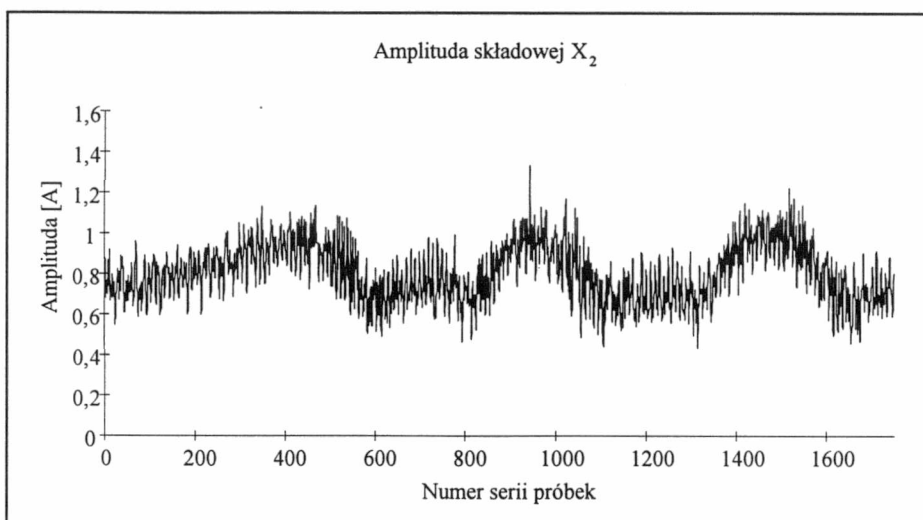
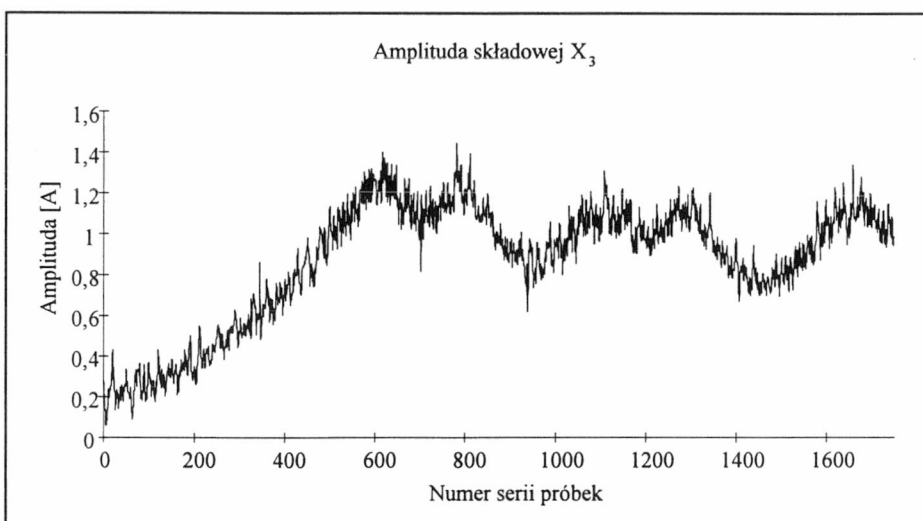


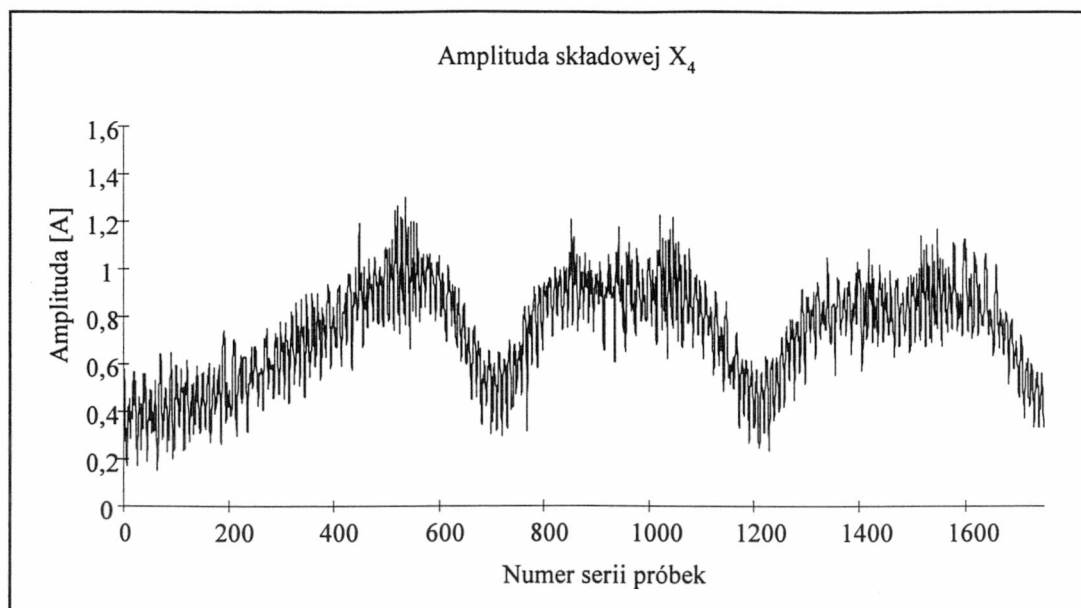
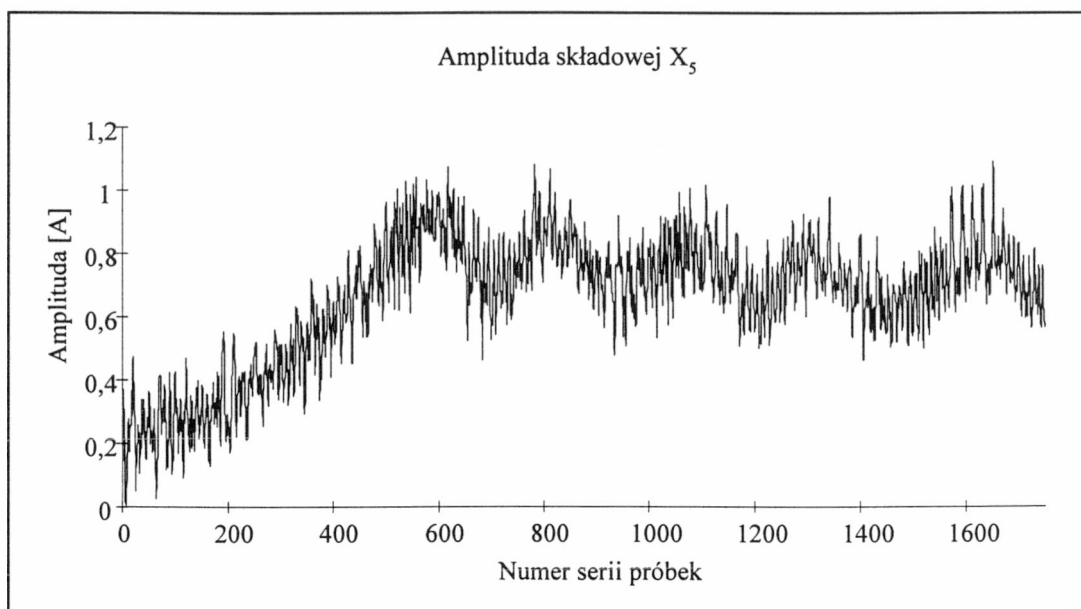
Rys. 6.15a. Przebieg amplitudy składowej X_1 dla przebiegu z rys. 6.12

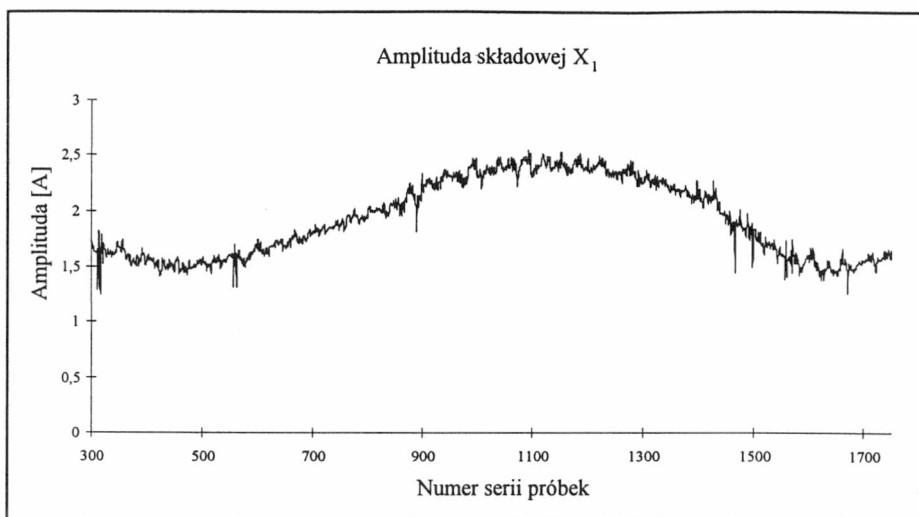
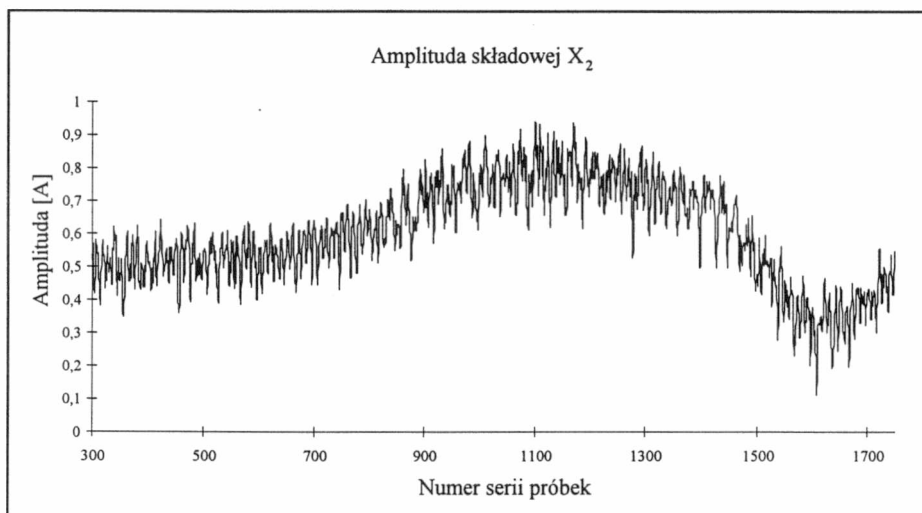
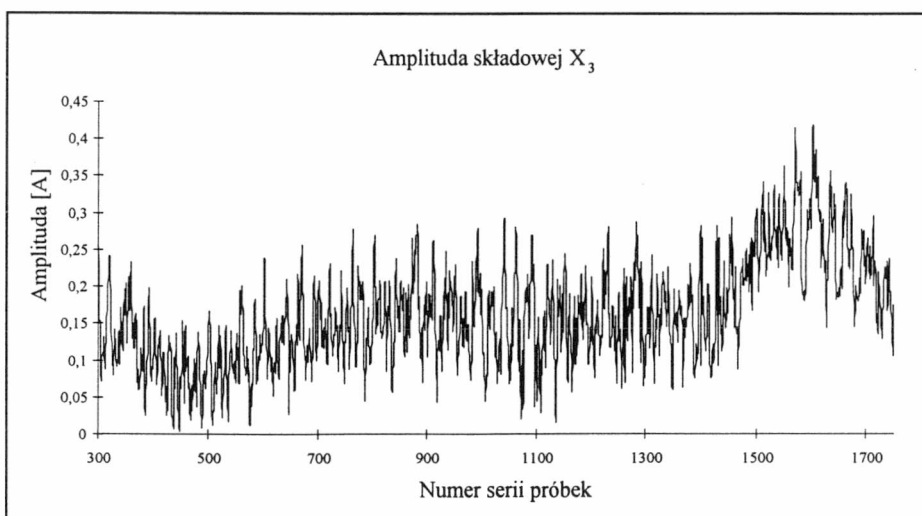


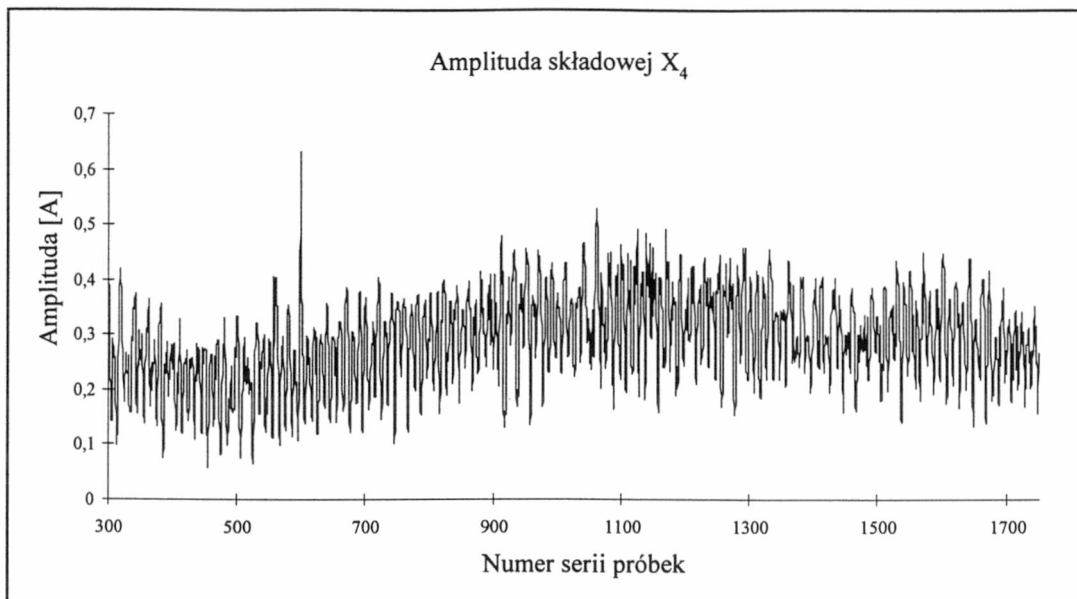
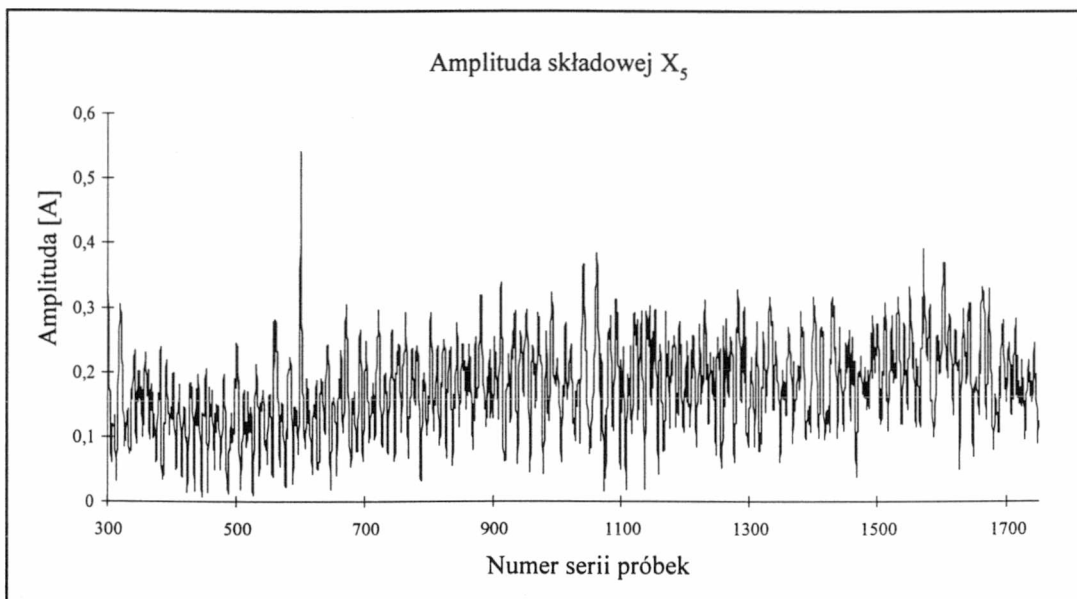
Rys. 6.15b. Przebieg amplitudy składowej X_2 dla przebiegu z rys. 6.12

Rys. 6.15c. Przebieg amplitudy składowej X_3 dla przebiegu z rys. 6.12Rys. 6.15d. Przebieg amplitudy składowej X_4 dla przebiegu z rys. 6.12Rys. 6.15e. Przebieg amplitudy składowej X_5 dla przebiegu z rys. 6.12

Rys. 6.16a. Przebieg amplitudy składowej X_1 dla przebiegu z rys. 6.13Rys. 6.16b. Przebieg amplitudy składowej X_2 dla przebiegu z rys. 6.13Rys. 6.16c. Przebieg amplitudy składowej X_3 dla przebiegu z rys. 6.13

Rys. 6.16d. Przebieg amplitudy składowej X_4 dla przebiegu z rys. 6.13Rys. 6.16e. Przebieg amplitudy składowej X_5 dla przebiegu z rys. 6.13

Rys. 6.17a. Przebieg amplitudy składowej X_1 dla przebiegu z rys. 6.14Rys. 6.17b. Przebieg amplitudy składowej X_2 dla przebiegu z rys. 6.14Rys. 6.17c. Przebieg amplitudy składowej X_3 dla przebiegu z rys. 6.14

Rys. 6.17d. Przebieg amplitudy składowej X_4 dla przebiegu z rys. 6.14Rys. 6.17e. Przebieg amplitudy składowej X_5 dla przebiegu z rys. 6.14

7. Wykrywanie zwarć łukowych w przesyłowych liniach elektroenergetycznych

7.1. Wprowadzenie

Wystąpieniu zwarcia w układzie elektroenergetycznym towarzyszy zmiana rozprywu prądów, które w określonych punktach układu mogą osiągać niebezpieczne dla urządzeń elektrycznych wartości. Ponadto występuje zmiana rozkładu napięć, wskutek czego w niektórych punktach układu mogą wystąpić nadmierne naprężenia dielektryczne, w innych zaś, napięcie może się tak zmniejszyć, że będzie wpływało zakłócająco na pracę urządzeń lub maszyn elektrycznych.

Jedno- i trójfazowe samoczynne ponowne załączanie (SPZ) linii elektroenergetycznej, po spowodowanym zwarcie wyłączeniu, jest bardzo skutecznym sposobem poprawy pewności pracy systemu elektroenergetycznego. Jednofazowe SPZ jest stosowane w liniach najwyższych napięć, natomiast trójfazowe w sieciach rozdzielczych. Po wystąpieniu zwarcia w linii, ta linia lub jedna faza tej linii jest wyłączana, a następnie, po czasie 0.5-1.5s ponownie załączana. Większość zwarć w liniach napowietrznych to zwarcia łukowe. Po wyłączeniu takiego zwarcia następuje dejonizacja strefy połukowej i ponowne załączenie linii umożliwia jej dalszą, normalną pracę. Ponowne załączenie na zwarcie trwałe może być jednak niebezpieczne dla niektórych urządzeń elektroenergetycznych. Aby zapobiec ponownemu załączeniu na zwarcie trwałe opracowano metody umożliwiające rozpoznawanie zwarć łukowych [16, 17].

T. Dalstein oraz B. Kulicke [16] sugerują osobną dla każdej z trzech faz wielowarstwową jednokierunkową sieć neuronową, która klasyfikuje zwarcia trwałe i łukowe, przy założeniu, że napięcie łuku jest zbliżone bardziej lub mniej do kształtu prostokątnego. Sieć była uczona na zbiorze około 3000 modeli zakłóceń łukowych oraz innych rodzajów awarii.

M. B. Djuric [17, 18] proponuje metodę, która bazuje na cyfrowym przetwarzaniu napięć i prądów na zaciskach zakłóconej linii. Metoda ta rozpoznaje napięciowy prostokątny impuls na zaciskach linii na podstawie analizy sinusowych składowych nieparzystych harmonicznych napięcia, które są w fazie z sinusowymi nieparzystymi

harmonicznymi prostokątnego impulsu. Wykazano, że pod uwagę mogą być brane tylko 3-e, 5-e, 7-e oraz 9-e harmoniczne napięcia analizowanego sygnału.

Stosowanie wyżej wymienionych metod jest niestety ograniczone ze względu na ich duży stopień skomplikowania. Abstrahując nawet od złożoności tych metod, założenie prostokątnego kształtu przebiegu napięcia łuku jest założeniem bardzo uproszczonym. W niniejszej pracy proponuje się nową metodę rozróżniania zwarć łukowych oraz trwałych [32].

Podczas zwarcia łukowego napięcie na początku linii jest sumą spadku napięcia na impedancjach linii oraz napięcia łuku. Napięcie łuku jest zniekształcone z powodu jego nieliniowej natury. Łuk ma bardzo mały wpływ na kształt przebiegu prądu, ponieważ wartość napięcia łuku jest o wiele mniejsza w stosunku do wartości napięcia danej fazy. Można zatem założyć sinusoidalny kształt przebiegu prądu.

W celu rozpoznania zwarcia łukowego proponuje się użyć stopień zniekształcenia krzywej napięcia na początku linii oraz rodzaj zmian tego zniekształcenia w miarę upływu czasu palenia się łuku. Rośnie przy tym spadek napięcia na łuku oraz stopień zniekształcenia napięcia na początku linii. W celu wyznaczenia stopnia zniekształcenia napięcia, należy najpierw wyznaczyć parametry podstawowej harmonicznej tego napięcia. Do wyznaczania tych parametrów zastosowano rekurencyjną sztuczną sieć neuronową. Oprócz estymowanych parametrów, z sieci uzyskać można błędy estymacji, które są dobrą informacją o stopniu zniekształcenia krzywej badanego napięcia. Pod wpływem ruchów powietrza łuk ma tendencję do wydłużania się. Wydłużanie się łuku wpływa na wzrost wartości błędów estymacji. Wzrost ten proponuje się zastosować jako kryterium dla rozróżniania zwarć łukowych od zwarć trwałych. Po otwarciu zwartej fazy obserwuje się wystąpienie łuku wtórnego [29, 49]. Z upływem czasu łuk ten wydłuża się, a jego długość może osiągnąć wartość nawet dziesięciokrotnie większą od swej wartości początkowej. Tak więc wzrost wartości błędów estymacji składowej podstawowej napięcia łuku wtórnego może być również użyty do wykrywania zwarcia łukowego.

7.2. Prezentacja metody

Podczas wystąpienia zwarcia łukowego w linii, początkowa wartość napięcia łuku może być ustalona z faktu, że gradient napięcia pozostaje stały i niezależny od prądu.

Wartość gradientu dla łuku elektrycznego w powietrzu wynosi 1.0 - 2.5 kV/m [15]. Minimalna długość łuku elektrycznego dla zwarć jednofazowych może być wyznaczona z długości łańcucha izolatorów.

Podczas trwania zakłócenia długość łuku w powietrzu ulega zmianie [29, 49, 30].

Wydłużanie spowodowane jest głównie przez:

- siły magnetyczne zależne od prądu zakłócenieniowego (zwarciovowego), lecz także:
- konwekcję plazmy i otaczającego powietrza, oraz
- wiatr.

Wydłużanie się czasu trwania zakłócenia wpływa na wzrost stopnia zniekształcenia krzywej napięcia, który proponuje się zastosować do wykrywania zwarcia łukowego. Wzrost zniekształcenia krzywej napięcia podczas trwania zakłócenia oznacza zwarcie łukowe. Jeśli łuk pierwotny nie wydłuża się, wówczas można wykorzystać łuk wtórny. Zwarcia łukowe nie zawsze muszą mieć charakter przemijający. Jeśli podczas trwania zwarcia łuk się nie wydłuża, to istnieje mała szansa pomyślnego ponownego załączenia linii.

W celu wyznaczenia stopnia zniekształcenia napięcia, należy najpierw wyznaczyć parametry podstawowej harmonicznej tego napięcia. W literaturze opisano różne cyfrowe metody wyznaczania w czasie rzeczywistym parametrów podstawowej składowej zniekształconych sygnałów [38, 19]. W ostatnich latach wzrasta nieustannie zainteresowanie metodami sztucznych sieci neuronowych. Do wyznaczania parametrów podstawowej składowej sygnału można zastosować rekurencyjne, adaptacyjne sieci neuronowe [9]. Zagadnienie formułuje się jako problem optymalizacyjny, który może być rozwiązany przez zastosowanie metody największego spadku.

Najprostsza sieć z dwoma neuronami została omówiona w rozdziale 7.4 (rys. 7.1). Sieć wyznacza nie tylko dwa parametry X_a i X_b , lecz także błędy estymacji e . Dla każdej wartości próbki badanego sygnału dostarczanej na wejście sieci, otrzymuje się odpowiedni błąd estymacji e_m . Suma kwadratów błędów estymacji w czasie jednego okresu jest miarą stopnia zniekształcenia krzywej napięcia i wyraża się wzorem:

$$D^2 = \sum_{m=1}^N \{e_m / U_n\}^2, \quad (7.1)$$

gdzie: U_n - napięcie znamionowe.

Po zakończeniu procedury wyznaczania stopnia zniekształcenia D^2 dla pierwszego okresu przebiegu napięcia, przesuwno okno próbkowania sieci NT o jedną połowę tego okresu, i całą procedurę powtarzano od nowa.

7.3. Algorytm działania sieci

Podstawową składową napięcia w systemie elektroenergetycznym można przedstawić w następujący sposób:

$$x(t) = X_a \sin(\omega T) + X_b \cos(\omega T), \quad (7.2)$$

gdzie:

- X_a, X_b - nieznane parametry podstawowej składowej sygnału,
- ω - znana pulsacja składowej podstawowej.

W rzeczywistości sygnał mierzony obarczony jest różnego rodzaju błędami, tak więc można go zapisać jako:

$$y(t) = x(t) + e(t), \quad (7.3)$$

gdzie:

$e(t)$ - jest to nieznaną błąd zawierający zakłócenia.

Z praktycznego punktu widzenia założono (rozdział 6.3), że sygnał $y(t)$ jest mierzony w skończonym przedziale czasu, w wyniku czego otrzymano N próbek mierzonego napięcia y_m :

$$y_m = X_a \sin(m\omega T) + X_b \cos(m\omega T) + e_m.$$

Podstawową czynnością w tym przypadku, jest sformułowanie odpowiedniej funkcji energetycznej $E(\mathbf{X})$, której minimum odpowiada szukanemu rozwiązaniu zadania optymalizacyjnego. Zagadnienie estymacji może być sformułowane w następujący sposób:

wyznaczyć wektor $\mathbf{X} = [X_a, X_b]$, który minimalizuje odpowiednio dobraną funkcję energetyczną $E(\mathbf{X})$.

W celu uproszczenia dalszych rozważań, przyjmuje się najczęściej stosowane kryterium najmniejszego błędu kwadratowego (3.13):

$$E(\mathbf{X}) = \frac{1}{2} \sum_{m=1}^N e_m^2.$$

Minimum tej funkcji otrzymuje się stosując algorytm optymalizacyjny największego spadku (3.9):

$$\frac{d\mathbf{X}}{dt} = -\frac{1}{\tau} \nabla E(\mathbf{X}), \quad \tau > 0, \quad (7.4)$$

gdzie:

τ - stała czasowa całkowania,

$$\nabla E(\mathbf{X}) = \left[\frac{\partial E(\mathbf{X})}{\partial X_a}, \frac{\partial E(\mathbf{X})}{\partial X_b} \right].$$

Równanie (7.4) może być zapisane w postaci układu równań różniczkowych:

$$\frac{dX_a}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \sin(m\omega T), \quad (7.5)$$

$$\frac{dX_b}{dt} = \frac{1}{\tau} \sum_{m=1}^N e_m \cos(m\omega T), \quad (7.6)$$

gdzie:

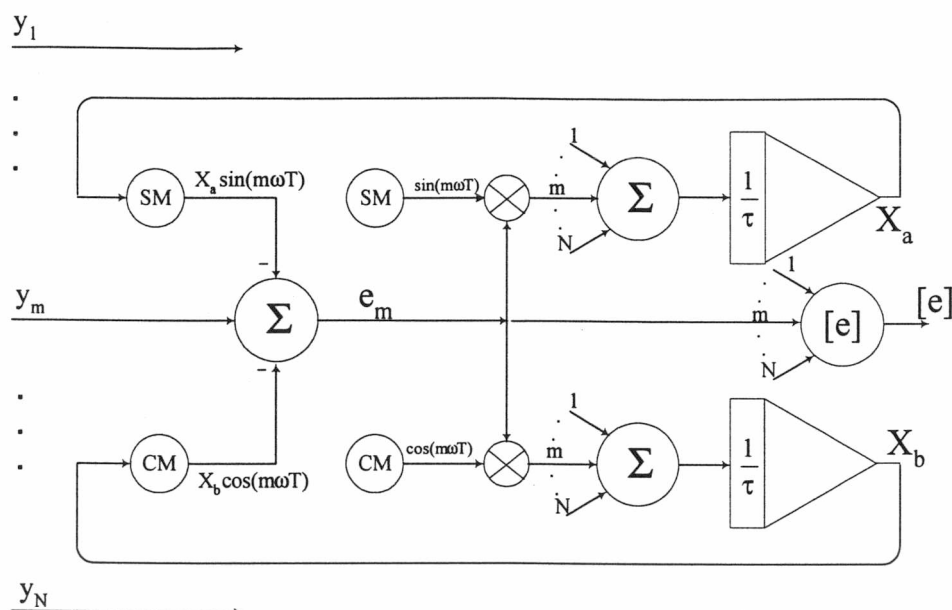
$$e_m = y_m - X_a \sin(m\omega T) - X_b \cos(m\omega T).$$

Rozpoznanie zwarcia łukowego odbywa się na podstawie analizy błędu estymacji e_m , który jest dobrym miernikiem stopnia zniekształcenia krzywej napięcia.

7.4. Implementacja algorytmu

Przedstawione powyżej równania: (7.5) oraz (7.6), stanowią podstawę do skonstruowania odpowiedniej sieci neuronowej (rys.7.1).

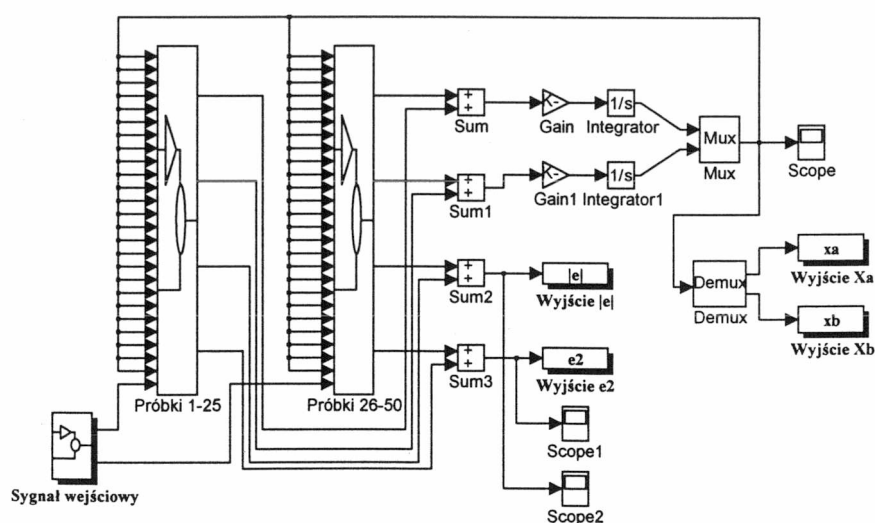
Na rysunku 7.1 przedstawiony jest fragment sztucznej, adaptacyjnej sieci neuronowej do wyznaczania wartości amplitud poszczególnych składowych badanego sygnału, będącego modyfikacją sieci z rys. 5.1. Opis i zasada działania tej sieci zawarta jest w rozdziale 5.5. Istotną różnicą w strukturze sieci, w porównaniu do sieci z rys. 5.1, jest wyprowadzenie dodatkowego kanału dla błędów estymacji, które następnie są poddawane odpowiedniej analizie.



Rys. 7.1. Fragment adaptacyjnej sieci neuronowej do wyznaczania parametrów podstawowej składowej zniekształconego sygnału napięcia

7.5. Wyniki badań

Opisany i omówiony w poprzednim rozdziale schemat sieci neuronowej został zaimplementowany, a następnie przebadany na komputerze z wykorzystaniem programu SIMULINK. Skonstruowany w tym programie schematy blokowy sieci przedstawiono na rys. 7.2.



Rys. 7.2. Implementacja układu sieci (rys. 7.1) w programie SIMULINK

W pierwszym etapie pracy przeprowadzono obszerne badania symulacyjne metody modelując zwarcie łukowe w linii 110 kV. Napięcie linii próbkowano z częstotliwością 2500 Hz. Do układu (rys. 7.1) doprowadzono $N = 25$ próbek napięcia. Rozpoznanie zwarcia łukowego odbywa się na podstawie analizy błędu estymacji e_m . Wyznaczano parametry podstawowej składowej napięcia w oparciu o próbki tego napięcia z przedziału równego połowie okresu podstawowej składowej. W miarę upływu czasu doprowadzano do układu (rys. 7.1) próbki z kolejnych połówek okresu. Wydłużanie się łuku powoduje, że wzrasta stopień zniekształcenia krzywej napięcia i rosną błędy estymacji e_m . Każdej próbce napięcia odpowiada jedna wartość tego błędu. Jako miernik stopnia zniekształcenia przyjęto, na tym etapie badań, sumę bezwzględnych wartości tych błędów z przedziału równego okresowi próbkowania (pół okresu). Niezależnie od położenia miejsca zwarcia w linii, suma 25 błędów estymacji $|e_m|$, odniesiona do napięcia fazowego linii wynosiła:

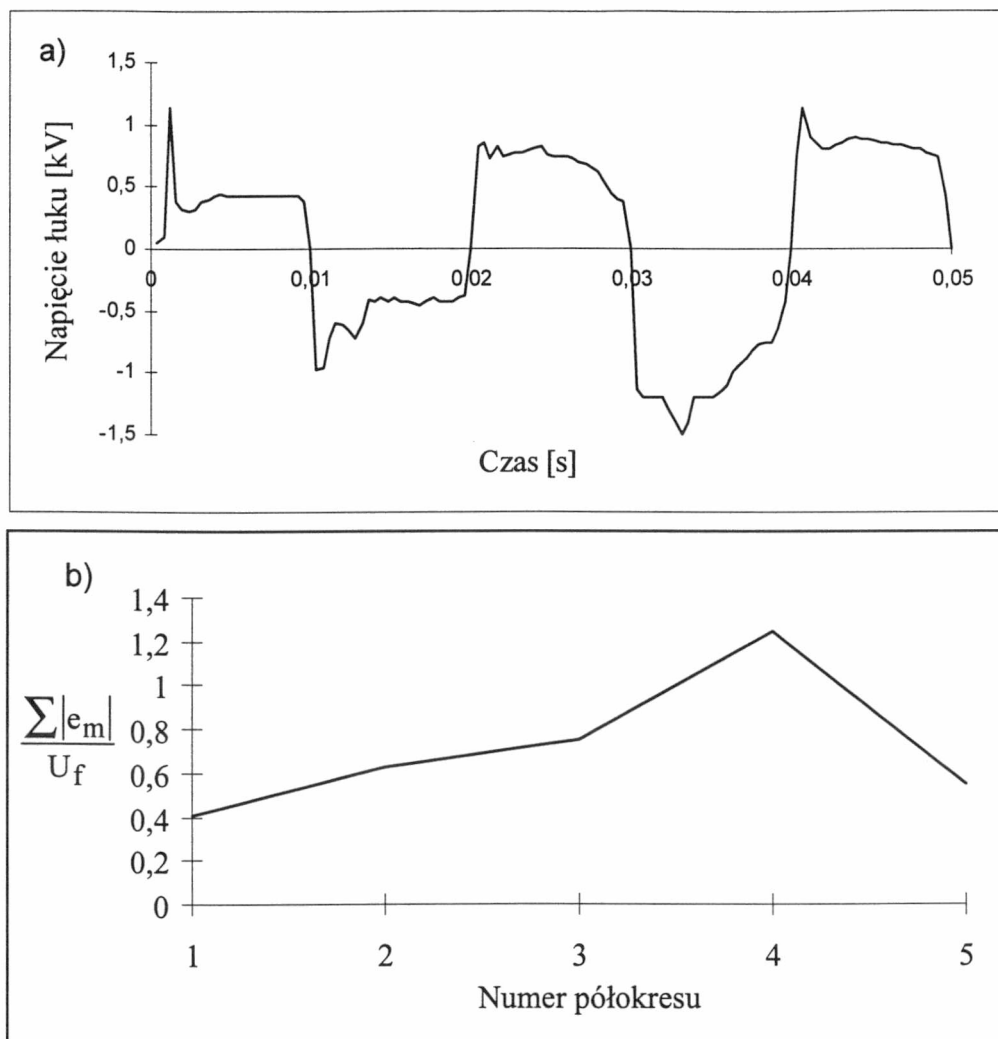
$$\frac{\sum_{m=1}^{25} |e_m|}{U_f} = 0.88 = \text{const.}$$

W tych wstępnych badaniach nie uwzględniano wydłużania się łuku.

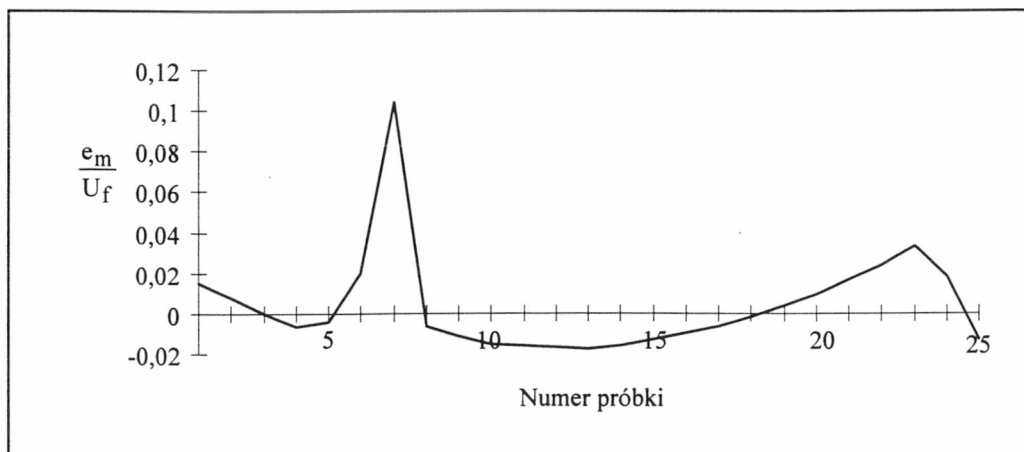
W dalszych badaniach wykorzystano zarejestrowane przebiegi napięć podczas zwarć łukowych w rzeczywistej linii 20 kV. Rysunki 7.3a i 7.5a przedstawiają przebiegi napięć łuku zwarciovego, podczas zwarć w odległości 2 km oraz 10 km od początku linii 20 kV. Na rysunkach 7.3b oraz 7.5b pokazano wyznaczone sumy błędów estymacji dla kolejnych półokresów przebiegu napięcia łuku. Obserwuje się wzrost wartości błędów, spowodowany wydłużaniem się łuku i wzrostem napięcia łuku. Zmniejszenie wartości sumy błędów w 5-tym półokresie (rys. 7.3b) spowodowane jest zgaśnięciem łuku i następnie ponownym jego zapłonem. Rysunki 7.4 i 7.6 przedstawiają przebiegi chwilowych wartości błędów estymacji e_m .

W kolejnym etapie badań wykorzystano wyniki testów łuku zwarciovego wykonane przez FGH (Forschungsgemeinschaft Hochspannungstechnik) w Mannheim (Niemcy). Przykładowo rys. 7.7 przedstawia przebiegi mierzonego napięcia łuku wzdłuż szeregu izolatorów na napięciu 20 kV, 110 kV oraz 220 kV. Przeprowadzono badania na rzeczywistych sygnałach napięciowych zarejestrowanych podczas testów

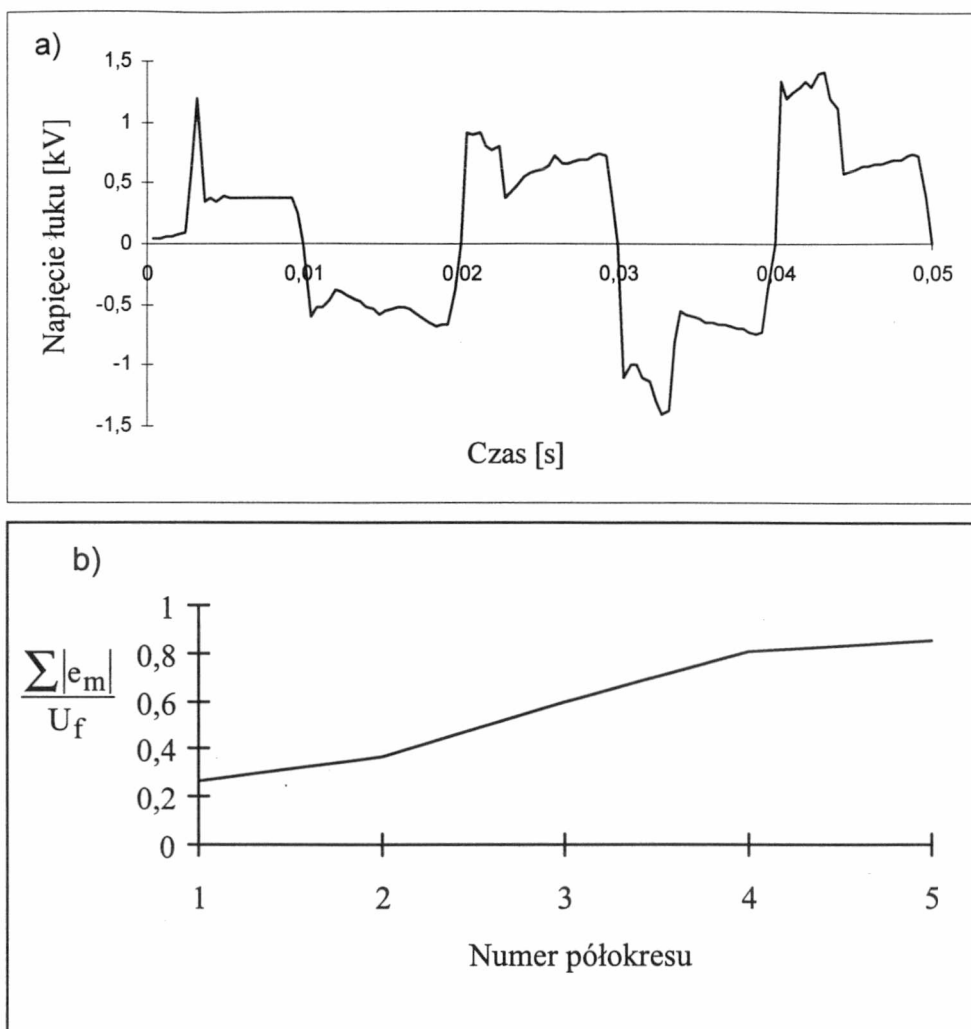
zwarciovych przez FGH [29, 30, 31], oraz na sygnałach z symulacji komputerowej otrzymanych z programu ATP (Alternative Transients Program) [62].



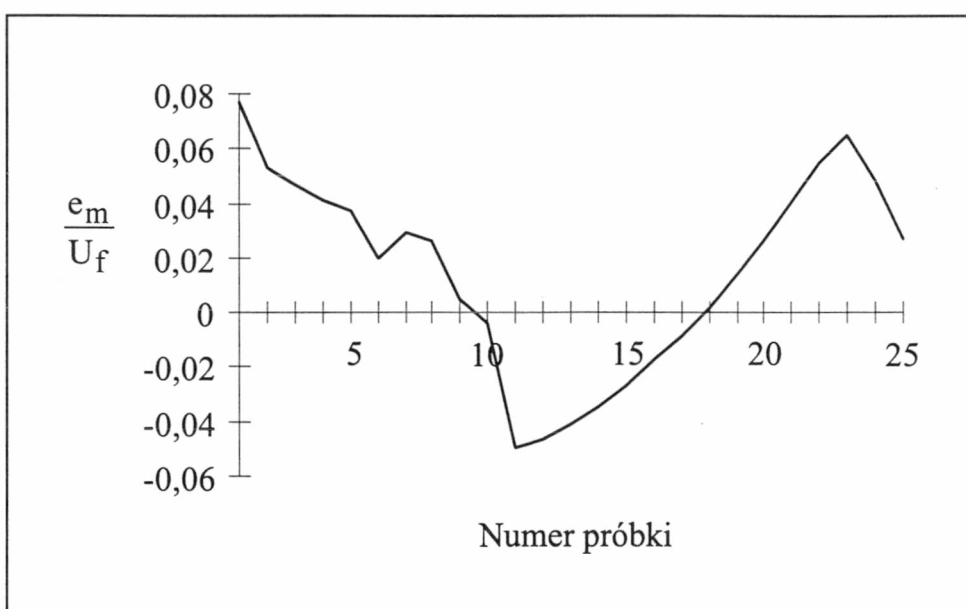
Rys. 7.3. Napięcie łuku (a) oraz suma błędów estymacji (b) podczas zwarcia w linii 20 kV w odległości 2 km



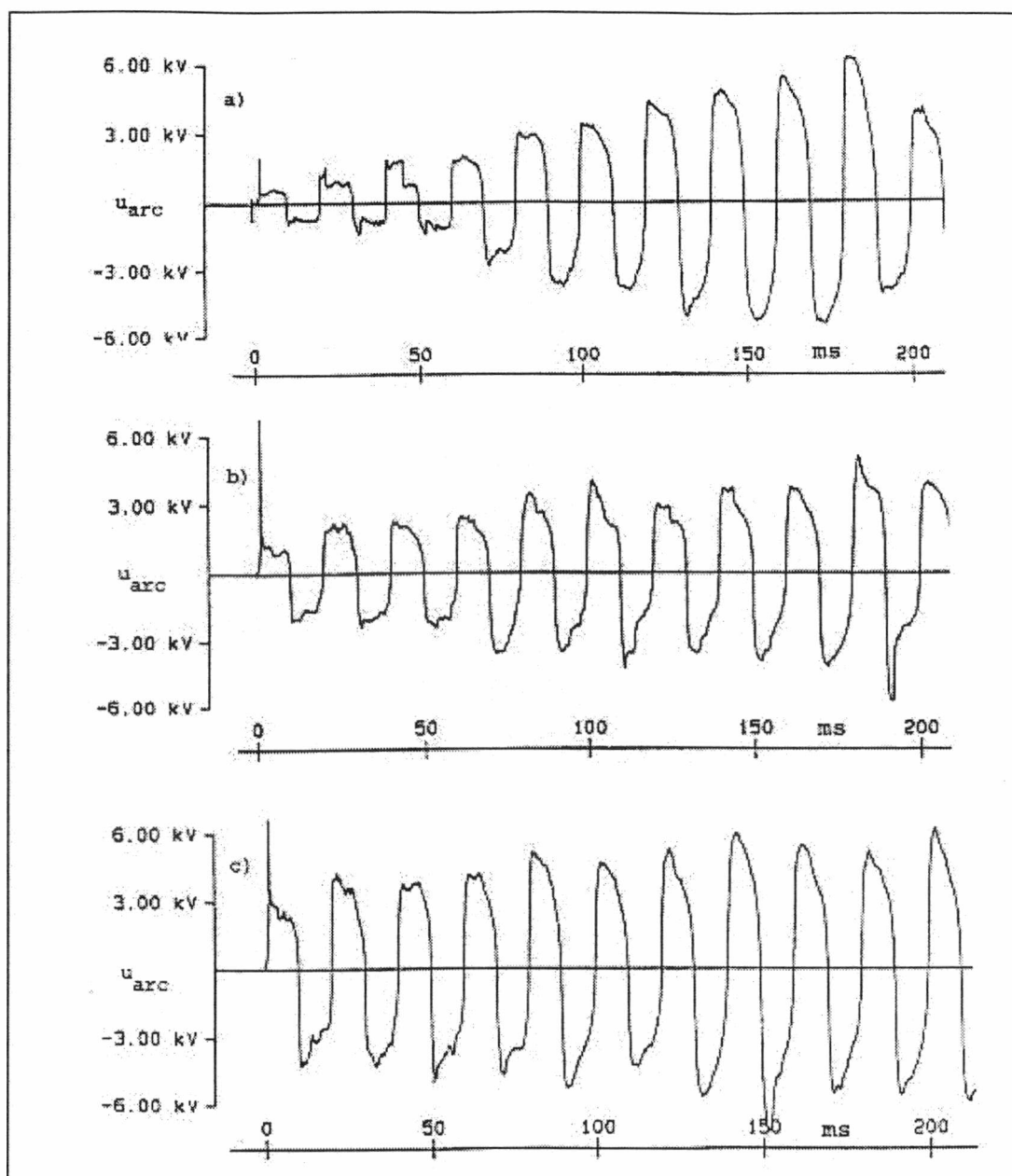
Rys. 7.4. Błędy estymacji podstawowej składowej napięcia łuku podczas zwarcia w linii 20 kV w odległości 2 km dla pierwszego półokresu przebiegu napięcia



Rys. 7.5. Napięcie łuku (a) oraz suma błędów estymacji (b) podczas zwarcia w linii 20 kV w odległości 10 km



Rys. 7.6. Błędy estymacji podstawowej składowej napięcia łuku podczas zwarcia w linii 20 kV w odległości 10 km dla piątego półokresu przebiegu napięcia

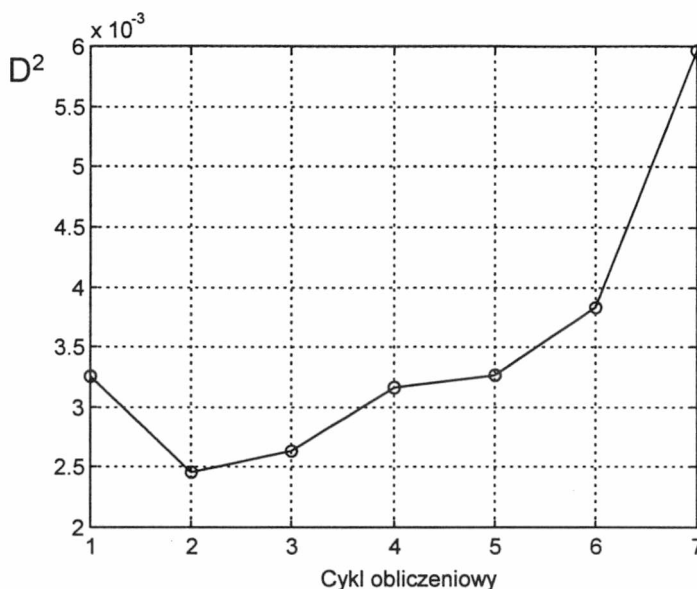


Rys. 7.7. Przebiegi napięcia łuku wzdłuż: a) 20 kV., b) 110 kV., 220 kV. układu izolatorów; długość początkowa łuku: 20 kV - 0.17 m, 110 kV - 1 m, 220 kV - 2.5 m; prąd łuku: a) 5.2 kA, b) 12.2 kA, c) 11.8 kA

Napięcie linii próbkowano z częstotliwością 2500 Hz. Na wejście sieci (rys. 7.1) dostarczano $N = 50$ próbek z jednego okresu przebiegu napięcia.

Rys. 7.8 przedstawia stopień zniekształcenia napięcia zgodnie ze wzorem (7.1) podczas zwarcia łukowego na linii przesyłowej 110 kV (rys. 7.7b). Przez pojęcie pierwszy cykl obliczeniowy należy rozumieć sytuację, w której do analizy brano próbki sygnału z pierwszego i drugiego półokresu; drugi cykl obliczeniowy - próbki z drugiego i trzeciego półokresu, itd. Z powodu występowania stanów nieustalonych oraz innych zjawisk, obserwuje się duże zakłócenia w pierwszym półokresie przebiegu napięcia.

W związku z tym, proponuje się przeprowadzać analizę stopnia zniekształcenia napięcia linii od drugiego lub trzeciego cyklu obliczeniowego, który upłynie od momentu powstania zakłócenia.

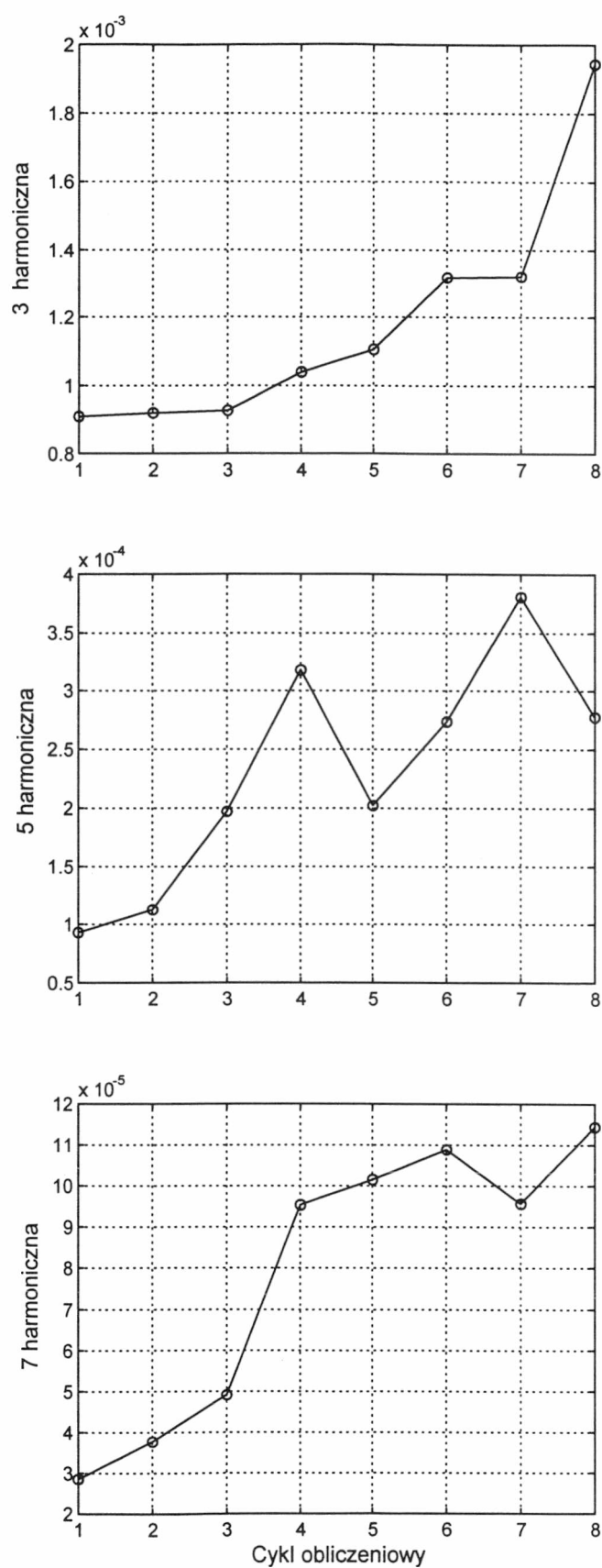


Rys. 7.8. Stopień zniekształcenia napięcia zgodnie z równaniem (7.1); zakłócenie na linii 110 kV (zgodnie z rys. 7.7b)

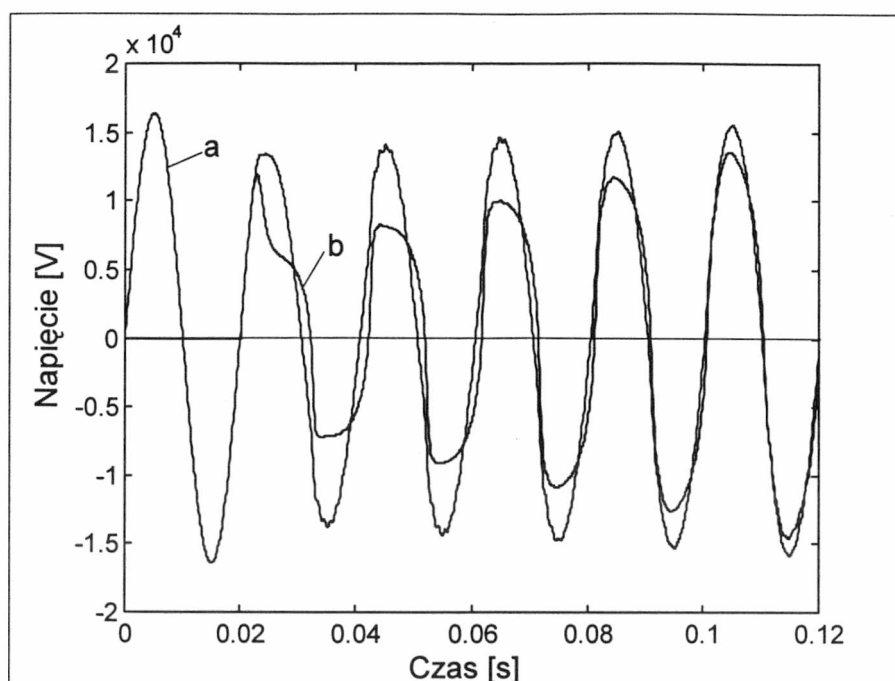
Podczas pierwszych sześciu, siedmiu półokresów trwania zakłócenia obserwuje się wyraźny wzrost wartości zniekształcenia napięcia. Po tym czasie długość łuku zazwyczaj pozostaje stała. Dodatkowo błędy estymacji sieci poddano analizie Fouriera, w celu wyznaczenia rozkładu wartości wyższych harmonicznych (rys. 7.9) dla cykli obliczeniowych opisanych wyżej. Na rys. 7.9 można zauważyć wyraźny wzrost 3-ej harmonicznej w ciągu pierwszych 8-u cykli trwania zakłócenia, zaś 7-ej harmonicznej - w ciągu 6-u cykli.

W ostatnim etapie pracy przeprowadzono symulację komputerową zwarcia łukowego z wykorzystaniem programu ATP [24, 30, 1, 25, 26]. Rys. 7.10 przedstawia przebieg napięcia na początku linii podczas trwania zwarcia, zaś rys. 7.11 stopień zniekształcenia napięcia dla tego przypadku.

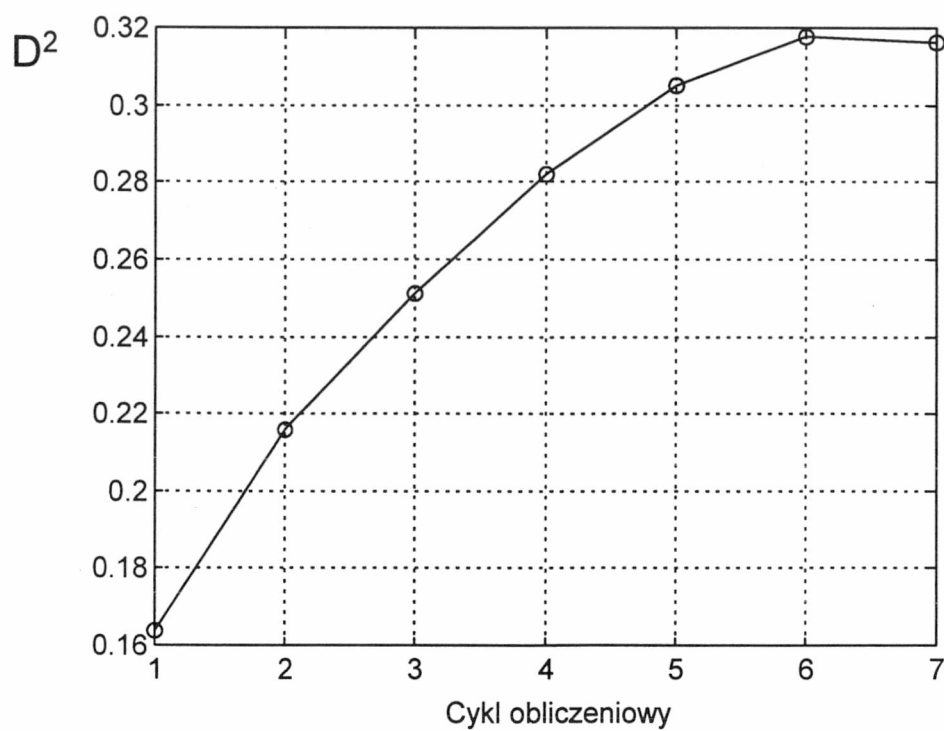
Dla przypadku, w którym łuk pierwotny nie wydłuża się, przeprowadzono symulację łuku wtórnego. Rys. 7.12 przedstawia przebieg wtórnego napięcia łuku, natomiast rys. 7.13 stopień jego zniekształcenia.



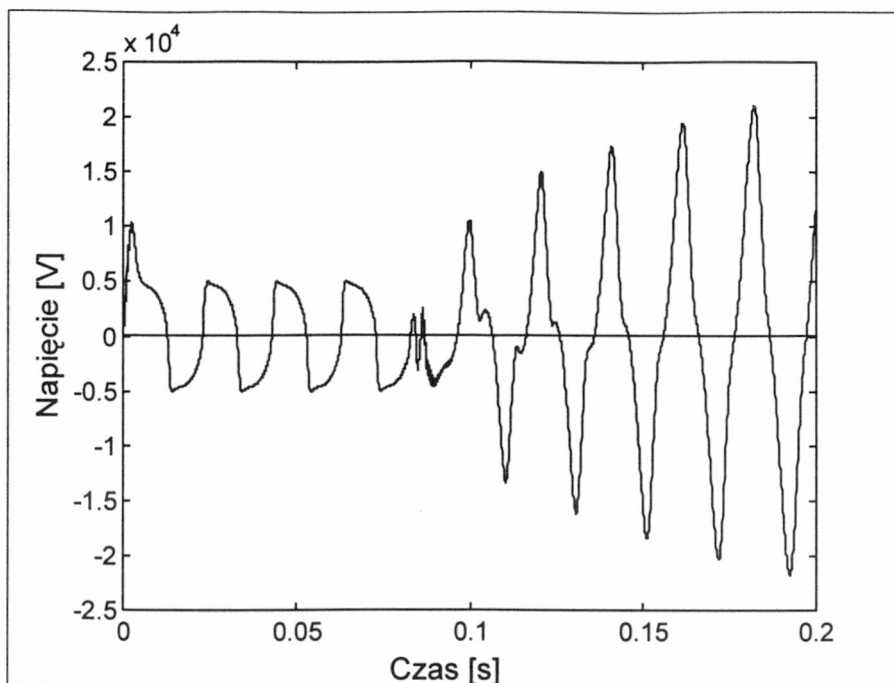
Rys. 7.9. Wyższe harmoniczne napięcia dla zwarcia w linii 110 kV



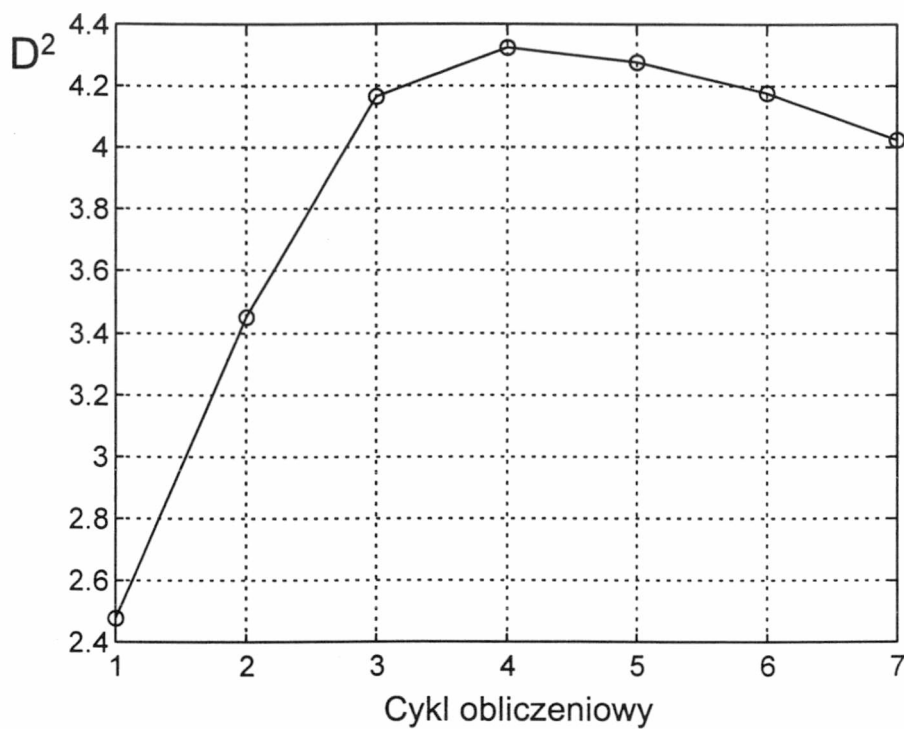
Rys. 7.10. Przebieg symulowanego napięcia łuku (a) oraz napięcia na początku linii (b)



Rys. 7.11. Zniekształcenie napięcia dla przebiegu (rys. 7.10)



Rys. 7.12. Napięcie łuku wtórnego



Rys. 7.13. Zniekształcenie napięcia łuku wtórnego (rys. 7.12)

Przeprowadzone badania potwierdziły, że przyrost błędów estymacji podstawowej składowej napięcia, spowodowany wydłużaniem się łuku, może stanowić kryterium rozróżniania zwarć łukowych od zwarć trwałych.

Wnioski

1. Opisane w rozprawie sieci neuronowe mogą być zastosowane do estymacji parametrów sygnałów w czasie rzeczywistym.
2. Sieci neuronowe optymalizujące mogą być projektowane na podstawie różnych kryteriów estymacji. Sieć zaprojektowana w oparciu o kryterium najmniejszego błędu kwadratowego ma najprostszą strukturę i odznacza się najkrótszym czasem obliczeń.
3. Czas obliczeń od momentu doprowadzenia do układu wszystkich próbek sygnału jest bardzo krótki. Zależy on od rodzaju sieci (kryterium estymacji) oraz od stałych czasowych integratorów. Stosując dostępne układy VLSI można osiągnąć czas obliczeń rzędu mikrosekund.
4. Sieci skonstruowane na podstawie nowego algorytmu opartego na kryteriach **TLS** i **RTLS**:
 - a) wykazały większą odporność na poziom zakłóceń szumami o rozkładzie normalnym oraz na poziom zakłóceń sygnału wyższymi harmonicznymi, w porównaniu z sieciami skonstruowanymi na podstawie standardowego kryterium LS;
 - b) dostarczają dokładniejszych wyników estymacji amplitudy w przypadku wystąpienia wahań częstotliwości w badanym sygnale - błędy nie przekraczają wartości 1 %;
 - c) czas konwergencji trajektorii wyznaczanych parametrów jest krótszy, niż w przypadku stosowania kryterium LS;
 - d) sieć oparta na kryterium TLS jest wrażliwa na zakłócenia impulsowe;
 - e) sieć skonstruowana na podstawie kryterium RTLS charakteryzuje się lepszą odpornością na zakłócenia impulsowe w porównaniu do kryterium TLS.
5. Dla wszystkich omówionych wyżej rodzajów sieci można stwierdzić, że:
 - a) dokładność estymacji amplitudy zależy od częstotliwości próbkowania oraz od długości okna próbkowania - wzrasta wraz ze wzrostem tych parametrów;
 - b) w trakcie przeprowadzonych badań najlepsze wyniki estymacji amplitudy dostarczała sieć o częstotliwości próbkowania $f_s = 2500$ Hz oraz oknie próbkowania $NT = 40$ ms;
 - b) wpływ na dokładność estymacji parametrów ma również odpowiedni dobór współczynnika uczenia sieci μ ; lepszą dokładność uzyskano dla: $\mu(t) = 1500 \exp(-50t)$, w porównaniu z $\mu(t) = \text{const}$.
6. Błędy estymacji, przy określonym kryterium estymacji i rodzaju zakłóceń, maleją wraz ze wzrostem długości okna pomiarowego NT. Liczba próbek, przy ustalonej długości okna, ma mniejszy wpływ na dokładność pomiaru.
7. Podczas pracy asynchronicznej maszyny synchronicznej prąd w uzwojeniach stojana może być rozłożony na składowe o różnych częstotliwościach, zależnych od poślizgu wirnika.
8. Dla identyfikacji stanu asynchronicznego możemy założyć pewną wartość poślizgu, która będzie osiągnięta podczas jego wzrostu.
9. Pojawienie się dodatkowych składowych prądu stojana X_2 , X_3 , X_4 oraz X_5 oznacza stan asynchroniczny maszyny.

10. Stosując sieć z $N = 30$ próbkami oraz oknem próbkowania $NT = 90$ ms można rozpoznać stan asynchroniczny już przy poślizgu 0.01.
11. W celu przeprowadzenia dalszych badań oraz wyeliminowania błędów przetwarzania istnieje potrzeba stworzenia sieci o większej rozdzielczości oraz większej dokładności.
12. Podczas zwarcia łukowego w linii elektroenergetycznej długość łuku może ulec zmianie. Wydłużanie się łuku podczas trwania zwarcia powoduje, że wzrasta napięcie łuku a tym samym stopień zniekształcenia napięcia linii. Prosta rekurencyjna sieć neuronowa umożliwia wyznaczanie parametrów podstawowej składowej napięcia i równocześnie stwarza możliwość odczytywania błędów estymacji tych parametrów. Wraz z wydłużaniem się łuku wzrastają błędy estymacji podstawowej składowej napięcia, będące różnicą między próbkami napięcia oraz wartościami wyznaczonej składowej podstawowej. Wzrost wartości zniekształcenia napięcia linii podczas trwania awarii wskazuje na zwarcie łukowe. W przypadku nie wydłużania się łuku pierwotnego należy wziąć pod uwagę łuk wtórny.

Literatura

- [1] **Al-Rawi A., Devaney M.:** Measurement of Secondary Arc Current in Transmission Lines Employing Single Phase Switching, *IEEE IMTC'97*, vol. 1, pp. 297-301, 1997.
- [2] **Alspector J., Gannet J. W., Haber S., Parker M. B., Chu R.:** A VLSI efficient technique for generating multiple uncorrelated noise sources and its application to stochastic neural networks, *IEEE Trans. Circ. Syst.*, vol. 38, pp. 109-123, 1991.
- [3] **Amari S. - I.:** Mathematical Foundations of Neurocomputing, *Proc. IEEE*, vol. 78, No. 9, pp. 1443-1463, 1990.
- [4] **Arun K. S.:** A unitary constrained total least squares problem in signal processing with more than one solution, *SIAM J. Matrix Anal. Appl.*, vol. 13, no. 3, pp. 728-745, 1992.
- [5] **Cichocki A., Kostyla P., Lobos T., Wacławek Z.:** Real time estimation of parameters of short circuit currents using neural networks, *6th Int. Symp. on Short Circuit Currents in Power Systems*, Liege, pp. 1.11.1-1.11.16, 1994.
- [6] **Cichocki A., Kostyla P., Lobos T., Wacławek Z.:** Neural Networks for Real-Time Estimation of Signals Encountered in Power Systems, presented at the *IMACS Int. Symp. on Signal Processing, Robotics and Neural Networks*, Lille (France), Conf. Proceedings, pp. 317-320, 1994.
- [7] **Cichocki A., Kostyla P., Lobos T., Wacławek Z.:** Neural networks for real time estimation of parameters of signals in power systems, *Int. Conf. on Intelligent System Application to Power Systems*, Montpellier, vol.2, pp. 749-756, 1992.
- [8] **Cichocki A., Lobos T.:** Adaptive Analogue Network for Real-Time Estimation of Basic Waveforms of Voltages and Currents, *IEE Proceedings - C*, vol. 139, No. 4, pp. 343-350, 1992.
- [9] **Cichocki A., Lobos T.:** Artificial Neural Networks for Real-Time Estimation of Basic Waveforms of Voltages and Currents, *IEEE Trans. on Power Systems*, vol. 9, No. 2, pp. 612-618, 1994.
- [10] **Cichocki A., Lobos T.:** Neural Networks for Real-Time Estimation of the Basic Waveforms of Voltages and Currents Encountered in Power Systems, *Int. J. Electronics*, vol. 74, No. 2, pp. 307-318, 1993.
- [11] **Cichocki A., Unbehauen R.:** *Neural Network for Optimization and Signal Processing*, Stuttgart, Teubner -Wiley, 1993.
- [12] **Cichocki A., Unbehauen R.:** Neural Network for solving systems of linear equations and related problems, *IEEE Trans. Circ. Syst.*, vol. 39, pp. 124-128, 1992.

- [13] **Cichocki A., Unbehauen R.:** Neural Network for solving systems of linear equations. Part II: Minimax and least absolute value problems, *IEEE Trans. Circ. Syst.*, vol. 39, pp. 619-633, 1992.
- [14] **Cichocki A., Unbehauen R.:** Simplified Neural Network for Solving Linear Least Squares and Total Least Squares Problems in Real Time, *IEEE Trans. On Neural Networks*, vol. 5, pp. 910-923, 1994.
- [15] **Clemens H., Rothe K.:** Schutztechnik in Elektroenergiesystemen, *Berlin: Verlag Technik*, pp. 173-174, 1991.
- [16] **Dalstein T., Kulicke B.:** Neural Network Approach to Fault Classification of High Speed Protective Relaying, *IEEE Trans. on Power Delivery*, vol. 10, no. 2, pp. 1002-1009, 1995.
- [17] **Djuric M. B., Terzija V. V., Skokljec I. A.:** Transmission Line Arcing Faults Recognition from the Voltage Signal, *11th PSCC*, Avignon (France), pp. 1037-1040, 1993.
- [18] **Djuric M. B., Radojevic Z. M., Terzija V. V.:** Digital Signal Processing Algorithm for Arcing Faults Detection on Transmission Lines, *12th Power Systems Computation Conference*, Dresden, Germany, vol. 2, pp. 681-687, 1996.
- [19] **Eichhorn K. F., Łobos T., Ruczewski P.:** Constrained Frequency Domain Algorithms for Determination of Parameters of Fundamental Sinewave of Signals, *IEE Proc.- C*, vol. 140, No. 6, pp. 477-480, 1993.
- [20] **Fa-Long L., Unbehauen R.:** *Applied Neural Networks for Signal Processing*, Cambridge University Press, 1997.
- [21] **Gill P., Murray W., Wright M.:** *Practical Optimization*, Academic Press, N. Y., 1981.
- [22] **Golub G. H., Van Loan C. F.:** An analysis of the total least squares problem, *SIAM J. Numer. Anal.* Vol. 17, pp. 883-893, 1980.
- [23] **Golub G. H., Van Loan C. F.:** *Matrix Computation*, Baltimore, Johns Hopkins Univ. Press, 1989.
- [24] **Hampel F. R., Ronchetti E. N., Stachel W. A.:** *Robust Statistics - The Approach Based on Influence Functions*, New York: Wiley, 1987.
- [25] **Johns A. T., Aggarwal R. K., Song Y. H.:** Improved Techniques for Modeling Fault Arcs on Faulted EHV Transmission Systems, *IEE - Proc.-Gener. Transm. Distrib.*, vol. 141, no. 2, pp. 148-154, 1994.
- [26] **Johns A. T., Al-Rawi M.:** Digital Simulation of EHV Systems Under Secondary Arcing Conditions Associated with Single-Pole Autoreclosure, *IEE Proc. C*, vol. 129, pp. 49-58, 1982.

- [27] **Kacprzak T., Ślot K.:** *Sieci neuronowe komórkowe*, Wydawnictwo Naukowe PWN, Warszawa, 1995.
- [28] **Kaczmarz S.:** Angenaherte Auflosung von Systemen linearer Gleichungen, *Bull. Acad. Poln. Sci. Lett. A*, vol. 35, pp. 355-357, 1937.
- [29] **Kizilcay M., Pniok T.:** Digital Simulation of Fault Arcs in Power Systems, *ETEP*, vol. 1, no. 1, pp. 55-60, January/February 1991.
- [30] **Kizilcay M., Koch K.-H.:** Numerical Fault Arc Simulation Based on Power Arc Tests, *ETEP*, vol. 4, No. 3, pp. 177-185, 1994.
- [31] **Koch K. H.:** Verzerrung der Meßgrößenspannung von Schutzrelais bei Lichtbogenkurzschlüssen, *Forschungsgemeinschaft für Hochspannungs - und Hochstromtechnik*, Report no. 8507, 1993.
- [32] **Koglin H. J., Lobos T.:** Recognition of Arcing Faults on Transmission Lines Using a Neural Network, *12th Power Systems Computation Conference*, Dresden, Germany, vol. 2, pp. 688-693, 1996.
- [33] **Koglin H. J., Lobos T., Waclawek Z.:** Neural Networks for Real -Time Estimation of Signal Parameters, *10th Int. Conference on Power System Protection (PSP'96)*, Bled, Slovenia, pp. 117-123, 1996.
- [34] **Korbicz J., Obuchowicz A., Uciński D.:** *Sztuczne sieci neuronowe - podstawy i zastosowania*, Akademicka Oficyna Wydawnicza PLJ, Warszawa, 1994.
- [35] **Kostyla P., Lobos T., Waclawek Z., Pytel J.:** Identification of asynchronous running of synchronous machines using a neural network, *XVII Seminarium z Podstaw Elektrotechniki i Teorii Obwodów*, Gliwice - Ustroń, Tom 2, ss. 497-502, 1994.
- [36] **Kucharski J.:** Badania nad wybranymi zjawiskami zakłóceniovymi i ruchowymi w systemie elektroenergetycznym dla potrzeb automatyki elektroenergetycznej, *Raport serii Sprawozdania nr 10*, Instytut Energoelektryki Politechniki Wrocławskiej, Wrocław, 1993.
- [37] **Kucharski J.:** Samowzbudzenie hydrogeneratora, *Raport serii Sprawozdania nr 22*, Instytut Energoelektryki Politechniki Wrocławskiej, Wrocław, 1995.
- [38] **Lobos T.:** Nonrecursive Methods for Real-Time Determination of Basic Waveforms of Voltages and Currents, *IEE Proc.- C*, vol. 136, No. 6, pp. 347-351, 1989.
- [39] **Mrozek B. i Z.:** *Matlab - uniwersalne środowisko do obliczeń naukowo-technicznych*, PLJ, Warszawa, 1996.
- [40] **O'Leary D. P.:** Robust regression computation using iteratively reweighted least squares, *SIAM J. Matrix Anal. Appl.*, vol. 11, pp. 466-480, 1990.
- [41] **Osowski S.:** *Sieci neuronowe*, WPW, Warszawa, 1994.

- [42] **Osowski S.:** *Sieci neuronowe w ujęciu algorytmicznym*, WNT, Warszawa, 1996.
- [43] **Osowski S.:** Neural Network for Estimation of Parameters of Sinewave, *Electronics Letters*, vol. 26, pp. 689-691, 1990.
- [44] **Osowski S.:** Neural Networks for Estimation of Harmonic Components in a Power System, *IEE Proc.- C*, vol. 139, No. 2, pp. 129-135, 1992.
- [45] **Osowski S.:** SVD Technique for Estimation of Harmonic Components in a Power System: a Statistical Approach, *IEE Proc.- C*, vol. 141, No. 5, pp. 473-479, 1994.
- [46] **Osowski S.:** Sztuczne sieci neuronowe - przegląd podstawowych struktur i metod uczenia, *XVI Seminarium z Podstaw Elektrotechniki i Teorii Obwodów*, Gliwice - Ustroń, Tom 1, ss. 31-52, 1993.
- [47] **Pytel J.:** *Elektromagnetyczne stany nieustalone*, Politechnika Wrocławska, Wrocław, 1983.
- [48] **Redfern M. A., Checksfield M. J.:** A Study into a New Solution for the Problems Experienced with Pole Slipping Protection, *IEEE Trans. on Power Delivery*, vol. 13, no. 2, April 1998.
- [49] **Song Y. H., Aggarwal R. K., Johns A. T.:** Digital Simulation of Fault Arcs on Long - Distance Compensated Transmission Systems with Particular Reference to Adaptive Autoreclosure, *ETEP*, vol. 5, no. 5, pp. 315-324, September/October 1995.
- [50] **Tadeusiewicz R.:** *Sieci neuronowe*, AOW, Warszawa, 1993.
- [51] **Tank D. W., Hopfield J. J.:** Simple neuron optimization networks: An a/d converter, signal decision circuits, and a linear programming circuit, *IEEE Trans. Circ. Syst.*, vol. CAS-33, pp. 533-541, 1986.
- [52] **Van Den Bos A.:** Nonlinear Least-Absolute Values and Minimax Model Fitting, *Automatica*, vol. 24, No. 6, pp. 803-808, 1988.
- [53] **Van Huffel S., Vandewalle J.:** The total least squares technique: computation, properties and applications, *SVD and Signal Processing*, E. F. Deprettere, North-Holland, pp. 189-207, 1988.
- [54] **Van Huffel S., Vandewalle J.:** Algebraic connections between the least squares and total least squares problems, *Numer. Math.*, vol. 55, pp. 431-449, 1989.
- [55] **Van Huffel S.:** On the significance of nongeneric total least squares problems, *SIAM J. Matrix Anal. Appl.*, vol. 13, pp. 20-35, 1992.
- [56] **Wacławek Z.:** *Wyznaczanie w czasie rzeczywistym parametrów sygnałów metodami sztucznych sieci neuronowych*, Rozprawa doktorska, Wrocław, 1996.

- [57] **Wang L.-X., Mendel J. M.:** Parallel structured networks for solving a wide variety of matrix algebra problems, *J. Parallel and Distributed Computing*, vol. 14, pp. 236-247, 1992.
- [58] **Wei M.:** Algebraic relations between total least squares and least squares problems with more than one solution, *Numer. Math.*, vol. 62, pp. 123-148, 1992.
- [59] **Widrow B., Lehr M.:** 30 years of adaptive neural networks: perceptron, madaline and back propagation, *IEEE Proc.*, vol. 78, pp. 1415-1442, 1990.
- [60] **Wright A., Christopoulos C.:** *Electrical Power System Protection*, Chapman & Hall, London, 1993.
- [61] **Yaozhong Ge, Fonghai Sui, Yuan Xiao:** Prediction Methods for Preventing Single-Phase Reclosing on Permanent Fault, *IEEE Trans. on Power Delivery*, vol. 4, No. 1, pp. 114-121, 1989.
- [62] **ATP (Alternative Transients Program),** Canadian / American EMTP User Group, 1987-92.
- [63] **MS EXCEL,** Microsoft Corporation, 1996.
- [64] **SIMULINK - User's Guide,** The Math Works, Inc., 1993.

