

Politechnika Warszawska
Wydział Elektroniki i Technik Informatycznych

Warszawa, 21 lutego 2017 r.

D z i e k a n a t

Uprzejmie informuję, że na Wydziale Elektroniki i Technik Informatycznych Politechniki Warszawskiej odbędzie się w dniu 7 marca 2017 r. publiczna obrona rozprawy doktorskiej

mgr. inż. Antoniego Wysockiego

temat: „Perceptronowe rekurencyjne sieci neuronowe w modelowaniu procesów dynamicznych i regulacji predykcyjnej”.

promotor – dr hab. inż. Maciej Ławryńczuk, prof. Politechniki Warszawskiej

recenzenci:

prof. dr hab. inż. Stefan Domek z Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie

prof. dr hab. inż. Marcin Witczak z Uniwersytetu Zielonogórskiego

Obrona odbędzie się w dniu 7 marca 2017 r. w sali 116 na Wydziale Elektroniki i Technik Informatycznych – Gmach im. Janusza Groszkowskiego, Warszawa, ul. Nowowiejska 15/19; początek godz. 10³⁰.

Po adresem: www.elka.pw.edu.pl/Wydzial/Rada-Wydzialu/Harmonogram-obron-doktorskich-streszczenia-i-recenzje zapewniony jest na stronie Wydziału dostęp do tekstów streszczenia rozprawy i recenzji, jak również do tekstu rozprawy umieszczonej w Bazie Wiedzy Politechniki Warszawskiej.

Dziekan



prof. dr hab. inż. Krzysztof Zaremba

mgr inż. Antoni Wysocki

Promotor

prof. nzw. dr hab. inż. Maciej Ławryńczuk

***Perceptronowe rekurencyjne sieci neuronowe
w modelowaniu procesów dynamicznych i regulacji predykcyjnej***

Streszczenie

Celem niniejszej pracy jest wszechstronne omówienie i porównanie różnych struktur rekurencyjnych sieci neuronowych oraz możliwości ich zastosowania do modelowania procesów dynamicznych i regulacji predykcyjnej. W pracy rozważa się rekurencyjną sieć perceptronową (RMLP), sieć Real Time Recurrent Network (RTRN), sieci Elmana (dwu- i trójwarstwową), sieć Jordana, sieć Elmana-Jordana, sieć RMLP-Elmana, sieci z własną rekurencją (dwu- i trójwarstwową) oraz sieci z rekurencją sąsiadującą (dwu- i trójwarstwową).

Po pierwsze, pokazano, iż rekurencyjne sieci neuronowe bazujące na perceptronie wielowarstwowym mogą znaleźć zastosowanie jako modele nieliniowych procesów dynamicznych. Uporządkowano i ujednolicono wzory dotyczące uczenia klasycznych sieci rekurencyjnych oraz wyprowadzono algorytmy uczenia sieci mniej popularnych w literaturze. Warto podkreślić, że w pracy rozważono modelowanie kilku typowych procesów dynamicznych inżynierii chemicznej, których właściwości statyczne i dynamiczne są nieliniowe oraz podjęto próbę wszechstronnego porównania właściwości rekurencyjnych sieci neuronowych pod względem dokładności otrzymanych modeli i liczby parametrów.

Po drugie, stwierdzono, że zastosowanie do modelowania klasycznych rekurencyjnych sieci neuronowych przy występowaniu w procesie znacznego opóźnienia, nie daje dobrych rezultatów, ponieważ otrzymane modele charakteryzują się dużą liczbą parametrów, ale dokładność modelowania pozostawia wiele do życzenia. Dlatego też zaproponowano prostą modyfikację struktur wszystkich omawianych sieci rekurencyjnych, umożliwiającą uwzględnienie opóźnienia. Dla rozważanych przykładów procesów dynamicznych okazało się, że zastosowanie zmodyfikowanej struktury sieci umożliwia łatwe otrzymanie modeli, których dokładność znacznie przewyższa dokładność struktur klasycznych i mają one jednocześnie znacznie mniej parametrów.

Po trzecie, pokazano, iż algorytmy regulacji predykcyjnej z cykliczną linearyzacją i bazujące na rozważanych rekurencyjnych sieciach neuronowych pozwalają na skuteczną regulację nieliniowych procesów dynamicznych. W pracy, dla kilku wybranych rekurencyjnych struktur sieci neuronowych, wyprowadzono szczegóły implementacji algorytmów regulacji predykcyjnej. Dla kilku reprezentatywnych procesów inżynierii chemicznej pokazano, że algorytmy regulacji predykcyjnej z linearyzacją charakteryzują się bardzo dobrą jakością regulacji, bliską algorytmom z nieliniową optymalizacją.

Szczecin, 9.12.2016 r.

Prof. dr hab. inż. Stefan Domek
Katedra Automatyki Przemysłowej i Robotyki
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

**RECENZJA ROZPRAWY DOKTORSKIEJ
DLA RADY WYDZIAŁU ELEKTRONIKI I TECHNIK INFORMACYJNYCH
POLITECHNIKI WARSZAWSKIEJ**

Tytuł rozprawy: **Perceptronowe rekurencyjne sieci neuronowe w modelowaniu procesów dynamicznych i regulacji predykcyjnej**

Autor rozprawy: **mgr. inż. Antoni Wysocki**

1. OGÓLNA CHARAKTERYSTYKA, ZAKRES I CEL ROZPRAWY

Jakie zagadnienie naukowe jest rozpatrzone w pracy /teza rozprawy/ i czy zostało ono dostatecznie jasno sformułowane przez autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?

Jednym z obszarów wiedzy, wzbudzającym od kilkadziesiąt lat duże zainteresowanie teoretyków i praktyków automatyki, jest wykorzystanie sztucznych sieci neuronowych, zwłaszcza rekurencyjnych, w modelowaniu i sterowaniu. Sieci neuronowe pozwoliły rozwiązać wiele praktycznych problemów charakterystycznych dla układów regulacji tzw. trudnych obiektów. Niestety, w wielu przypadkach proponowane rozwiązania miały charakter jednostkowy i nie wynikały z szerszych analiz bądź badań porównawczych. W sytuacji wielkiej różnorodności typów i struktur sieci neuronowych utrudnia to znacząco rozwój ich zastosowań.

Autor rozprawy postanowił wszechstronnie zbadać różne struktury rekurencyjnych sieci neuronowych pod kątem ich przydatności do modelowania procesów dynamicznych i regulacji predykcyjnej. Swoje zadanie ujął w postaci trzech powiązanych celów naukowo-badawczych, mających udowodnić sformułowane w pracy tezy:

– po pierwsze postanowił porównać różne struktury rekurencyjnych sieci neuronowych bazujących na perceptronie wielowarstwowym pod względem liczby parametrów i jakości modelowania oraz wykazać, że modele nieliniowych procesów dynamicznych w postaci takich sieci charakteryzują się bardzo dużą dokładnością, przewyższającą modele liniowe oraz modele z sieciami nierekurencyjnymi;

– po drugie zaproponował autorską modyfikację klasycznych struktur rekurencyjnych sieci neuronowych, umożliwiającą modelowanie procesów ze znacznym opóźnieniem a następnie sprawdził efektywność struktur zmodyfikowanych na tle rozwiązań klasycznych;

– po trzecie wykorzystał zaproponowane struktury rekurencyjnych sieci neuronowych do syntezy kilku typów algorytmów regulacji predykcyjnej, pokazując ich skuteczność w regulacji silnie nieliniowych procesów dynamicznych.

Proponowane rozwiązania zweryfikował przy pomocy licznych badań symulacyjnych, przeprowadzonych z wykorzystaniem modeli kilku reprezentatywnych (benchmarkowych) procesów dynamicznych wykorzystywanych w inżynierii chemicznej.

Uważam, że podjęcie tematu rozprawy doktorskiej z tego zakresu było celowe i potrzebne, zarówno ze względów poznawczych, teoretycznych jak i praktycznych a postawione cele i jasno sformułowane tezy określają istotne i aktualne na tle obecnego stanu wiedzy, a zarazem oryginalne, zadanie badawcze.

2. ZAWARTOŚĆ MERYTORYCZNA ROZPRAWY

Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł / w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle /świadczący o dostatecznej wiedzy autora. Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?

Recenzowana rozprawa liczy 172 strony, zawiera 64 rysunki oraz 25 tabel. Podzielona jest na osiem rozdziałów, poprzedzonych streszczeniami w języku polskim i angielskim oraz uzupełnionych spisem literatury.

W rozdziale pierwszym, Autor omówił założenia, cel, zakres i tezy pracy. Wymienił również siedem prac autorskich (lub współautorskich z promotorem rozprawy) ściśle związanych z tematyką rozprawy, opublikowanych głównie w materiałach ważnych konferencji naukowych, wydanych przez znane wydawnictwa. Świadczy to o Jego dużej aktywności badawczej i umiejętności prezentowania własnych dokonań naukowych. Ponieważ publikacje powstały w latach 2013 – 2016, a więc równoległe z rozprawą doktorską, stanowią ważne uzupełnienie rozważań z rozprawy i są w kilku jej miejscach (np. na stronach 27, 56, 147) przywołane przez Autora, jako źródło szczegółowych danych.

W kolejnych dwóch rozdziałach Autor przypomniał w sposób uporządkowany ogólnie znane i będące podstawą dalszych rozważań zagadnienia dotyczące sieci neuronowych, oraz przedstawił szczegółowe zależności wyprowadzone w sposób systematyczny dla konkretnych typów i struktur sieci rekurencyjnych. I tak:

– w rozdziale drugim opisał precyzyjnie jednaście różnych struktur rekurencyjnych sieci neuronowych bazujących na perceptronie wielowarstwowym. Dla każdej sieci pokazał jej schemat strukturalny oraz wyznaczył relację między jej strukturą a liczbą parametrów modelu. W podsumowaniu wymienił również inne, mniej popularne struktury, wykazując ogromne rozeznanie w tematyce sieci rekurencyjnych;

– w rozdziale trzecim przypomniał gradientowe algorytmy uczenia sieci rekurencyjnych a następnie przedstawił szczegółowe wzory wyliczania gradientów dla omawianych sieci.

W początkowej części rozdziału czwartego Autor omówił dwa procesy chemiczne, neutralizacji i polimeryzacji, podając równania opisujące ich właściwości statyczne i dynamiczne. Podobnie w rozdziale piątym przypomniał ogólną zasadę regulacji predykcijnej i omówił sposoby realizacji algorytmów predykcyjnych dla procesów nieliniowych, w tym znane algorytmy suboptymalne. Omówione modele i algorytmy wykorzystano później do przeprowadzenia badań symulacyjnych.

Zasadniczą, oryginalną część rozprawy stanowi moim zdaniem druga część rozdziału czwartego oraz rozdziały szósty i siódmy, w których Autor przedstawił sposoby implementacji rozważanych struktur neuronowych sieci rekurencyjnych w omówionych wcześniej modelach nieliniowych procesów chemicznych oraz w nieliniowych suboptymalnych regulatorach predykcyjnych, dokonał analizy ich skuteczności i złożoności numerycznej oraz przeprowadził obszerne badania symulacyjne, uzupełnione o krytyczną analizę uzyskanych wyników.

W rozdziale ósmym Autor podsumował zawartość pracy i uzyskane wyniki. Sformułował również liczne wnioski natury ogólnej i szczegółowej oraz wskazał na możliwe dalsze kierunki badań.

W zakończeniu rozprawy podał spis literatury obejmujący 95 pozycji, w większości trafnie dobranych i poprawnie cytowanych w tekście rozprawy.

3. OGÓLNA OCENA ROZPRAWY

Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?

Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?

Czy autor wykazał umiejętność poprawnego i przekonującego przedstawienia uzyskanych przez siebie wyników /zwięzłość, jasność, poprawność redakcyjna rozprawy/?

Jakie są słabe strony rozprawy i jej główne wady?

Jaka jest przydatność rozprawy dla nauk technicznych?

Autor rozprawy zrealizował postawione cele w sposób kompleksowy, używając właściwej metodyki badań, przejrzyście przywołując źródła inspiracji i założenia proponowanych metod. Przeprowadzona analiza źródeł literaturowych przedmiotu, obszerność rozważań, duża liczba eksperymentów i krytyczna ocena uzyskanych wyników zasługują na pochwałę i stanowią moim zdaniem główne osiągnięcie Autora.

Do najważniejszych dokonań Doktoranta zaliczam:

1. systematyczny i przejrzysty opis podstawowych typów sieci neuronowych z oceną ich przydatności do modelowania procesów dynamicznych;
2. wyprowadzenie i zebranie w jednolitej formie szczegółowych wzorów i zależności opisujących poszczególne struktury sieci;
3. analizę teoretyczną złożoności parametrycznej rozważanych modeli neuronowych;
4. zaproponowanie modyfikacji rozważanych struktur sieci ułatwiającej modelowanie procesów o dużych opóźnieniach bez konieczności zwiększania złożoności parametrycznej struktur;

5. zbadanie wpływu typu sieci i liczby jej parametrów (dla przyjętej struktury) na jakość modelowania wyrażoną przez błędy modelu dla zbiorów danych uczących, weryfikujących i testowych;
6. przeprowadzenie dużej liczby badań symulacyjnych.

Dużym walorem pracy jest również bardzo szczegółowa prezentacja uzyskanych wyników. Oprócz czytelnych przebiegów czasowych pokazanych na licznych rysunkach, Autor zestawiał w wielu tabelach najważniejsze wyniki, w tym liczby parametrów modelu dla większości badanych struktur sieci oraz wskaźniki jakości modelowania dla różnych zbiorów danych.

Podczas lektury rozprawy nasunęło mi się kilka uwag, o charakterze częściowo dyskusyjnym, ale wartych komentarza Doktoranta:

1. zaproponowana i sformułowana w drugiej tezie rozprawy modyfikacja struktur sieci polegająca na uzupełnieniu klasycznych struktur o dodatkowe opóźnienia w sygnale sterującym i uzyskanie w ten sposób porównywalnej jakości modelowania przy znacznie mniejszej liczbie parametrów sieci (str. 153 rozprawy) jest z punktu widzenia teorii sterowania dość oczywista (m.in. powszechnie znane jest modelowanie obiektów inercyjnych wysokich rzędów o wielu parametrach jednoparametrowym zastępczym opóźnieniem). Potwierdza to sam Autor podając takie oceny już na wstępie pracy, np. na stronach 20, 22 i 27, przed jakąkolwiek analizą symulacyjną.
2. jak Autor skomentuje efekt umiejscowienia dodatkowych elementów opóźniających na wejściach sieci, jeśli opóźnienia w modelowanym procesie dotyczyłyby zmiennych stanu bądź wyjść (jak w przykładzie 7.3 na stronie 149)? Czy opóźnienie pomiarowe 100 s w tym przykładzie wynika z uwarunkowań technologicznych (o których nie pisano w rozdz. 4.1) czy zostało wprowadzone wyłącznie dla pokazania skuteczności proponowanej modyfikacji?
3. na stronie 156 rozprawy słusznie zauważono, że proponowana modyfikacja sieci rekurencyjnej wymaga znajomości opóźnienia pomiarowego. Czy Autor przeprowadził jakies badania symulacyjne wrażliwości proponowanych zmodyfikowanych struktur sieci na błąd określenia opóźnienia, zwłaszcza w przypadku przeszacowania jego wartości?
4. reaktor neutralizacji w rozdziale 4 opisany został mało dokładnie. Na przykład q_1 na rys. 4.1 oznacza zasadę a w tekście kwas, we wzorach (4.1), (4.2) i (4.3) nie opisano oznaczeń, podobnie w tabeli 4.1. Dlaczego przyjęto takie a nie inne skalowanie sygnałów? Dlaczego nie wspomniano o wersjach 1x1 i 2x2 badanych w rozdziałach 7.2 i 7.3? Dlaczego nie wspomniano o opóźnieniach pomiarowych (patrz uwaga 2)?
5. podobne uwagi dotyczą reaktora polimeryzacji z rozdz. 4.2. Zmienne stanu w tabeli 4.11 nie zostały opisane i nie mają wymiaru fizycznego, podobnie sygnały wejściowe! Nie opisano skrótu NAMW (Number Average Molecular Weight). Które wartości są właściwe: $y = D_1/D_0$ czy $y = (D_1(t)/D_0(t) - D_{10}/D_{00})/10000$? Podobnie $u = F_1$ czy $u(t) = (F_1(t) - F_{10}) \cdot 100$? Niejednoznaczny opis reaktorów pojawia się również na str. 94 gdzie w opisie wykresów pisze się „o efekcie niedociągnięcia wartości pH” a na rysunkach pokazano NAMW;

6. ze względów formalnych algorytm uczenia sieci na str. 54 powinien być uzupełniony w kroku 3. o warunek stopu: „... albo osiągnięta została założona maksymalna liczba iteracji, algorytm kończy swoje działanie”. Z tych samych względów krok 4. algorytmu na str. 125 powinien być uzupełniony o zapis „... w przeciwnym przypadku przechodzi się do kroku 5” a krok 4.4. o zapis „... iteracje są przerywane i przechodzi się do kroku 5. W przeciwnym przypadku ...”.
7. na str. 127 Autor podał kluczowy dla regulacji predykcyjnej wzór (6.3) na pośredni pomiar zakłóceń i zarazem na modelowanie zakłóceń, zakładając ich stałość w horyzoncie predykcji – wzór (6.2). Wzór (6.3) jest wielokrotnie wykorzystywany w dalszej części rozprawy a nawet powtórzony na str. 136, jako (6.43). Czy proponowany zapis obarczony jest błędem redakcyjnym, czy Autor świadomie go używał zamiast poprawnego moim zdaniem wzoru $d(k) = y(k) - y_{mod}(k|k-1)$?

I jeszcze dwie subiektywne uwagi natury redakcyjnej:

1. zgadzam się z Autorem, że perceptronowa sieć rekurencyjna jest bardzo szeroko opisana w literaturze. Skoro jednak Autor zdecydował się na osobny rozdział 3.2.1, to warto byłoby moim zdaniem przytoczyć szczegółowe wzory na gradienty dla tej sieci dla ułatwienia lektury rozprawy;
2. podrozdział 4.3 dotyczy modelowania reaktora polimeryzacji. Moim zdaniem z pewnością byłoby więc gdyby był włączony do podrozdziału 4.2 z numeracją 4.2.3. Podobnie nawiązując do tytułu rozdziału 4 fragment rozprawy ze stron 152 – 155, z rysunkami 7.9 – 7.13 i tabelami 7.2, 7.3, mógłby być przeniesiony do rozdziału 4.1.

4. UWAGI SZCZEGÓŁOWE

Praca napisana jest poprawnym i zrozumiałym językiem oraz bardzo starannie zredagowana, z małą liczbą drobnych błędów literowych bądź edytorskich, np.:

1. na str. 17 powinno być „dodatkowo przyjmuje się $n_u + 1$ parametrów $n_{b_i} \geq \tau_i$ oraz $n_a > 0$ ”;
2. na stronach 29 i 33 podano, że dla sieci Jordana stosunek liczby parametrów bezpośrednio związanych z rekurencją sieci do całkowitej liczby parametrów (zakładając $n_u = 1$) wynosi 1:3 a powinno być jest mniejszy od 1:4. Dobrą wartość podano na str. 35;
3. w tabeli 2.1 na str. 47 dla $K, n_a, n_b = 5, 1, 1$ powinno być odpowiednio 21 oraz 47;
4. na stronie 73 napisano o wygenerowaniu trzech zbiorów danych po 1800 próbek każdy a powinno być po 1200 (jak na rysunkach umieszczonych na następnych stronach);
5. we wzorze (5.3) na str. 114 zamiast symboli normy powinny być nawiasy. We wzorze brakuje również istotnego ograniczenia $\Delta u(k+p|k) = 0$ dla $N_u \leq p \leq N$;
6. na stronie 115 za wzorem (5.6) powinno być $\|x\|_A^2 = x^T A x$;

7. we wzorze (5.4) na str. 115 pojawiły się bez komentarza współczynniki wagowe μ_m chociaż nie było ich we wcześniejszym sformułowaniu funkcji kosztu (5.2);
8. we wzorze (5.14) na str. 118 powinno być n_{a_m} zamiast n_a^m ;
9. we wzorze (5.21) na str. 119 powinno być $\mathbf{y}^{min} \leq \mathbf{G}(k)\Delta u(k) + \mathbf{y}^0(k) \leq \mathbf{y}^{max}$;
10. na str. 122 powinno być $\mathbf{H}(k)$ zamiast $\mathbf{H}^1(k)$ – to nie jest wersja iteracyjna algorytmu;
11. we wzorze (5.40) na str. 124 powinno być raczej $y^{zad}(k-p)$ a nie $y^{sp}(k-p)$;
12. odwołania do zadania optymalizacji (5.7) na stronach 117 i 121 powinny być zamienione na (5.3) lub (5.13). Wzór (5.7) nie opisuje zadania optymalizacji tylko funkcję kosztu;
13. sformułowanie na str. 113 „Zmienne decyzyjne (5.1) oblicza się w taki sposób, aby prognozowane uchyby regulacji były jak najmniejsze” powinno być zmienione na np.: „Zmienne decyzyjne (5.1) oblicza się w taki sposób, aby prognozowane uchyby regulacji były optymalne w sensie przyjętej funkcji kosztu”.

Zauważone błędy w żadnym stopniu nie obniżają mojej wysokiej oceny rozprawy.

5. WNIOSKI KOŃCOWE

Do której z następujących kategorii Recenzent zalicza rozprawę:

- a/ nie spełniająca wymagań stawianych rozprawom doktorskim przez obowiązujące przepisy*
- b/ wymagająca wprowadzenia poprawek i ponownego recenzowania*
- c/ spełniająca wymagania*
- d/ spełniająca wymagania z wyraźnym nadmiarem*
- e/ wybitnie dobra, zasługująca na wyróżnienie*

W moim odczuciu rozprawa ma wybitnie poznawczy charakter i pokazuje duże potencjalne możliwości wykorzystania rekurencyjnych sieci neuronowych do modelowania właściwości statycznych i dynamicznych złożonych procesów przemysłowych o wielu wejściach i wyjściach oraz do budowy efektywnych algorytmów regulacji predykcyjnej takich procesów.

Dokonania Autora przedstawione w rozprawie uważam za oryginalne i stanowiące zauważalny wkład w wykorzystanie teorii neuronowych sieci rekurencyjnych w automatyce a zadanie naukowe określone w tytule i celach rozprawy za jasno sformułowane a następnie poprawnie rozwiązane przy użyciu właściwych metod matematycznych oraz potwierdzone na drodze symulacji wykonanych przy pomocy nowoczesnych narzędzi numerycznych. Autor w sposób kompleksowy zebrał rozproszone w literaturze światowej informacje, porównał wady i zalety poszczególnych typów sieci, zaproponował własną modyfikację klasycznych rozwiązań zwiększającą obszar stosowalności rozważanych sieci przy zachowaniu umiarkowanego stopnia złożoności ich struktury, uzasadnił celowość użycia rozważanych sieci do syntezy suboptymalnych algorytmów sterowania predykcyjnego oraz podał sposób ich realizacji obliczeniowej. Tym samym wykazał, że posiada niezbędną wiedzę w zakresie dyscypliny naukowej *automatyka i robotyka* oraz że ma duże predyspozycje do prowadzenia badań

naukowych, wcześniej potwierdzone publikacjami autorskimi wymienionymi w wykazie literatury.

Podsumowując stwierdzam, że przedstawiona mi do recenzji praca doktorska mgr. inż. Antoniego Wysockiego pt. *Perceptronowe rekurencyjne sieci neuronowe w modelowaniu procesów dynamicznych i regulacji predykcyjnej* spełnia z nadmiarem wymagania Ustawy z dnia 14 marca 2003 roku o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. Nr 65, poz. 595, z późn. zm.) stawiane rozprawom doktorskim, jak również kryteria zwyczajowo przyjęte w tym zakresie w krajowym środowisku naukowym w dyscyplinie *automatyka i robotyka*, i wnioskuję o jej przyjęcie oraz dopuszczenie do publicznej obrony.



prof. dr hab. inż. Marcin Witczak
Instytut Sterowania i Systemów Informatycznych
Wydział Informatyki, Elektrotechniki i Automatyki
Uniwersytet Zielonogórski
ul. Prof. Z. Szafrana 2
65-516 Zielona Góra

09.12.2016

***KWESTIONARIUSZ - RECENZJA ROZPRAWY DOKTORSKIEJ DLA RADY
WYDZIAŁU ELEKTRONIKI I TECHNIK INFORMACYJNYCH
POLITECHNIKI WARSZAWSKIEJ***

Tytuł rozprawy: Perceptronowe rekurencyjne sieci neuronowe w modelowaniu procesów dynamicznych i regulacji predykcyjnej

Autor rozprawy: Antoni Wysocki

1. Jakie zagadnienie naukowe jest rozpatrzone w pracy /teza rozprawy/ i czy zostało ono dostatecznie jasno sformułowane przez autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?

Problemem naukowym stanowiącym przedmiot recenzowanej rozprawy doktorskiej jest opracowanie oraz analiza struktur i metod perceptronowych sieci neuronowych dla potrzeb modelowania procesów dynamicznych i regulacji predykcyjnej.

Bieżący rozwój technologiczny oraz związane z nim zaawansowane procesy technologiczne posiadające silnie nieliniowy charakter rodzą potrzebę konstruowania coraz bardziej wyrafinowanych technik sterowania. Wymagają one odpowiedniej wiedzy o sterowanym procesie, które zazwyczaj sprowadza się do znajomości jego modelu matematycznego. Niestety, często okazuje się, że modele analityczne (budowane na podstawie praw fizyki, chemii, itd.) nie odzwierciedlają w pełni zachowania rzeczywistego procesu lub posiadana wiedza nie pozwala na zbudowanie ich odpowiedniej struktury. Remedium na tego typu problemy są modele zachowań, których struktura nie odzwierciedla rzeczywistego procesu, a jedynie jego zachowanie wejściowo wyjściowe. W związku z powyższym faktem, w literaturze nazwa się je "czarnymi skrzynkami". Niewątpliwie sztuczne sieci neuronowe rozważane w recenzowanej rozprawie należą właśnie do powyższej grupy modeli.

Powyższe uwarunkowania przekształciły teorię modelowania neuronowego ze sztuki projektowania zadowalająco działającego modelu do nowoczesnej gałęzi nauki jaką jest dzisiaj. Na przestrzeni ostatnich kilkudziesięciu lat zostało opracowanych wiele różnego rodzaju struktur sieci neuronowych. W kontekście zastosowań modelowania dynamicznych procesów przemysłowych, szczególnie atrakcyjne wydają się być rekurencyjne sieci neuronowe będące przedmiotem rozprawy. Oprócz możliwości modelowania samego procesu dają one możliwość zaprojektowania układu sterowania o pożądanej jakości. W kontekście układów sterowania, na szczególną uwagę zasługują: sterowanie optymalne, sterowanie predykcyjne, sterowanie adaptacyjne, sterowanie

odporne oraz sterowanie ślizgowe. W ostatnich latach pojawiają się również rozwiązania hybrydowe łączące wspomniane klasyczne metody sterowania z metodami obliczeń inteligentnych, takimi jak, np. sztuczne sieci neuronowe. Podejścia hybrydowe dają możliwość łączenia zalet powyższych rozwiązań przy jednoczesnej eliminacji ich wad. W recenzowanej rozprawie, rozwiązanie hybrydowe bazuje na rekurencyjnych sieciach i sterowaniu predykcyjnym. Zagadnienie projektowania strategii sterowania predykcyjnego jest bardzo ważnym i stale aktualnym problemem teorii sterowania. Od lat 60tych ubiegłego stulecia, techniki sterowania predykcyjnego są stale rozwijane oraz doskonalone. W literaturze pokazuje się również znaczącą liczbę przykładów ich praktycznego zastosowania pokazujących przejrzyste korzyści płynące z implementacji sterowania predykcyjnego.

Przedstawiona rozprawa skupia się na opracowaniu struktur rekurencyjnych sieci neuronowych oraz bazujących na nich strategii sterowania predykcyjnego.

Celem przedstawionych badań było udowodnienie następujących tez:

- Rekurencyjne sieci neuronowe bazujące na perceptronie wielowarstwowym zastosowane jako modele nieliniowych procesów dynamicznych charakteryzują się bardzo dużą dokładnością, przewyższającą modele liniowe oraz sieci nierekurencyjne.
- Korzystne jest wprowadzenie dodatkowego opóźnienia do klasycznych rekurencyjnych sieci neuronowych, co umożliwi modelowanie procesów ze znacznym opóźnieniem.
- Algorytmy regulacji predykcyjnej z cykliczną linearyzacją i bazujące na rekurencyjnych sieciach neuronowych pozwalają na skuteczne regulacje nieliniowych procesów dynamicznych.

Powyższe tezy sformułowane są w przejrzysty sposób i jasno odzwierciedlają zadania badawcze postawione w rozprawie. Sama rozprawa ma charakter teoretyczny i obejmuje następujące zagadnienia:

- Przegląd, analiza i opracowanie struktur rekurencyjnych sieci neuronowych;
- Symulacyjna analiza efektywności modelowania opracowanych rekurencyjnych sieci neuronowych;
- Opracowanie metod sterowania predykcyjnego z zastosowaniem rekurencyjnych sieci neuronowych;
- Symulacyjna analiza efektywności opracowanych strategii sterowania predykcyjnego bazującego na rekurencyjnych sieciach neuronowych.

2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł / w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle /świadczący o dostatecznej wiedzy autora. Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?

Przedstawiona do recenzji rozprawa doktorska zawiera 8 rozdziałów i bibliografię. Podstawę pracy stanowią rozdziały od 3'ego do 7'ego, które prezentują oryginalne wyniki naukowe Doktoranta w zakresie projektowania struktur, algorytmów uczenia oraz strategii sterowania bazujących na rekurencyjnych sieciach neuronowych. Oryginalne wyniki pracy naukowej doktoranta dotyczą również zastosowania opracowanych metod oraz technik modelowania i sterowania dla wybranych procesów przemysłowych. W szczególności, rozdział 1 stanowi wstęp rozprawy, w którym Autor formułuje cele i tezy pracy doktorskiej. W drugim rozdziale omówiono różne struktury rekurencyjnych sieci neuronowych bazujących na perceptronie wielowarstwowym. Rozdział 3 przedstawia algorytmy uczenia sieci rekurencyjnych. Rozdział 4 poświęcony jest zastosowaniom rekurencyjnych sieci neuronowych do modelowania procesów dynamicznych. W rozdziale piątym omówiono ogólną zasadę regulacji predykcyjnej oraz sformułowania wybranych nieliniowych algorytmów regulacji predykcyjnej. Rozdział 6 poświęcono omówieniu implementacji algorytmów regulacji predykcyjnej dla wybranych sieci rekurencyjnych. W rozdziale 7 przedstawia się wyniki symulacji algorytmów predykcyjnych opartych na rekurencyjnych sieciach neuronowych. Rozdział 8 podsumowuje rozprawę i przedstawia kierunki dalszych badań.

Wykaz bibliograficzny obejmuje 96 pozycji, które bardzo dobrze odzwierciedlają istniejący stan wiedzy w zakresie modelowania i sterowania neuronowego. Autor wykorzystał powyższe pozycje do skonstruowania przejrzystego opisu stanu wiedzy w zakresie modelowania i sterowania neuronowego. Na szczególną uwagę zasługują rozdziały 2 i 3, których części przeglądowe stanowią znakomite kompendium wiedzy w zakresie badań poruszanych w rozprawie. Świadczy to o bardzo dobrej wiedzy Autora dotyczącej teorii i zastosowania przemysłowych sztucznych sieci neuronowych. Ważnym podkreślenia jest również fakt, że Autor wyciąga prawidłowe wnioski z przeprowadzonej analizy literaturowej, której efektem są klarowne zadania badawcze.

3. Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?

Przedmiotem pracy badawczej było zaproponowanie nowatorskich metod modelowania i sterowania neuronowego. Zgodnie z zaprezentowanym w pracy, obecnym stanem wiedzy, w rozważanym obszarze badawczym nie rozpatrywano dotychczas tak sformułowanych metod modelowania i sterowania. Autor zauważył i przeanalizował szereg interesujących aspektów problemu, które były pomijane lub stanowiły marginalny element w innych opracowaniach. Stąd też konieczne okazało się rozwiązanie wielu cząstkowych zadań.

Należy zaznaczyć, że do realizacji celów pracy autor użył właściwych metod przy odpowiednio sformułowanych założeniach. W szczególności, podstawowym założeniem był brak znajomości modelu rozważanego procesu. Jest to założenie w pełni uzasadnione, a jego konsekwencją jest konieczność zastosowania modeli typu "czarna skrzynka", do których zalicza się sztuczne sieci neuronowe. Kolejne założenie dotyczyło faktu związanego z dynamiką systemu uniemożliwiającego jego efektywne modelowanie z zastosowaniem klasycznych sieci perceptronowych. Uzasadnia to w pełni konieczność zastosowania rekurencyjnych sieci neuronowych, którą przyjęto w rozprawie. Ostatnie z założeń dotyczyło konieczności sterowania procesem przy określonych parametrach jakościowych oraz spełnieniu ograniczeń zmiennych wejściowych i wyjściowych. Naturalnym rozwiązaniem przy tak sformułowanych założeniach jest zastosowanie sterowania predykcyjnych, którego hybrydową wersję, bazującą na rekurencyjnych sieciach neuronowych zaproponowano w rozprawie.

4. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?

Zaproponowane metody umożliwiają rozwiązanie problemów modelowania i sterowania neuronowego, dla których istniejące w literaturze rozwiązania nie dają zadowalających rezultatów. Świadczy to jednoznacznie o osiągnięciu przez Autora celu postawionego we wstępnej części pracy, zdefiniowanego w postaci Tezy 1, 2 i 3.

Dorobek naukowy autora rozprawy obejmuje 7 publikacji, w tym 1 artykuł w czasopiśmie indeksowanym przez Thomson Reuters (15pkt wg listy MNiSW):

- Wysocki, M. Ławryńczuk, Elman Neural Network for Modelling and Predictive Control of Delayed Dynamic Systems, w: „Archives of Control Sciences”, t. 26(LXII), 2016, ss. 5–30.

Pozostałe 6 publikacji zostało opublikowane w materiałach konferencji krajowych i międzynarodowych:

- A. Wysocki, M. Ławryńczuk, An investment strategy for the Stock Exchange using neural networks, w: „2013 Federated Conference on Computer Science and Information Systems”, FedCSIS 2013, ss. 183–190.
- A. Wysocki, M. Ławryńczuk, On Choice of the Sampling Period and the Horizons in Generalized Predictive Control, w: „Recent Advances in Automation, Robotics and

Measuring Techniques”, Springer, t. 267, Heidelberg 2014, ss. 328–339.

- A. Wysocki, Skuteczne uczenie sieci neuronowej Elmana, w: „Aktualne problemy automatyki i robotyki”, Exit, Warszawa 2014, ss. 754–764.
- A. Wysocki, M. Ławrynczuk, Predictive Control of a Multivariable Neutralisation Process Using Elman Neural Networks, w: „Advances in Intelligent Systems and Computing”, Springer, t. 350, Heidelberg 2015, ss. 335–344.
- A. Wysocki, M. Ławrynczuk, Jordan neural network for modelling and predictive control of dynamic systems, w: „2015 20th International Conference on Methods and Models in Automation and Robotics (MMAR)”, ss. 145-150.
- A. Wysocki, M. Ławrynczuk, Two- and Three-Layer Recurrent Elman Neural Networks as Models of Dynamic Processes, w: „Challenges in Automation, Robotics and Measurement Techniques”, Springer, t. 440, Heidelberg 2016, ss. 165–175.

Do najważniejszych osiągnięć rozprawy, stanowiących samodzielny i oryginalny dorobek autora należy zaliczyć:

- Opracowanie efektywnych struktur i metod uczenia rekurencyjnych sieci neuronowych o minimalnej liczbie parametrów wewnętrznych;
- Opracowanie struktur i metod uczenia rekurencyjnych sieci neuronowych dla procesów z opóźnieniami;
- Opracowanie strategii sterowania predykcyjnego bazującej na rekurencyjnych sieciach neuronowych;
- Implementacja i symulacyjna weryfikacja opracowanych technik modelowania i sterowania neuronowego dla reaktora neutralizacji oraz reaktora polimeryzacji.

Powyższe osiągnięcia, jak również fakt ich opublikowania w postaci wyżej wymienionych pozycji literaturowych jednoznacznie potwierdzają ich znaczenie dla rozwoju dziedziny modelowania i sterowania neuronowego.

5. Czy autor wykazał umiejętność poprawnego i przekonującego przedstawienia uzyskanych przez siebie wyników /zwięzłość, jasność, poprawność redakcyjna rozprawy/?

Praca przygotowana jest bardzo starannie pod względem edytorskim. Proponowane rozwiązania w zakresie sterowania i modelowania neuronowego przedstawiane są z zastosowaniem odpowiedniego aparatu matematycznego wspomaganego przez liczne ilustracje i wyniki symulacji komputerowych. Świadczy to jednoznacznie o wysokich umiejętnościach autora w zakresie poprawnego i przekonującego przedstawiania wyników. Uwzględniając wymienione osiągnięcia naukowo-badawcze oraz fakt ich opublikowania w wyżej wymienionych pozycjach, uważam że mgr inż. Antoni Wysocki zrealizował cel rozprawy, udowodnił postawione tezy oraz wykazał się wiedzą i umiejętnością samodzielnego rozwiązywania trudnych problemów teoretycznych i praktycznych szeroko rozumianej współczesnej automatyki i robotyki w zakresie modelowania i sterowania neuronowego.

6. Jakie są słabe strony rozprawy i jej główne wady?

Całą rozprawę oceniam bardzo pozytywnie, jednak niektóre spostrzeżenia i wnioski Autora wymagają dalszych wyjaśnień. W szczególności:

1) W rozdziale 2 Autor stwierdza, że: "*W przypadku wszystkich sieci neuronowych rozważa się proces wielowymiarowy o n_u wejściach i jednym wyjściu. Chcąc uzyskać model procesu o n_u wejściach i n_y wyjściach należy zastosować n_y takich sieci.*" Powyższe stwierdzenie wydaje się dość zaskakujące w kontekście bogactwa sieci neuronowych o wielu wejściach i wielu wyjściach dostępnych w literaturze. Wydaje się również, że powyższe stwierdzenie będzie prawdziwe jedynie dla procesów, w których zachodzi silna separowalność wyjść.

2) Proces reaktora neutralizacji opisany równaniami (4.1)-(4.3) posiada liniowe równania stanu i nieliniowe równania wyjścia. Wydaje się, że posiadając taką wiedzę można by zaprojektować rekurencyjną sieć neuronową z liniową częścią dynamiczną i nieliniową częścią statyczną. Czy takie rozwiązanie nie zmniejszyłoby liczby parametrów modelu i nie przyspieszyłoby procesu uczenia?

3) W rozdziale 4 Autor zdawkowo stwierdza, że "Każdy ze zbiorów danych powinien być liczny oraz reprezentatywny, by nauczona sieć neuronowa mogła właściwie modelować każda możliwą sytuację powstałą w rzeczywistym procesie (tzn. dla różnych punktów pracy)." Wydaje się, że w tym przypadku można było przeprowadzić bardziej szczegółową analizę problemu rozpoczynając od prac dotyczących planowania eksperymentu dla sztucznych sieci neuronowych, np.:

- "Optimal training strategies for locally recurrent neural networks." *Journal of Artificial Intelligence and Soft Computing Research* 1 (2011).
- "Toward the training of feed-forward neural networks with the D-optimum input sequence." *IEEE Transactions on neural networks* 17.2 (2006): 357-373.

4) Zastosowanie modelu liniowego (4.4) dla celów porównawczych nie wydaje się być najlepszym rozwiązaniem. Czy nie prościej byłoby (mając ogólną postać (4.1)-(4.3)) dokonać identyfikacji modelu opisanego w przestrzeni stanów, z zastosowaniem np. Matlab System Identification Toolbox?

7. Jaka jest przydatność rozprawy dla nauk technicznych?

Osiągnięcia przedstawione w rozprawie są bardzo istotne dla rozwoju nauk technicznych w kontekście teorii i zastosowań rekurencyjnych sieci neuronowych. Szczególnie wysoko oceniam opracowane algorytmy sterowania predykcyjnego bazujące na rekurencyjnych sieciach neuronowych. Mogą one mieć szerokie zastosowanie w układach diagnostyki i sterowania zaawansowanymi procesami technologicznymi. Mają również stymulujący wpływ na dalszy rozwój teorii w tym zakresie, szczególnie w zakresie analizy zbieżności i stabilności układów tego typu.

8. Do której z następujących kategorii Recenzent zalicza rozprawę:

- a/ nie spełniająca wymagań stawianych rozprawom doktorskim przez obowiązujące przepisy
- b/ wymagająca wprowadzenia poprawek i ponownego recenzowania
- c/ spełniająca wymagania
- d/ spełniająca wymagania z wyraźnym nadmiarem
- e/ wybitnie dobra, zasługująca na wyróżnienie

Reasumując, podniesione wyżej uwagi krytyczne i komentarze nie wpływają jednak na wysoką ocenę oryginalnych i opublikowanych osiągnięć naukowo-badawczych, zasadniczych wyników zawartych w recenzowanej pracy oraz jej ogólną pozytywną ocenę. Przedstawione wyniki stanowią niewątpliwie rozwiązanie znaczącego problemu naukowego. Stwierdzam zatem, że przedstawiona przez magistra inżyniera Antoniego Wysockiego rozprawa spełnia warunki określone w artykule 13 ust. 1 i ust. 2 Ustawy o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki z dnia 14 marca 2003 roku i wnioskuje o dopuszczenie go do publicznej obrony.

