

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FAKULTA STROJNÍHO INŽENÝRSTVÍ

ÚSTAV MECHANIKY TĚLES

Ing. Jiří Krejsa

Simulační modelování technických objektů pomocí neuronových sítí

Simulation Modelling of Technical Objects Using Artificial Neural Networks

PhD Thesis

vědní obor

INŽENÝRSKÁ MECHANIKA

Školitel: RNDr. Ing. Tomáš Březina, CSc.

Oponenti: Prof. Drhab. Inž. Waldemar Oleksiuk

Prof. Ing. Zdeněk Krupka, CSc.

Doc. Ing. Čestmír Ondrušek, CSc.

TU Warszaw

VA Brno

FEI VUT Brno

Datum obhajoby: 27. 11. 2000

© 2001 J. Krejsa
ISBN 80-214-1795-1

OBSAH

1. ÚVOD	6
2. FORMULACE PROBLÉMU A CÍLŮ JEHO ŘEŠENÍ	6
3. ANALÝZA SOUČASNÉHO STAVU ŘEŠENÉ PROBLEMATIKY ..	7
4. PREZENTACE A ANALÝZA VÝSLEDKŮ	8
3.1. Modelování soustav pomocí neuronových sítí	8
4.1 Volba sledovaných parametrů	9
4.2 Modelovaná soustava.....	9
4.3 Použité neuronové sítě	10
4.4 Implementační detaily.....	11
4.5 Tréninkové a testovací výsledky	11
4.6 Automatické navrhování sítí.....	16
4.7 Analýza počátečního náhodného nastavení parametrů sítě.....	18
4.8. Doporučení při nasazení neuronových sítí pro modelování dynamických soustav	20
4.9. Ověření doporučujících zásad.....	22
5. ZÁVĚR	24
LITERATURA	26
SEZNAM PUBLIKACÍ AUTORA	27
SUMMARY	30
ŽIVOTOPIS.....	31

1. ÚVOD

Disertační práce je zaměřena na použití umělých neuronových sítí pro úlohu simulačního modelování dynamických soustav. Neuronové sítě se díky výraznému rozvoji v posledních letech jeví jako vhodná metoda pro tuto úlohu díky některým výhodným vlastnostem, především nepotřebnosti matematického modelu. Existují ovšem i nevýhodné vlastnosti, zvláště obtížná verifikace výsledků, náročnost procesu učení a přítomnost experta, který na základě zkušeností určí, jaký druh sítě použít. Předložená práce analyzuje vliv některých parametrů podstatných při nasazení neuronových sítí na zadanou úlohu a uvádí doporučující zásady, které by měly pomoci jejich použití.

Práce obsahuje teoretický přehled použitých paradigmat neuronových sítí a soubor parametrů podstatných pro chování sítí u dané úlohy, který je dále zkoumán. Použita jsou klasická i netradiční paradigmata dopředných i rekurentních sítí (různé varianty Back-propagation algoritmu, kaskádová korelace, sítě s radiálními bázemi, Elmanův model, Jordanův model, rekurentní kaskádová korelace, rekurentní sítě s radiálními bázemi, a další). Na datech popisujících chování jednoduché soustavy je vliv vybraných parametrů analyzován z různých hledisek, především přesnosti, generalizační schopnosti, výpočetní náročnosti.

Jako podstatný se projevil vliv počátečního náhodného nastavení vah a prahů sítě, který je zpracován pomocí statistických metod na souboru 400 sítí shodné třídy se stejnou topologií a nastavením všech parametrů kromě počátečního nastavení vah a to pro reprezentanta dopředných i rekurentních sítí. V práci je také diskutována možnost automatického generování sítí pomocí genetického programování, které se ukazuje jako slibná metoda do budoucna, která je ovšem doposud neúměrně výpočetně náročná.

Doporučující zásady pro nasazení neuronových sítí v obdobných úlohách jsou hlavním závěrem a přínosem celé práce a jsou ověřeny na složitějších soustavách - soustavě s hysterezní tuhostí a soustavě typu MIMO, pro které sítě navržené dle doporučujících zásad prokázaly dobrou shodu s požadovanými výsledky.

2. FORMULACE PROBLÉMU A CÍLŮ JEHO ŘEŠENÍ

Cílem práce je analyzovat použití umělých neuronových sítí pro simulační modelování technických objektů s výrazným dynamickým chováním, a to s využitím dopředných i rekurentních sítí s klasickými (kompletní vrstvené) i obecnými topologiemi, nalézt veličiny podstatné pro toto použití a posoudit jejich vliv, případně nalézt optimální metodu z hlediska použitelnosti a přesnosti řešení a také z hlediska výpočetní náročnosti.

Dílčí cíle lze formulovat v následujících bodech:

1. Analyzovat současný stav poznání o použití neuronových sítí pro simulační modelování technických objektů.
2. Navrhnout množinu veličin podstatně ovlivňujících použití neuronových sítí z hlediska jejich adaptačních a generalizačních vlastností.
3. Vybrat a implementovat perspektivní paradigmat neuronových sítí. Uvážit i rekurentní topologie, nestandardní topologie a evoluční neuronové sítě. Zpracovat interface mezi implementacemi paradigmat a simulačním software.
4. Na úlohách simulačního modelování detailně porovnat jejich adaptační a generalizační vlastnosti a analyzovat vliv veličin ovlivňujících tyto vlastnosti, včetně výpočetní složitosti nutné k dosažení vyhovující úrovně generalizace.
5. Získané zkušenosti vyhodnotit a navrhnout metodiku optimálního nasazení paradigmat.

3. ANALÝZA SOUČASNÉHO STAVU ŘEŠENÉ PROBLEMATIKY

V úlohách technické praxe tvoří významnou část soustavy s výrazným dynamickým chováním. Běžné techniky modelování se zabývají především takovými soustavami, které lze definovat lineárními vztahy. Pro jejich modelování se používá ověřených technik založených na lineární algebře, teorii komplexních čísel a teorii obyčejných lineárních diferenciálních rovnic.

V současné době ale nejsou k dispozici modelovací techniky pro velkou množinu nelineárních časově proměnných soustav, které by vyhovovaly požadavkům na stabilitu, robustnost a na dobrou dynamickou odezvu. Jednou z možných modelovacích technik je použití umělých neuronových sítí. Neuronové sítě jsou obsáhlou třídou metod umělé inteligence, jejíž vznik se datuje okolo roku 1950, kdy se objevují první práce na téma umělých neuronových sítí. Skutečný rozvoj nastává v roce 1986 přelomovým článkem D.E. Rumelharta, který představil v současné době velmi užívanou metodu zpětného šíření.

Neuronové sítě jsou používány v různých oblastech, pro svoji rychlost se v technické praxi osvědčují především v oblasti modelování, simulace a řízení. Návrh každého řídicího systému je založen na modelu řízené soustavy. Neuronové modely řízených soustav jsou používány jako základ pro vytvoření regulátoru především v případech, kdy se řízená soustava chová jako časově spojitý nelineární dynamický systém, nebo má pouze částečně známou, případně zcela neznámou strukturu a/nebo parametry.

Dobře navržená a natrénovaná neuronová síť má celou řadu kladných vlastností, především robustnost a poměrně dobrou odolnost proti šumu, dále pak velmi krátkou dobu odezvy natrénované sítě. Hlavní výhodou použití neuronových sítí je ale to, že odstraňují nutnost vytvoření matematického modelu dané soustavy. V

technické praxi se často vyskytují velmi složité soustavy, pro které je vytvoření matematického modelu obtížné. Téměř vždy je ale možné určit parametry soustavy nebo její odezvu zjistit experimentálně. Neuronová síť může potom na základě takto získaných údajů vytvořit model daného zařízení, případně může být těchto údajů použito přímo při konstrukci neuroregulátoru.

Obecně má použití neuronových sítí také jisté nevýhody. Jsou to jednak dlouhé výpočtové časy při tréninku sítě, zvláště u rozsáhlejších sítí, které jsou pro složitější systémy nezbytné. Velmi obtížně řešitelným problémem je principiální nemožnost verifikace výsledků sítě. Širšímu nasazení pro reálné úlohy technické praxe brání také nutnost přítomnosti experta na neuronové sítě při řešení takové úlohy, což není vždy možné. Expert určuje nejen trénovací algoritmus a typ sítě, ale například topologii, parametry tréninku a další detaily.

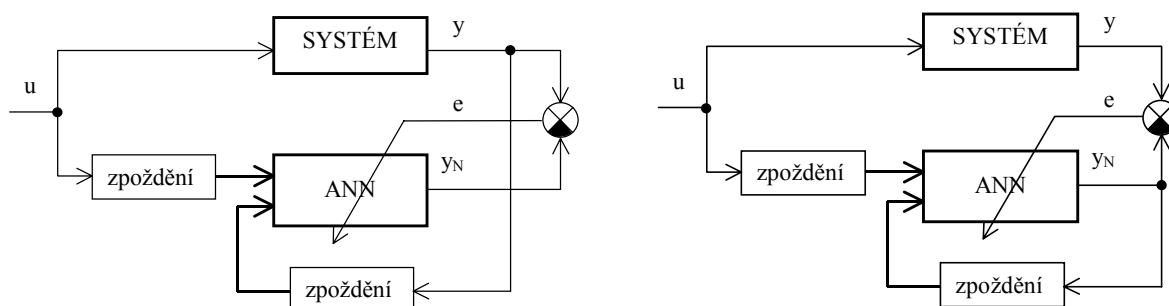
Pro použitelnost neuronových sítí v dané problematice jsou rozhodující její adaptační a generalizační vlastnosti. Ty jsou ovlivňovány celou řadou faktorů. V literatuře lze nalézt řadu prací srovnávajících vliv úzce vymezené množiny parametrů sítí na jejich užité vlastnosti. Tyto práce se ale většinou zabývají pouze dvěma algoritmy učení a jsou často účelové, snažící se dokázat „sílu“ jednoho algoritmu na úkor druhého, a není provedeno komplexní zhodnocení všech podstatných faktorů, které konečný výkon sítí ovlivňují. Cílem disertační práce je zhodnotit vliv jednotlivých faktorů komplexně a navrhnout z hlediska použitelnosti optimální metodu konstrukce neuronové sítě pro tento typ úloh.

4. PREZENTACE A ANALÝZA VÝSLEDKŮ

3.1. Modelování soustav pomocí neuronových sítí

V této práci jsou analyzovány neuronové sítě použité pro simulační modelování nelineárních dynamických soustav. Experimentální část je zaměřena na vytvoření neuronových modelů uvedené třídy soustav v podobě její identifikace na základě získaných dat o jejím chování.

Neuronové sítě jsou vhodné pro modelování silně nelineárních soustav jak typu SISO, tak MIMO. Schéma simulované identifikace pomocí dopředné neuronové sítě (blok ANN) je na obr. 1.



Obr. 1. Sériově-paralelní identifikace

paralelní identifikace.

V sériově-paralelním případě je vedena zpětná vazba ze systému, v paralelním případě z neuronové sítě. Při identifikaci BIBO soustavy je sériově-paralelní identifikace výhodnější, protože všechny signály použité pro identifikaci jsou omezené. Naproti tomu u paralelní identifikace není zaručena (díky zpětné vazbě z modelu) konvergence a stabilita [12]. U experimentů byla použita sériově-paralelní identifikace ve fázi tréninku sítě. U rekurentních sítí nebyla použita zpoždění, jelikož rekurentní síť má vnitřní dynamiku díky vnitřní zpětné vazbě.

4.1 Volba sledovaných parametrů

Cílem práce je analyzovat použití neuronových sítí při modelování dynamických soustav. Základním úkolem je volba parametrů, jejichž vliv bude v experimentech sledován. Na základě literatury a zkušeností byly vybrány následující parametry:

- Typ sítě (dopředná / rekurentní)
- Topologie
- Učící algoritmus
- Počáteční nastavení parametrů sítě (vah a prahů)

Parametr *typ sítě* zahrnuje jednak základní rozdělení na sítě dopředné a rekurentní a dále určitou třídu sítí. Parametr *učící algoritmus* zahrnuje pouze konkrétní použité paradigma, nikoliv jeho vnitřní nastavení.

Hodnotící kritéria

Vliv jednotlivých parametrů ovlivňuje chování sítě, které je možné sledovat z různých hledisek. Vzhledem k povaze úlohy byla uvažována následující hlediska:

- Přesnost dosažených výsledků
- Výpočetní a implementační náročnost

Přesnost dosažených výsledků se vztahuje především na přesnost sítě při testování na testovacích datech. Výpočetní náročnost zahrnuje počet operací vztažených na tréninkovou epochu, počet epoch, a paměťovou náročnost při tréninku a běhu sítě.

4.2 Modelovaná soustava

Pro testování vlivu jednotlivých parametrů u konkrétních implementací vybraných paradigmat byla zvolena soustava popsaná rovnicí (1)

$$y(k) = 0.1y(k-1) + 0.7y(k-2) + 0.8 \sin u(k) + 0.3 \sin 2u(k), \quad (1)$$

kde k je časový krok, u je vstup a y výstup soustavy. Pro generování tréninkových dat bylo použito buzení soustavy

$$u(t) = \sin(t) \cdot 1.6 \sin\left(1.8t + \frac{\pi}{3}\right). \quad (2)$$

Pro vytvoření trénovací množiny bylo použito vzorkování vstupního i výstupního signálu s délkou časového kroku $\Delta k = 0.05 \text{ sec}$.

4.3 Použité neuronové sítě

Pro analýzu různých paradigmat neuronových sítí byly vybrány následující třídy sítí. Za třídou sítě je vždy uvedena zkratka, pod kterou bude třída dále označována

A) Dopředné sítě

A1) Úplná třívrstvá síť

- Základní algoritmus Back-propagation s momentem a adaptivním krokem učení (GDX)
- Resilient Back-propagation (Rprop)
- Algoritmy konjugovaných gradientů
 - a) Fletcher-Reeves Update (CGF)
 - b) Polak-Ribiere Update (CGP)
 - c) Powell-Beale Restart (CGB)
 - d) Poměrný (Scaled) Konjugovaný Gradient (SCG)
- Kvazi-Newtonovské algoritmy
 - a) Broyden, Fletcher, Goldfarb a Shanno Algoritmus (BFGS)
 - b) One Step Secant Algoritmus (OSS)
- Levenberg-Marquardtův algoritmus (LM)
- Síť s radiální bází
 - a) Neomezený počet skrytých neuronů (RBF)
 - b) Omezený počet skrytých neuronů (RBF)

A2) Kaskádová korelace

- Neomezený počet vrstev (CC1)
- Omezený počet vrstev (CC2)

B) Rekurentní sítě

- Elmanův model (ELM)
- Jordanův model (JOR)
- Rekurentní kaskádová korelace (RCC)
- Rekurentní RBF
 - a) Rekurentní středy (RRBFC)
 - b) Lineární rekurze (RRBFL)
 - c) Adaptivní filtr (RRBFA)

Všechny použité algoritmy mají různé parametry tréninku (hodnoty různých konstant, terminální kritéria apod.), které ovšem nebyly předmětem zkoumání.

Shrnutí použitých sítí

Přehled jednotlivých tříd sítí spolu s počtem vstupů, počtem skrytých neuronů jednotlivých tříd a adaptovaných parametrů (vah a prahů) je uveden v tabulce 1.

Třída	# vstupů	# skrytých neuronů	# vah a prahů
Úplná dopředná třívrstvá síť (GDX, Rprop, CFG, CGP, CGB, SCG, BFGS, OSS, LM, RBFB)	3	3, 5, 10, 20	16, 26, 51, 101
Úplná dopředná třívrstvá síť (RBFE)	3	2000	10001
Úplná dopředná třívrstvá síť (RBFA)	3	37	186
Kaskádová korelace (CC1)	3	8	76
Kaskádová korelace (CC2)	3	3, 5	19, 34
Elmanův model (ELM)	2	3, 5, 10, 20	22, 46, 141, 481
Jordanův model (JOR)	2	3, 5, 10, 20	16, 26, 51, 101
Rekurentní kaskádová korelace (RCC)	2	7	61
Rekurentní RBF - rekurentní středy (RRBFC)	2	3, 5, 10, 20	22, 46, 141, 481
Rekurentní RBF - lineární rekurze (RRBFL)	2	3, 5, 10, 20	16, 26, 51, 101
Rekurentní RBF - adaptivní filtr (RRBFA)	2	3, 5, 10, 20	21, 35, 70, 140

Tab. 1. Počet parametrů jednotlivých typů sítí

V dalším textu jsou jednotlivé třídy označovány pouze pomocí zkratk spolu s počtem skrytých neuronů - např. rekurentní RBF s použitím adaptivního filtru a 3 neurony ve skryté vrstvě je označena RRBFA_3.

4.4 Implementační detaily

Vzhledem k rozsahu použitých paradigmat nebylo možné použít pro experimenty jednotnou platformu. Přehled použitých programových prostředků uvádí tabulka 2.

Použitý software	Třída
Matlab 5.2-5.3, Neural Network Toolbox 3.0	GDX, Rprop, CFG, CGP, CGB, SCG, BFGS, OSS, LM, RBFE, RBFA, ELM
Stuttgart Neural Network Simulator (SNNS)	GDX, RBFB, CC1
Úprava zdrojového kódu public domain	CC1, CC2, RCC
Vlastní implementace	GDX, RBFA, RBFB, CC1, CC2, JOR, RRBFC, RRBFL, RRBFA

Tabulka 2. - Přehled použitých programových prostředků

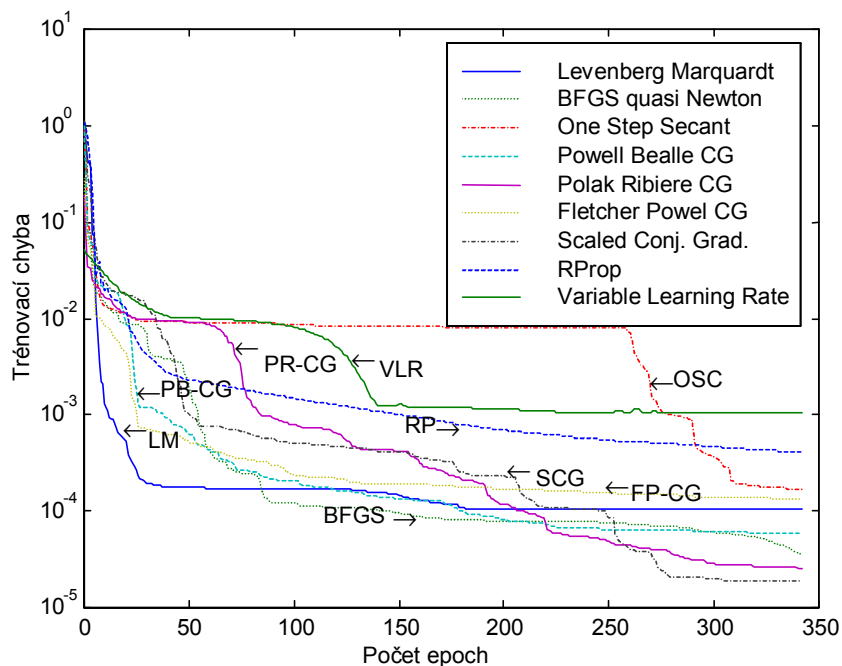
Pro všechny výpočty byly použity osobní počítače s procesorem Intel Pentium II / 400 Mhz, operačním systémem Windows NT 4.0 a operační paměť 128 a 256 MB.

4.5 Tréninkové a testovací výsledky

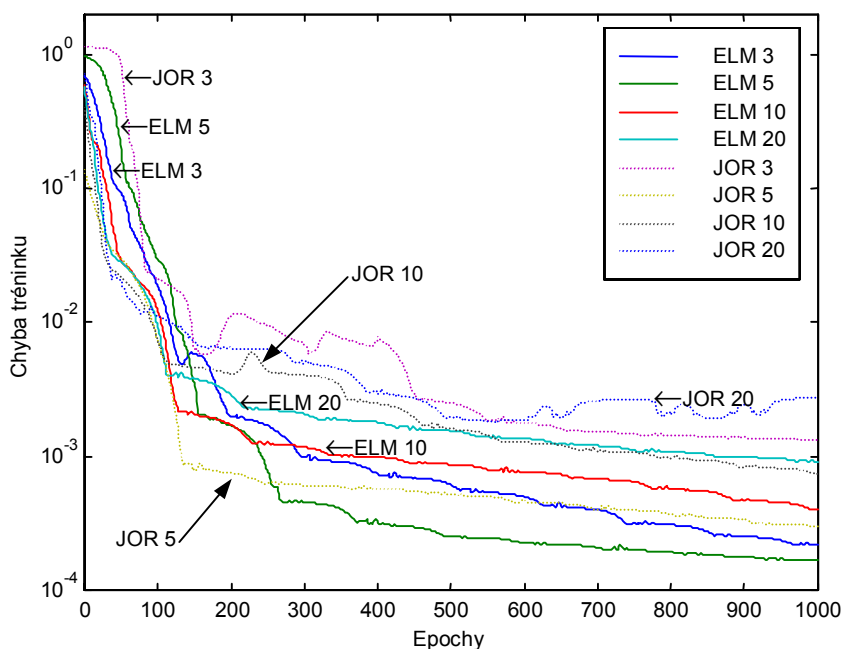
4.5.1 Trénink

Pro dopředné sítě obsahovala tréninková data tři vstupy a jeden výstup. Vstupem byla hodnota buzení v daném časovém kroku $u(k)$ a hodnota výstupu systému ve dvou předchozích časových krocích $y(k-1), y(k-2)$. Výstupem byla požadovaná hodnota výstupu systému $y(k)$. Pro rekurentní sítě byla použita pouze hodnota buzení $u(k)$ a hodnota výstupu systému v minulém časovém kroku $y(k-1)$. Pro částečnou eliminaci počátečního nastavení vah bylo trénováno vždy deset sítí stejné třídy, které se lišily počátečním nastavením vah. Jako reprezentant dané třídy byla

brána síť s druhým nejlepším tréninkovým výsledkem. Pro ilustraci průběhu chyby tréninku jsou uvedeny průběhy pro úplné dopředné třívrstvé topologie se 3 skrytými neurony (Obr. 2.) a některé průběhy pro rekurentní síť (Obr. 3.).



Obr. 2. Průběh tréninku pro dopředné síť se 3 neurony ve skryté vrstvě



Obr. 3. Průběh chyby tréninku pro Elmanovu a Jordanovu rekurentní síť.

Koncová hodnota chyby

Ve výsledcích úplných vrstvených dopředných topologií lze vysledovat zlepšující se schopnost učení se zvyšujícím se počtem skrytých neuronů. Nejhorších výsledků v

tréninku dosahoval algoritmus GDX, nízké hodnoty chyby dosáhly všechny ostatní algoritmy, nejlepších BFGS a LM algoritmus. U sítí s RBF je výsledek přesné konstrukce (RBFE) dle očekávání v tréninku dobrý, síť RBFB naopak dosahují výsledků nejhorších ze všech použitých tříd. U kaskádové korelace jsou výsledky tréninku v porovnání s úplnými dopřednými vrstvenými sítěmi poněkud horší.

U rekurentních sítí obecně je zřejmé že vliv počtu skrytých neuronů je zde nižší než u dopředných sítí, navíc síť s vysokým počtem (20) dosahují horších výsledků než síť menší. U Elmanova a Jordanova modelu jsou shodně nejlepší síť s pěti neurony ve skryté vrstvě, rekurentní síť s radiálními bázemi dosahují nejlepších výsledků při 10 neuronech ve skryté vrstvě.

Výpočetní náročnost

U úplných dopředných topologií je nejméně náročný algoritmus Rprop. Druhým nejméně náročným algoritmem je základní Back-propagation. Algoritmy konjugovaných gradientů a kvazi-Newtonovské algoritmy jsou zhruba stejně náročné, OSS algoritmus je poněkud méně náročný než algoritmus BFGS. Nejnáročnější je Levenberg-Marquardtův algoritmus. Je nutné mít ovšem v patrnosti, že složitější algoritmy současně rychleji konvergují.

Sítě s radiálními bázemi bez omezení počtu skrytých neuronů vyžadují největší množství operací ze všech použitých algoritmů, náročnost modelu RBFA je také značná. Naopak RBF síť s předem určeným počtem neuronů jsou řádově nejméně náročné ze všech uvedených algoritmů. Dopředná kaskádová korelace je průměrně náročná. Vzhledem k principu tvorby sítě nemůže být u modelů kaskádové korelace a RBFE a RBFA uveden počet operací na parametr a váhu.

Rekurentní sítě jsou překvapivě méně náročné na počet operací oproti sítím dopředným. To je způsobeno jednak menším počtem vstupů sítě a dále tím, že u rekurentních sítí nejsou používány relativně složité algoritmy typu konjugovaných gradientů nebo kvazi-Newtonovské algoritmy, ale většinou modifikace základního gradientního algoritmu.

4.5.2. Testování

Pro generování testovacích dat bylo použito buzení soustavy podle rovnice (3).

$$u(t) = \left(4 \sin \left(t + \frac{\pi}{3} \right) \right) (0.7 \sin(0.2t + 7)) (2 \sin(0.12t + 4)) (\text{saw}(0.7t)). \quad (3)$$

Pro testovací množinu bylo stejně jako pro trénovací množinu použito vzorkování vstupního i výstupního signálu s délkou časového kroku $\Delta k = 0.05 \text{ sec}$.

Pro vyjádření kvality chování sítě na testovacích datech byly vytvořeny dva parametry. *Maximální testovací chyba* je absolutní hodnota největšího rozdílu mezi výstupem sítě a požadovanou hodnotou vztažená k maximální hodnotě testovacích dat v celém intervalu testování. *Průměrná testovací chyba* je vypočtena jako suma absolutních hodnot chyby v každém kroku, podělena počtem kroků a dále dělená

maximální hodnotou výstupu a násobená stem. Průměrná testovací chyba pro všechny kombinace struktur sítí a učících algoritmů je zobrazena v tabulce Tab. 3.

Nejlepší hodnoty při testování maximální testovací chyby dosáhla úplná dopředná síť s 20 neurony na skryté vrstvě, trénovaná algoritmem Levenberg-Marquardt (LM). Její maximální chyba činí pouhých 3.7 %. Jako druhá nejlepší se ukázala síť stejného typu s pouze 5 neurony na skryté vrstvě, trénovaná algoritmem konjugovaných gradientů s Polak-Ribierovým update (CGP). Sítě s radiálními bázemi mají chybu vysokou, kaskádová korelace dosahuje průměrných výsledků, chyby u sítí s větším počtem neuronů jsou poměrně dobré. Rekurentní sítě vykazují u maximální testovací chyby horší výsledky než sítě dopředné, u sítí s menším počtem neuronů ve skryté vrstvě jsou ale kompaktnější.

Počet skrytých neuronů	3 SN	5 SN	10 SN	20 SN	Spec
Třída					
GDX	43.642	1.349	1.386	2.435	-
Rprop	6.728	31.710	1.387	1.315	-
CGF	8.644	0.718	3.142	1.144	-
CGP	4.936	0.308	2.087	1.149	-
CGB	30.890	37.539	0.395	0.852	-
SCG	1.002	0.744	0.864	1.024	-
BFGS	40.271	2.258	0.623	0.293	-
OSS	0.651	1.072	3.482	0.354	-
LM	4.585	36.995	0.291	0.078	-
RBFE	-	-	-	-	7.201
RBFA	-	-	-	-	8.014
RBFB	35.010	21.414	7.410	6.133	-
CC	8.914	4.214	-	-	2.054
RCC	-	-	-	-	5.612
ELM	7.191	5.894	27.828	34.210	-
JOR	8.471	6.412	8.011	37.088	-
RRBFC	8.814	6.133	5.447	26.121	-
RRBFL	8.135	4.031	3.800	31.021	-
RRBFA	13.422	7.813	3.182	19.203	-

Tab. 3. Průměrná testovací chyba [%]

U hodnoty průměrné testovací chyby je výsledek nad 10 % možné považovat za nepřijatelný, za akceptovatelné sítě s chybou kolem 5 % a za vynikající sítě s chybou pod jedno procento. U dopředných sítí dosáhla nejlepší hodnoty dopředná vrstvená síť s 20 skrytými neurony trénovaná algoritmem Levenberg-Marquardt - 0.078 %. Chyby pod 1 % dosáhla celá řada sítí s dopřednou topologií, u algoritmu One Step Secant (OSS) dosáhla této chyby síť s 3 neurony ve skryté vrstvě - 0.651 %.

Nejstabilnější z dopředných sítí je třída trénovaná algoritmem poměrných konjugovaných gradientů (SCG), tyto sítě dosáhly průměrné testovací chyby kolem

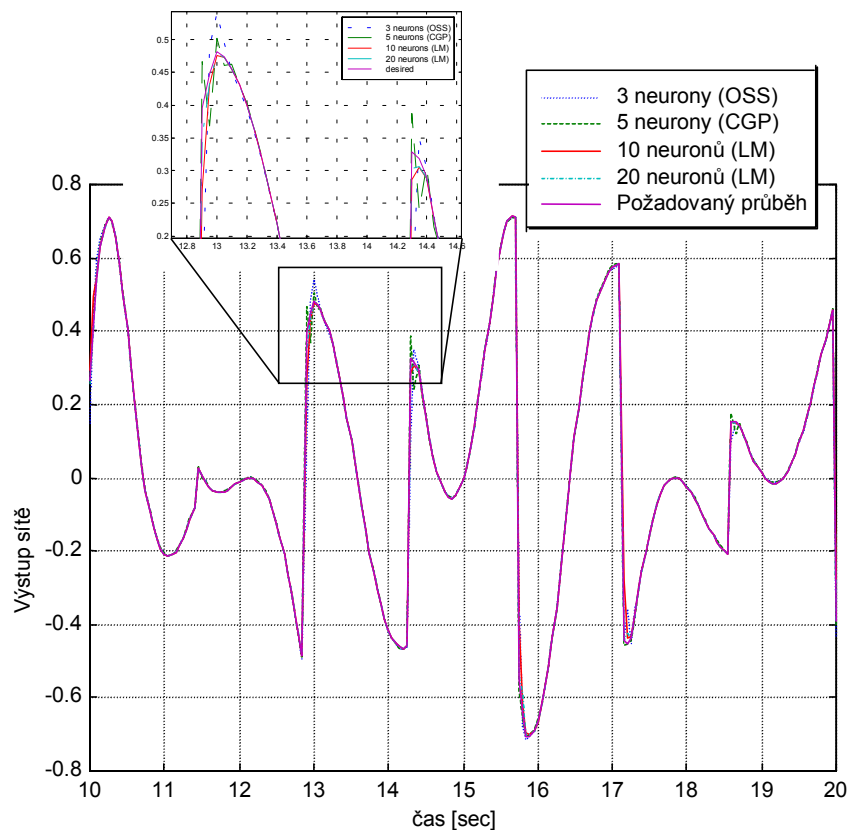
jednoho procenta pro všechny variace. Obdobně stabilní jsou i chyby sítí trénovaných algoritmem One Step Secant (OSS) a pro sítě s 5 a více skrytými neurony také třídy trénované algoritmy konjugovaných gradientů.

Výsledky rekurentních sítí z hlediska průměrné testovací chyby dobře korelují s hodnotami maximální testovací chyby. Rekurentní sítě dosahují horších výsledků než sítě dopředné, většina nejlepších výsledků rekurentních sítí je na hranici akceptovatelnosti. Pro rekurentní sítě již opět platí pravidlo známe z literatury - příliš velký počet neuronů ve skryté vrstvě zhoršuje generalizaci sítí.

Dopředné sítě

Na obr. 4. jsou uvedeny průběhy nejlepších dopředných vrstvených sítí trénovaných různými algoritmy. Kromě požadovaného průběhu je uveden vždy nejlepší zástupce sítě pro určitý počet neuronů ve skryté vrstvě. Jsou to:

- 3 neurony ve skryté vrstvě - One Step Secant algoritmus (OSS)
- 5 neuronů ve skryté vrstvě - algoritmus konjugovaných gradientů - Polak-Ribiere update (CGP)
- 10 a 20 neuronů ve skryté vrstvě - Levenberg-Marquardtův algoritmus (LM)



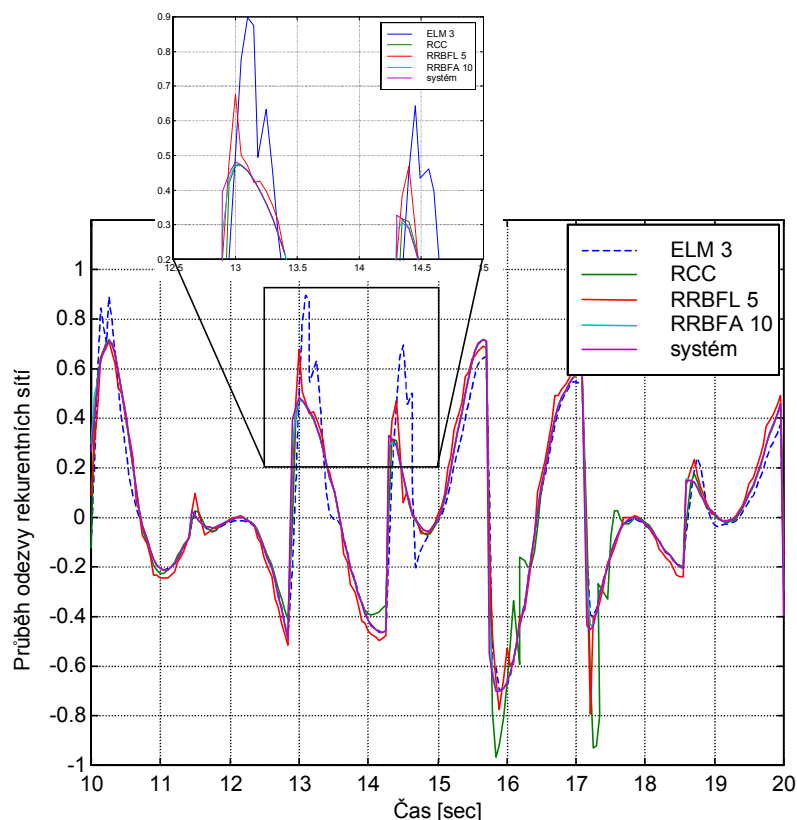
Obr. 4. Srovnání odezvy nejlepších dopředných sítí s požadovaným průběhem

Rekurentní sítě

Obdobně jako u dopředných sítí jsou na Obr. 5. uvedeny průběhy nejlepších rekurentních sítí s různým počtem skrytých neuronů ve srovnání s požadovaným výstupem modelovaného systému. Vzhledem k tomu že všechny rekurentní sítě s

20-ti neuronů ve skryté vrstvě dosáhly průměrné testovací chyby 19 - 37 %, není v grafu uvedena žádná z těchto sítí, místo ní je zobrazen průběh odezvy rekurentní kaskádové korelace. Obr. 5. tak kromě požadovaného výstupu obsahuje odezvy těchto rekurentních sítí:

- Elmanův model se 3 skrytými neurony (ELM_3; 7,19 %)
- Rekurentní RBF s lineární rekurzí a 5 skrytými neurony (RRBFL_5; 4,22 %)
- Rekurentní RBF s filtrem a 10 skrytými neurony (RRBFA_10; 3,13 %)
- Rekurentní kaskádová korelace (RCC; 5,62 %)



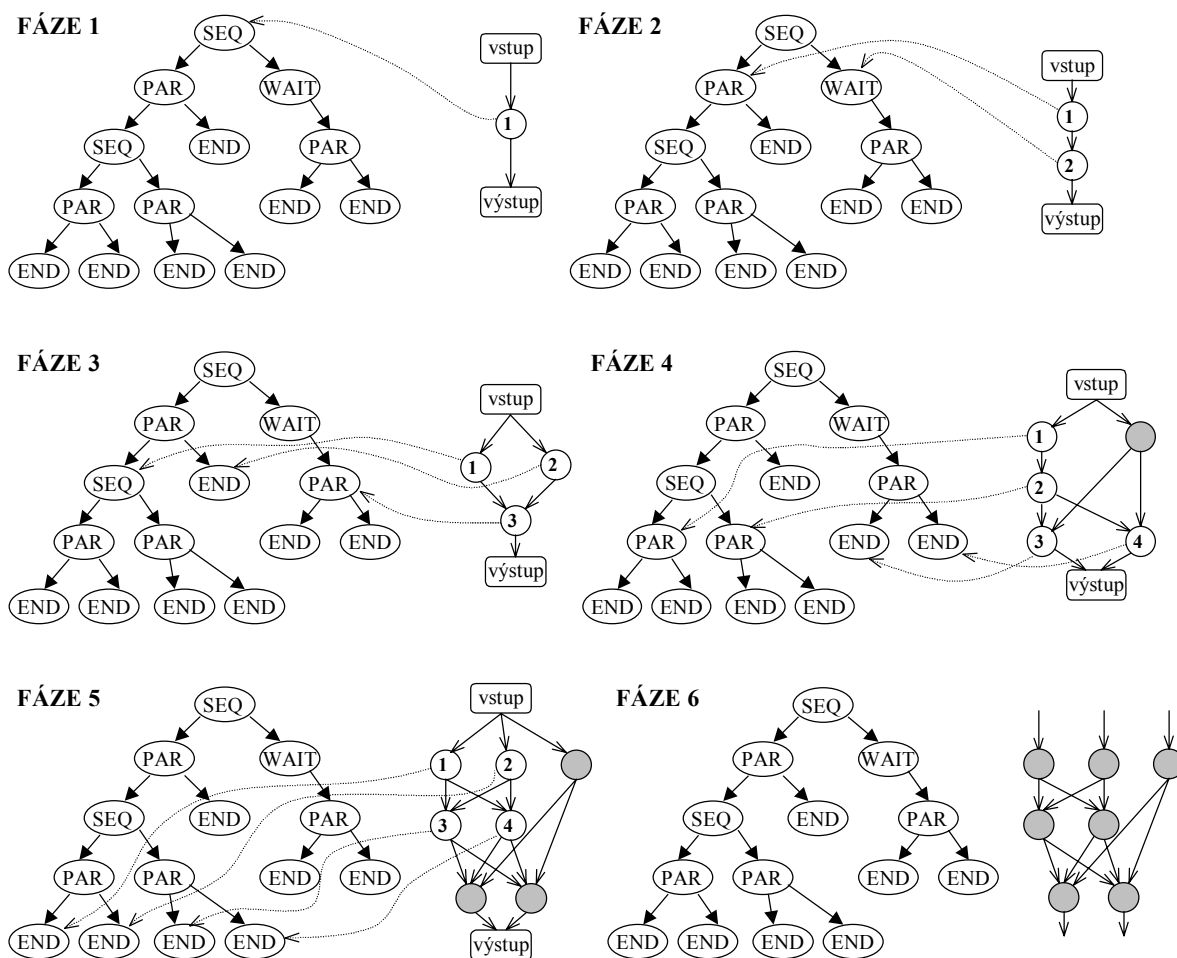
Obr. 5. Srovnání odezvy nejlepších rekurentních sítí s požadovaným průběhem

4.6 Automatické navrhování sítí

Jedním z problémů při konstrukci neuronové sítě je odhad její struktury. Ideální je metoda, kdy je celá síť navržena obecně, automaticky, s topologií optimální pro řešenou úlohu. Možným řešením je metoda založená na genetickém programování. Podkladem je strojová generace topologie sítí, založená na genetickém algoritmu. Základním problémem úlohy je zakódování topologie sítí do chromozomu.

Metoda používá celulárního kódování. Pravidel gramatiky je užito na uzly orientovaného grafu s uspořádanými hranami - buňky. Každá buňka může číst celulární kód a v závislosti na přečteném symbolu se může dělit, měnit svoje vnitřní parametry a nakonec se stane neuronem. Nově vzniklé buňky opět mohou číst symboly, které tak představují instrukce pro vývoj topologie. Buňka se rozvíjí podle dalších symbolů tak dlouho, dokud její vývoj není ukončen. Příklad zpracování je na

obrázku 6. Na levé straně je uveden gramatický strom popisující síť, napravo aktuální stav sítě.



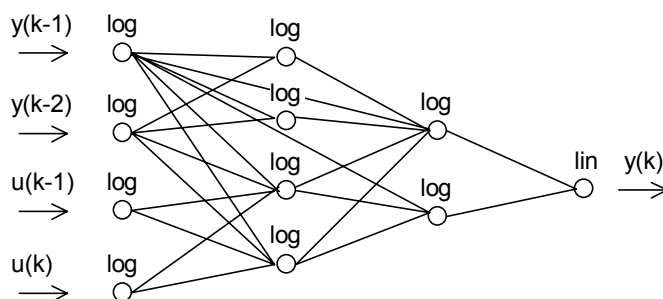
Obr. 6. Schéma vývoje sítě (použití symbolů SEQ, PAR, END, WAIT)

Pro vývoj topologie sítě je podstatných následujících pět symbolů:

- symbol sériového dělení SEQ vytváří dvě buňky z jedné (výstupní arita 2). První takto vzniklá buňka dědí vstupní hrany, druhá výstupní hrany původní buňky.
- symbol paralelního dělení PAR vytváří opět dvě buňky z jedné (arita rovna 2). Obě vzniklé buňky ale dědí jak vstupní, tak výstupní propojení původní buňky.
- symbol prořezání CUT odstraní z generované sítě předepsanou hranu.
- symbol čekání WAIT pouze připravuje buňku na další zpracování.
- ukončovací symbol END znamená, že se buňka stává neuronem.

Gramatické stromy jsou metodou chápány jako chromozomy. Na uvedených strukturách jsou definovány běžné genetické operátory. Jakmile je provedeno křížení a mutace, je podle získaného chromozomu vyvinuta příslušná neuronová síť. Tato síť je dále trénována obvyklým způsobem gradientním algoritmem. Trénink je omezen pouze na několik desítek iterací. Struktura jedné ze sítí navržených uvedeným algoritmem pro danou úlohu je zobrazena na obr. 7. Při testování dosáhla

tato síť maximální testovací chyby 25,126 % a průměrné testovací chyby 8,471 %. Tyto hodnoty nejsou v porovnání s výsledky ostatních tříd sítí výrazně horší ani lepší. Celkový počet parametrů výsledné sítě je $(22+7 = 29)$, což zhruba odpovídá dopředné síti s úplnou vrstvenou topologií s 5 neurony ve skryté vrstvě. Výpočetní náročnost je ovšem mnohonásobně vyšší. Při uvážení těchto skutečností a také implementační náročnosti popsaného algoritmu je zřejmé, že nevýhody prozatím převažují. Do budoucna se ovšem metoda jeví jako slibná především pro snadnou paralelizovatelnost, která může přinést podstatné zvýšení rychlosti a tím odstranění největšího záporu metody.



Obr. 7. Struktura automaticky navržené sítě

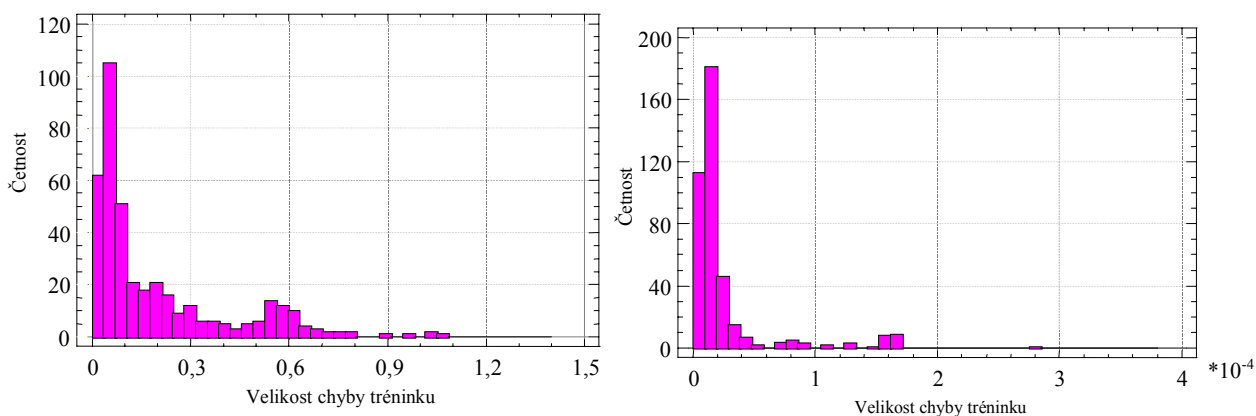
4.7 Analýza počátečního náhodného nastavení parametrů sítě

Jedním z faktorů, který významně ovlivňuje průběh tréninku i celkové chování neuronových sítí, je počáteční nastavení adaptovaných parametrů sítě - vah a prahů. Vzhledem k množství tříd sítí, které bylo v této práci použito byly vybrány pouze dvě třídy: jedna (SCG s topologií 3-5-1) jako reprezentant dopředných sítí a druhá (Elmanův model 2-5-1) jako reprezentant sítí rekurentních. Cílem analýzy bylo určit jak ovlivňuje počáteční nastavení parametrů sítě průběh tréninku a chování sítě a zda bylo oprávněné použití druhé nejlepší z deseti sítí se stejnými parametry.

Obě sítě byly trénovány na stejných datech, se stejnými parametry algoritmu učení a terminálními kritérii. Velikost statistického souboru byla zvolena o hodnotě 400 realizací. Na souboru získaných dat byla provedena základní statistická analýza. Průběh směrodatné odchylky a rozpětí pro model SCG_5 ukázal, že v průběhu tréninku se hodnoty chyby sítě k sobě přibližují, což je možné interpretovat tak, že většina sítí spadá do stejného lokálního extrému chybové plochy. U Elmanova modelu směrodatná odchylka i rozpětí po počátečním prudkém poklesu zůstávají prakticky konstantní, nedochází tedy k výraznému přeskupování jednotlivých realizací.

Histogramy rozložení chyby v iteracích - SCG

Bližší rozbor je možný z histogramů jednotlivých řezů statistického souboru - histogramu rozložení chyby v určité iteraci (epoše). Histogramy pro první epochu (netrénované sítě) a pro závěrečnou epochu tréninku jsou uvedeny na obrázku 8. Pro histogramy bylo použito 40 tříd, pro hustotu šířka intervalu 15 %.



Obr. 8. Histogramy chyby tréninku pro epochy 1 a 1000 (SCG)

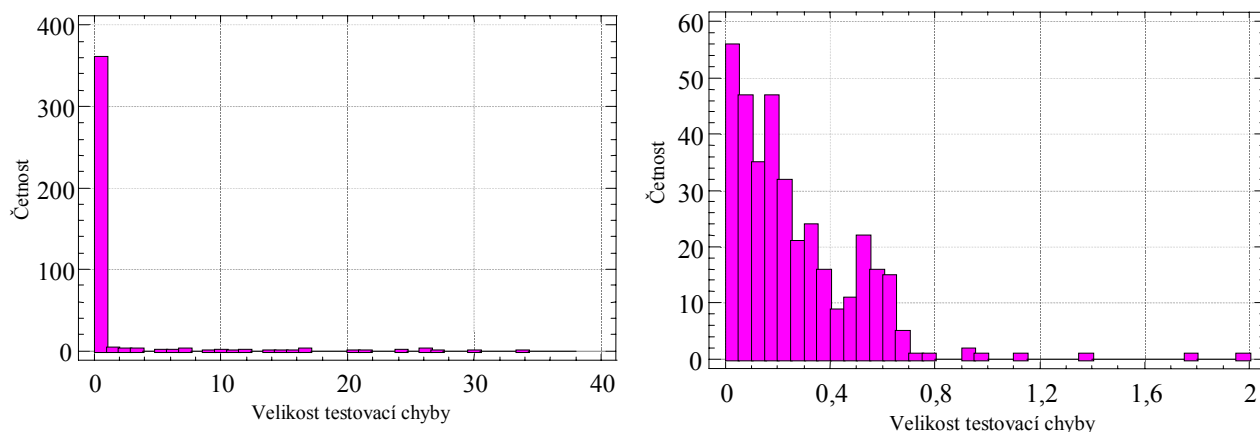
Ještě před tréninkem jsou body reprezentované množinou vah a prahů na chybové ploše rozmístěny v několika skupinách. Jde o vícemodální rozdělení pravděpodobnosti chyby. Je zřejmé že i po ukončení tréninku lze síť rozdělit do několika skupin charakterizovaných modalitami rozdělení pravděpodobnosti.

Vývoj chyby během tréninku

Předmětem dalšího zkoumání byl popis vývoje chyby během tréninku (vztahu mezi rozložením jednotlivých iterací). Jako popisující veličina byl vybrán podíl iterací. Bylo zjištěno že podíl iterací pro síť SCG vykazuje lognormální rozdělení. Toto rozdělení bylo testováno pomocí Kolmogorovova testu a testu na Chí-kvadrát, které lognormální rozdělení potvrdily. Pro Elmanův model se nepodařilo najít transformace rozdělení pravděpodobnosti iterací na standardní rozdělení.

Testování

Pro síť SCG byla zpracována statistická analýza také pro výsledky testování sítě. Vzhledem k tomu, že skupina sítí dosáhla výsledku testovací chyby v rozmezí 3-40, je v histogramu i při počtu 40 tříd většina sítí soustředěna v jedné oblasti. Obr. 9. proto uvádí histogram s omezenou horní hranicí na 2,0 (odpovídající zprava useknutému rozdělení pravděpodobnosti chyby).



Obr. 9. Histogramy testovací chyby (celkový a s horním omezením) - SCG

Z histogramů je zřejmé že při testování je většina sítí soustředěna do oblasti chyby 0,0-0,4; další výrazná skupina je v oblasti chyby 0,4-0,7.

Vztah mezi testovací chybou a tréninkem sítí byl vyjádřen pomocí koeficientů korelace pro testování a poslední tréninkovou epochu. Obdobně jako u analýzy tréninku byl nejprve použit Pearsonův koeficient korelace, ale vzhledem k asymetrii všech realizací byl aplikován také Spearmanův koeficient pořadové korelace. Pearsonův koeficient korelace činil 0,8970, Spearmanův koeficient 0,6296. Testování a trénink jsou tedy významně závislé.

Ověření metodiky použité pro trénink sítí

Při experimentech byla použita vždy druhá nejlepší síť z deseti realizací jedné třídy jako její reprezentant, aby byl omezen vliv počátečního nastavení parametrů. V práci byl tento předpoklad ověřen, tedy byla určena pravděpodobnost, že druhá nejlepší z deseti realizací bude "dost dobrá" (její chyba bude nižší než určený limit). Byly určeny tabulky pravděpodobností, že alespoň i z m sítí (kde $i=1,\dots,10$ a $m=10$) dosáhne limitní hodnoty chyby. Tato limitní hodnota byla určena na základě předchozích zkušeností a to různě pro model SCG a model ELM.

Z uvedených výsledků je zřejmé, že při výběru druhé nejlepší sítě z deseti realizací lze pro danou třídu sítí určit hodnotu chyby, které síť s velkou pravděpodobností dosáhne. Konkrétní mezní hodnota chyby sítě, již lze pokládat za dostatečně malou, je závislá na třídě sítí. Otázkou zůstává, jakým způsobem určit, že konkrétní realizace sítě je "dostatečně dobrá". S obecnou platností takový pojem nelze zavést. Je možné pouze stanovit takové kritérium, že za "dobrou" síť bude považována síť s alespoň tak nízkou hodnotou chyby jako 25 % sítí při neomezeném počtu realizací. V takovém případě je pravděpodobnost, že alespoň dvě z deseti realizací budou "dobré" 0,7546. Při změně kritéria na 30 % sítí je tato pravděpodobnost již 0,861. Použití druhé nejlepší z deseti realizací je otázka volby konkrétní hodnoty kritéria. Na základě výsledků uvedených v této kapitole byl předpoklad použití druhé nejlepší sítě z deseti vyhodnocen jako oprávněný.

4.8. Doporučení při nasazení neuronových sítí pro modelování dynamických soustav

Z provedených experimentů plyne několik závěrů které jsou shrnuty v doporučeních při nasazení neuronových sítí pro modelování dynamických soustav.

Volba typu sítě

V experimentech dopředné sítě překonaly sítě rekurentní prakticky ve všech sledovaných kritériích. Dopředné sítě dosáhly obecně větší přesnosti, jsou méně náročné na implementaci i na výpočtový čas.

Počet vrstev

U dopředných sítí lze doporučit jedinou skrytou vrstvu (při uvažování vrstvených topologií). Přestože některé práce [7] ukazují dobré zkušenosti i s větším počtem skrytých vrstev, v experimentech se zlepšení chování takových sítí neprojevovalo, naopak, roste výpočetní náročnost a snížena je také stabilita sítí.

Počet neuronů skryté vrstvy

Volba tohoto parametru je pro použití sítě klíčová - ovlivňuje generalizační schopnosti sítě i celkový počet adaptovaných parametrů a tím výpočetní náročnost sítě. Provedené experimenty prokázaly u rekurentních sítí chování známé z literatury. To znamená, že při příliš nízkém počtu neuronů skryté vrstvy síť dosahuje špatných výsledků v tréninku, při zvyšování počtu skrytých neuronů se zlepšuje do jisté meze jak schopnost tréninku tak generalizační vlastnosti, po jejím překročení se zlepšují výsledky tréninku, ale zhoršují se generalizační schopnosti.

U dopředných sítí ovšem toto chování nebylo prokázáno. Naopak, v závislosti na použitém algoritmu učení se u většiny sítí generalizační schopnosti nezhoršily ani při poměrně vysokém počtu neuronů ve skryté vrstvě. Toto chování je zřejmě specifické pro danou úlohu a nelze je zobecňovat pro obecné nasazení neuronových sítí při modelování soustav s dynamickým chováním.

Obecně lze doporučit počet skrytých neuronů odpovídající nejméně dvojnásobku počtu vstupních neuronů. Řešením může být také nasazení sofistikovaných algoritmů, které počet neuronů skryté vrstvě určují víceméně automaticky - viz. dále.

Algoritmus učení

Algoritmus učení souvisí v některých případech i s použitou topologií - konstruktivistické algoritmy mají již v sobě obsažen pokus o optimalizaci topologie pro řešenou úlohu. Z těchto algoritmů dosahuje dobrých výsledků kaskádová korelace. Její nevýhodou je složitější implementace a tendence ke tvorbě velmi hlubokých sítí, které mohou vést k nestabilnímu chování. Sítě RBF jsou velmi rychlé a v některých variantách jsou také schopny automaticky zvolit počet skrytých neuronů. Jejich přesnost je ovšem ještě horší než u kaskádové korelace. Algoritmus uvedený ve zvláštní kapitole o automatickém generování neuronových sítí pomocí genetického programování řeší problém s volbou počtu neuronů či topologie obecně. Nicméně výpočetní náročnost této metody je dosud příliš velká. Zbývají tedy gradientní algoritmy a jejich modifikace. Z hlediska přesnosti je možné doporučit Levenberg-Marquardtův algoritmus. Tento algoritmus je ovšem bohužel velmi náročný výpočetně a to jednak počtem operací, ale především paměťovými nároky. Z hlediska přesnosti a zároveň náročnosti vycházejí z ostatních algoritmů vítězně především algoritmus poměrných konjugovaných gradientů (SCG) a metoda One Step Secant (OSS). Obě metody také prokazují relativně vysokou "nezávislost" na počtu neuronů skryté vrstvy.

Vliv počátečního nastavení parametrů sítě

Experimenty prokázaly, že počáteční nastavení parametrů má na chování sítě zásadní vliv, často větší než volba algoritmu učení. Je proto nutné síť o určité topologii trénovat několikrát pro různé počáteční nastavení vah a prahů a poté zvolit síť nejlepší. Na základě provedených výpočtů se jako rozumný kompromis mezi výpočetní náročností a kvalitou výsledků ukázal počet deseti sítí.

Shrnutí

Jako první přiblížení je vhodné zvolit úplnou dopřednou třívrstvou síť. Síť je vhodné trénovat algoritmem SCG nebo OSS. Pokud to paměťové možnosti dovolují, lze použít Levenberg-Marquardtovu metodu. Síť je vhodné trénovat v deseti exemplářích lišících se počátečním nastavením vah a prahů a pro další práci použít nejlepší exemplář.

4.9. Ověření doporučujících zásad

Použití neuronových sítí pro danou úlohu bylo ověřeno na soustavách složitějších než soustava použitá v hlavní části této práce. Konkrétně se jedná o SISO soustavu s nelineární tuhostí a MIMO soustavu se dvěma vstupy a dvěma výstupy.

4.9.1. Soustava s nelineární tuhostí

Pro ověření metodiky byla vybrána soustava s nelineární tuhostí

$$m\ddot{x} + b\dot{x} + \bar{k}(x, \dot{x}) = F(t), \quad (4)$$

kde m je hmotnost a b tlumení. Nelineární funkce tuhosti $\bar{k}(x, \dot{x})$ je definována podle (5)

$$\bar{k}(x, \dot{x}) = \alpha kx + (1 - \alpha)kdw, \quad (5)$$

kde α , k a d jsou kladná čísla. Proměnná w je definována následující diferenciální rovnicí:

$$\dot{w} = \frac{1}{d} \left(a\dot{x} - \beta|\dot{x}||w|^{p-1}w - \chi\dot{x}|w|^p \right), \quad (6)$$

kde a , β a χ jsou kladná čísla a p je liché číslo. Tento model simuluje chování s nelineární hysterezí.

Byla navržena dopředná neuronová síť následujících vlastností:

- topologie 5-6-1 (5 vstupních neuronů - buzení v časových krocích k , $k-1$, $k-2$; výstup soustavy v časových krocích $k-1$, $k-2$, 6 skrytých neuronů, 1 výstupní).
- Algoritmus poměrných konjugovaných gradientů

V souladu s výše uvedenými závěry bylo trénováno 10 sítí, lišících se pouze náhodným nastavením vah a prahů na počátku. Z těchto sítí byla vybrána síť s nejnižší tréninkovou chybou. Po provedení tréninku byla síť testována na testovacích datech.

Při testování odezva sítě odpovídala požadované odezvě s dostatečnou přesností. Vyšší hodnoty chyby sítě jsou patrné v oblastech změny orientace signálu, nicméně i u největšího překmitu v oblasti 26-28 vteřiny je chyba menší než $4e-3$.

Závěrem lze pro soustavu s nelineární tuhostí konstatovat, že se podařilo neuronovou síť naučit požadovanému chování a neuronový model tak použít pro případné další úlohy (řízení, atd.).

4.9.2. MIMO soustava

Neuronová síť je schopná modelovat i soustavy typu MIMO. Pro ověření navržených doporučení byla testována soustava podle [12], popsané vztahem (7)

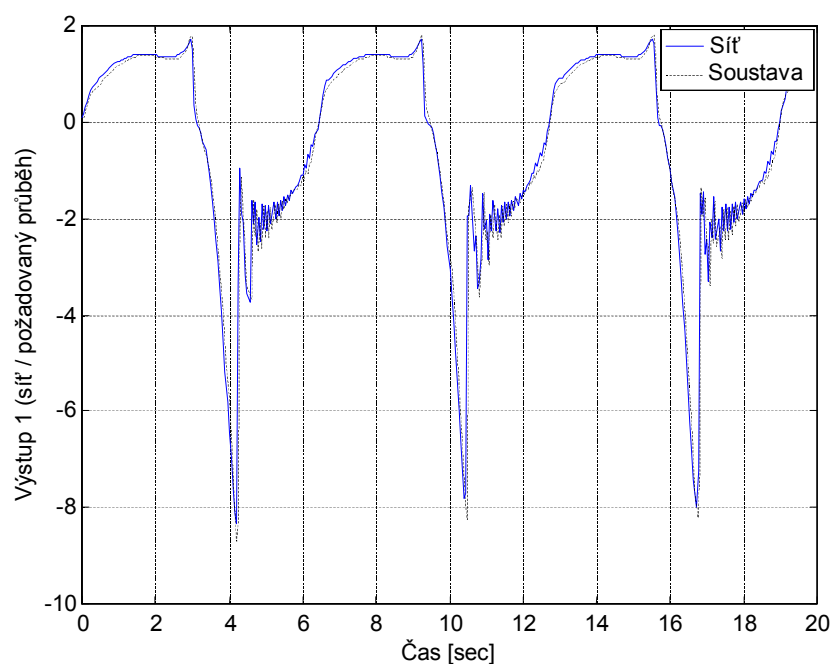
$$\begin{bmatrix} y_1(k+1) \\ y_2(k+1) \end{bmatrix} = \begin{bmatrix} \frac{y_1(k)}{1+y_2^2(k)} \\ \frac{y_1(k)y_2(k)}{1+y_2^2(k)} \end{bmatrix} + \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix}, \quad (7)$$

kde k je časový krok, u_1 je buzení prvního vstupu soustavy, u_2 je buzení druhého vstupu soustavy a y_1, y_2 je první, respektive druhý výstup soustavy. Jako tréninkových dat bylo využito odezvy soustavy na náhodné buzení po dobu 1000 vteřin s délkou časového kroku pět setin vteřiny.

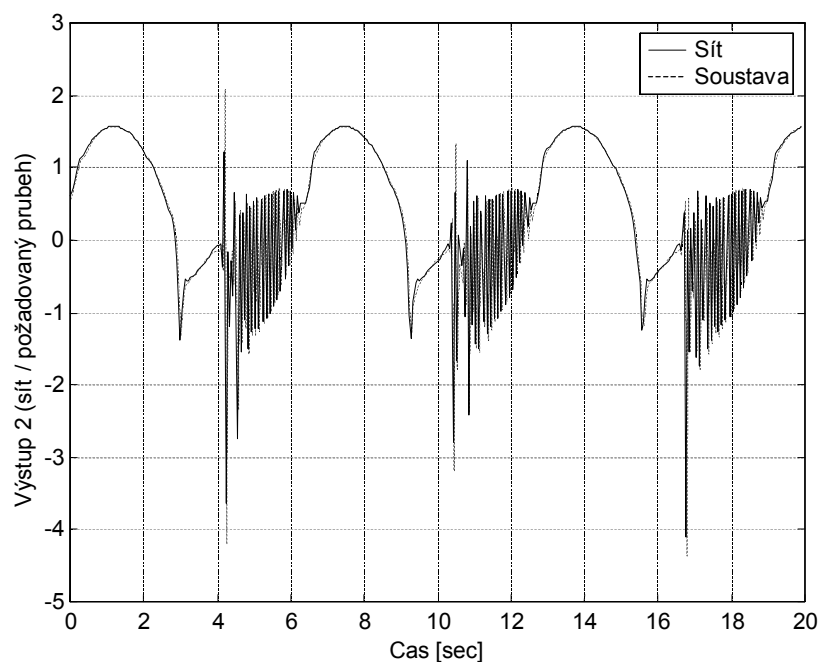
Byla navržena dopředná neuronová síť následujících vlastností:

- topologie 8-10-2 (8 vstupních neuronů- obě buzení v časových krocích $k, k-1$; oba výstupy soustavy v časových krocích $k-1, k-2$, 10 skrytých a 2 výstupní - výstupy soustavy v časovém kroku k)
- Algoritmus poměrných konjugovaných gradientů

Po provedení tréninku byla síť testována na testovacích datech. Bylo použito buzení sinovou křivkou pro první vstup a kosinovou pro druhý. Porovnání požadovaných výstupů soustavy a odezev sítě je uvedeno na obrázku 10 a 11. Obr. 10. uvádí časový průběh odezvy sítě a výstupu systému pro první výstup y_1 , obr. 11. uvádí srovnání pro výstup y_2 . Z grafů je zřejmé, že výstup sítě je téměř totožný s požadovaným průběhem, a lze konstatovat, že se podařilo neuronovou síť naučit požadovanému chování.



Obr. 10. Výstup 1 - Porovnání požadovaného průběhu a odezvy sítě



Obr.11. Výstup 2 - Porovnání požadovaného průběhu a odezvy sítě

5. ZÁVĚR

V předložené práci jsou uvedeny výsledky analýzy použití umělých neuronových sítí pro úlohu modelování soustav s dynamickým chováním. Práce je uvedena teoretickým přehledem použitých paradigmat neuronových sítí. Je navržen soubor parametrů podstatných pro chování sítí u dané úlohy a tento je dále zkoumán. Na datech popisujících chování jednoduché soustavy je analyzován vliv jednotlivých parametrů na kvalitu sítí z různých hledisek, především přesnost, generalizační schopnosti, výpočetní náročnost. Je rozebrán vliv počátečního náhodného nastavení vah a prahů sítě. Dále je diskutována možnost automatického generování sítí pomocí genetického programování. Na závěr je uvedeno několik doporučujících zásad pro nasazení neuronových sítí v obdobných úlohách a tyto zásady jsou ověřeny na složitějších soustavách.

Zadání disertační práce bylo splněno ve stanoveném rozsahu a je možné jej shrnout v následujících bodech:

- První fáze byla věnována vyhledání perspektivních paradigmat neuronových sítí použitých pro úlohu modelování dynamických soustav a příbuzných typů úloh.
- Další práce byly zaměřeny na zpracování jednotlivých paradigmat. Byly použity systémy Matlab, SNNS a vlastní implementace.
- Souběžně s implementacemi byl prováděn výběr parametrů podstatných pro řešení úlohu. Jako sledované parametry byly vybrány:
 - typ sítě (dopředná / rekurentní)
 - topologie sítě
 - algoritmus učení

- počáteční nastavení vah a prahů sítě

Jako kritéria určující kvalitu sítě byla vybrána přesnost odezvy sítě na testovacích datech a výpočetní a implementační náročnost.

- d) Hlavní částí práce bylo provedení numerických experimentů. Implementovaná paradigmatu byla použita na jednoduché úloze a byla zpracována srovnávací studie vlivu sledovaných parametrů na kritéria určující kvalitu sítě. Zvláštní pozornost byla věnována vlivu počátečního nastavení vah, který byl zkoumán na souboru 400 sítí pro rekurentní i dopředný model. Tento soubor byl statisticky zpracován a práce uvádí charakteristiky vlastního vývoje tréninku a vztah mezi tréninkovými a testovacími výsledky. Dále byly určeny pravděpodobnosti dosažení určité limitní hodnoty koncové trénovací chyby pro i z n sítí.
- e) Samostatný oddíl se věnuje principu, implementaci a výsledkům sítí s topologií generovanou automaticky pomocí genetického programování. Metoda prokazuje slibné výsledky, ale v současné době je příliš výpočetně náročná. Pokud je k dispozici dostatek výpočetních prostředků, je výhodná pro automatický návrh optimální struktury sítě.
- f) Na základě provedených experimentů bylo navrženo několik doporučujících zásad při nasazení neuronových sítí v úloze simulačního modelování soustav s dynamickým chováním, které představují hlavní závěr práce. Za nejdůležitější výsledky je možné považovat následující konstatování:
- Dopředné sítě prokazují lepší chování než sítě rekurentní.
 - Jako algoritmus učení při uvážení jak přesnosti výsledků a rychlosti učení, tak implementační a výpočetní náročnosti algoritmu je nejvhodnější některá varianta gradientního algoritmu, především metoda poměrných konjugovaných gradientů nebo One Step Secant metoda, případně v závislosti na paměťových možnostech Levenberg-Marguardtova metoda.
 - Zcela zásadní vliv na kvalitu výsledků má počáteční nastavení parametrů sítě - vah a prahů. Tento vliv je podstatnější než použitý algoritmus učení. Je vhodné vytvořit alespoň deset sítí stejných parametrů lišících se pouze v počátečním nastavení vah a prahů a z těchto sítí použít síť s nejlepšími testovacími výsledky.
 - Některé výhodné vlastnosti sofistikovanějších algoritmů (automatické určení topologie u kaskádové korelace nebo u sítí generovaných genetickým programováním) nepřevažují vyšší implementační náročnost, případně vysoké nároky na výpočetní čas.
- g) Doporučující zásady byly ověřena na dvou úlohách - soustavě s hysterezní tuhostí a soustavě typu MIMO. Pro obě úlohy byly použity neuronové sítě navržené na základě doporučujících zásad, a vykazují v obou případech dobré výsledky.

Závěrem lze říci, že hlavním **teoretickým přínosem práce** je zhodnocení nejpoužívanějších paradigmat neuronových sítí použitých na úloze modelování dynamické soustavy. Při studiu odborné literatury nebyla obdobná studie nalezena, některé zdroje uvádějí pouze zhodnocení vybraných algoritmů učení z hlediska rychlosti konvergence, což pro reálné využití není dostačující.

Uvedené zhodnocení sítí, podpořené o statistický rozbor, vyústilo v doporučující zásady - viz. předchozí odstavec f). Doporučení byla ověřena a představují hlavní **praktický přínos práce**.

LITERATURA

- [1] Broomhead D.S., Lowe D. Multivariable functional interpolation and adaptive networks, *Complex Systems*, 1988, vol. 2, pp. 321-355.
- [2] Bury V. K. *Statistical models in applied science*, John Wiley & Sons, New York, 1975.
- [3] Dennis J.E., Schnabel R. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Englewood Cliffs, NJ, Prentice Hall, 1983.
- [4] Elman J.L. Finding Structure in Time, *CRL Tech. Report 8801*, Univ. of California at San Diego, Center for Research in Language, 1988.
- [5] Fahlman S.E. The Recurrent Cascade-Correlation Architecture, *Tech. Report CMU-CS-91-100*, Carnegie Mellon University, Pittsburgh, USA, 1991.
- [6] Gruau F. "Neural network synthesis using cellular encoding and the genetic algorithm", *PhD. Thesis*, Claude Bernard University, Lyon, 1994.
- [7] Haykin S. *Neural Networks - A Comprehensive Foundation*, Macmillian College Publishing Company, NJ, USA, 1994.
- [8] Janíček P., Ondráček E., Porada V. Identifikace objektů a systémů v mechanice, lékařství, biomechanice a kriminalistice, *Inženýrská mechanika*, 1996, Vol. 3., č.2, pp. 71-88.
- [9] Kosko B. *Neural Networks and Fuzzy Systems*, Prentice Hall International, 1992.
- [10] Koza J.R. *Genetic programming: On the programming of computers by mean of natural selection*, MIT press, 1992.
- [11] Moller M. F. A scaled conjugate gradient algorithm for fast supervised learning, *Technical report*, Computer Science Department, University of Aarhus, 1990.
- [12] Narendra K.S., Parthasarathy K. Identification and Control of Dynamical Systems Using Neural Networks, *IEEE Transactions on Neural Networks*, 1990, vol. 1, no. 1, pp. 4-27.
- [13] Riedmiller M, Braun H. A direct adaptive method for faster Backpropagation learning: The RPROP Algorithm, *Proceedings of the IEEE International Conference on Neural Networks*, 1994, pp. 987-996.

- [14] Rumelhart D.E., Hinton G.E., Williams R.J. Learning internal representations by error propagation, *In Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, 1986, Vol.1, MA, MIT Press.
- [15] Werbos P.J. Backpropagation through time: What it does and how to do it, *Proceedings of the IEE 90*, 1990, pp. 1550-1560.
- [16] Williams R.J., Zipser D. A Learning Algorithm for Continually Running Fully Recurrent Neural Networks, *Neural Computation*, 1989, No. 1, pp 270-280.

SEZNAM PUBLIKACÍ AUTORA

- [A1] Dumek L., Grove T., Raudenský M., Krejsa J. Novel Approaches to the IHCP: Neural Networks, *Inverse Problems in Engineering Mechanics*, Edited by H.D.Bui, A.A.Balkema, Rotterdam, Brookfield, 1994, pp. 411-416.
- [A2] Raudenský M., Horský J., Krejsa J., Sláma L. Výzkum chlazení povrchů s vysokými teplotami a vývoj inverzních úloh, *Engineering Mechanics '95*, Svratka, CZ, 1995.
- [A3] Raudenský M., Horský J., Krejsa J. Usage of Neural Network for Coupled Parameter and Function Specification of Inverse Heat Conduction Problem, *National Heat Transfer Conference*, Portland, USA, 1995.
- [A4] Raudenský M., Horský J., Krejsa J. Usage of Neural Network for Coupled Parameter and Function Specification of Inverse Heat Conduction Problem, *International Communications in Heat and Mass Transfer*, 1995.
- [A5] Březina T., Sláma L., Krejsa J. Genetic Algorithms: Three examples of the Use, *Workshop of Tempus JEP 4143-94*, Praha, 1995.
- [A6] Sláma L., Krejsa J., Ondrůšek Č., Kratochvíl C., Březina T. Rozdílné přístupy k modelování elektromechanických pohonných soustav, *Výpočtová Mechanika 95*, Perning, CZ, 1995.
- [A7] Roupec J., Hažmuk I., Krejsa J. Campus Computer Network at the Faculty of Mechanical Engineering in the Technical University of Brno and its further development, *INVEX - CCT'95*, Brno, CZ, 1995.
- [A8] Raudenský M., Horský J., Krejsa J., Tseng I. Experimental Study of Heat Transfer on Rotating Roll cooled by Water Nozzles, *ISROMAC 6*, Oahu, USA, 1996.
- [A9] Sláma L., Raudenský M., Horský J., Březina T., Krejsa J. Application of Genetic Algorithm for Evaluation of Quenching Test by Inverse Task with Unknown Time Constant of Sensor, *AIENG 96*, Florida, USA, 1996.
- [A10] Krejsa J., Sláma L., Horský J., Raudenský M., Pátíková B. The Comparison of the Traditional and Non-classical Methods Solving the Inverse Heat Conduction Problem, *Heat Transfer 96*, Udine, Italy, 1996.

- [A11] Krejsa J., Slavík J., Grove T. Použití neuronové sítě typu Back-Propagation pro řešení inverzní úlohy pružnosti, *Engineering Mechanics '95*, Svratka, CZ, 1996.
- [A12] Sláma L., Krejsa J., Kratochvíl C. Different Approaches in Modelling of the Electromechanical Drive Systems, *GAMM 96*, Praha, CZ, 1996.
- [A13] Raudenský M., Horský J., Krejsa J., Sláma L. Usage of Artificial Intelligence Methods in Inverse Problem for Estimation of Material Parameters, *International Journal of Numerical Methods for Heat and Fluid Flow*, USA, 1996.
- [A14] Sláma L., Raudenský M., Horský J., Březina T., Krejsa J. Evaluation of Quenching Test of Rotating Roll with Unknown Time Constant of Sensor Using Genetic Algorithm, *Mendel 96*, Brno, CZ, 1996.
- [A15] Roupec J., Krejsa J. Dominance and Recessivity in Genetic Algorithms, *Mendel 96*, Brno, CZ, 1996.
- [A16] Raudenský M., Horský J., Krejsa J., Sláma L. Non-classical approaches to the Inverse Problems, *General Workshop MMSP'96*, Davos, Switzerland, 1996.
- [A17] Březina T., Krejsa J., Kratochvíl C., Sláma L. Approximation and prediction of the non-linear system's behaviour via artificial neural network, *Days of Mechanics*, Varna, Bulgaria, 1996.
- [A18] Březina T., Krejsa J., Kratochvíl C., Sláma L. Modelling of the electromechanical system via neural network, *Výpočtová Mechanika 96*, Pernink, Czech Republic, 1996.
- [A19] Březina T., Krejsa J., Žalud P. Neural Network: Application on Electromechanical System's Modeling, *Engineering Mechanics 97*, CZ, 1997.
- [A20] Březina T., Krejsa J., Slavík J. Some possibilities of Neural Network Application, *Mechatronics and Robotics 97*, 1997, Brno, pp. 15-20.
- [A21] Krejsa J., Březina T., Doležal, Sláma L. Establishing the Behaviour of Polymerization Unit via Neural Network, *Mechatronics and Robotics 97*, 1997, Brno, pp. 37-42.
- [A22] Slavík J., Sláma L., Březina T., Kratochvíl C., Krejsa J. Application of Simulated Annealing and Genetic Algorithms to the Optimization of Vibration Isolation Systems, *Mechatronics and Robotics 97*, 1997, Brno, pp. 93-98.
- [A23] Slavík J., Sláma L., Krejsa J., Kratochvíl C. Genetic Algorithms and Simulated Annealing Used for Design Optimization of Damped Supports, *3NWGA*, ed. by Jarmo T. Alander, 1997, FAIS, pp. 255-264.
- [A24] Březina T., Krejsa J., Kratochvíl C. The Contribution to Automatized Design of Neural Networks for Non-linear Dynamic Systems Simulation, *Mechatronika 97*, 1997, Poland.

- [A25] Březina T., Krejsa J., Kratochvíl C. The Automatized Design of Neurocontrollers using Genetic Algorithm, *EIS'98 Symposium*, Tenerife, Spain, 1998.
- [A26] Krejsa J., Březina T. Simulace nelineárního dynamického systému evoluční neuronovou sítí, *Colloquium Dynamics of Machines 98*, 1998, Prague, CZ.
- [A27] Březina T., J. Krejsa J., Kratochvíl C., Slavík J. Modelling of Dynamic Systems using Neural Networks generated by Genetic Algorithms, *Mechanics in Design (MID '98)*, 1998, Nottingham, UK, pp 689 - 698.
- [A28] Doležal R., Krejsa J., Březina T. Srovnání klasického PI regulátoru s fuzzy a neuro-fuzzy regulátory, *Engineering Mechanics 98*, Svratka, 1998.
- [A29] Březina T., Krejsa J. A Few Notes about Neural Nets Generated via Cellular Grammar, *Mendel 98*, Brno, ČR, 1998.
- [A30] Březina T., Krejsa J. Neural Networks Used for Dynamic Systems Simulation and Neurocontrollers Design, *Seventh conference on nonlinear vibrations, stability and dynamics of structures*, Blacksburg, Virginia, USA, 1998.
- [A31] Doležal R., Kotek V., Krejsa J. Simulační modelování nelineárních dynamických systémů, *Colloquium Dynamics of Machines 99*, 1999, Praha, pp. 53-58.
- [A32] Doležal R., Krejsa J. Užití metod umělé inteligence pro řízení dynamických systémů, *Aplikovaná mechanika*, 1999, Brno, pp 35-42.
- [A33] Krejsa J., Doležal R. Neural networks for Non-linear System Identification, *Engineering Mechanics 99*, 1999, vol. 3., pp. 791-796.
- [A34] Doležal R., Krejsa J., Březina T. Adaptivní identifikace nelineárních soustav pomocí neuronových sítí, *Engineering Mechanics 99*, 1999, vol. 3., pp. 765-770.
- [A35] Krejsa J., Doležal R. Recurrent Neural Networks for Non-linear System Identification, *Zeszyty Naukowe Katedry Mechaniki Stosowanej*, 11/99, IIIrd. *International Seminar of Applied Mechanics*, 1999, Wisla, pp. 37-40.
- [A36] Březina T., Krejsa J., Kratochvíl C., Doležal R. Neural Nets Evolved by GA - Simulation and Control of Dynamic Systems, *19th Computers in Engineering Conference*, 1999, Alexis Park Resort and Spa, Las Vegas, Nevada, CD-ROM.
- [A37] Kratochvíl C., Kotek V., Březina T., Krejsa J. Modelling and Analysis of Dynamic Properties of Drive Systems with Gears, *17th Biennial Conference on Mechanical Vibration and Noise*, 1999, Alexis Park Resort and Spa, Las Vegas, Nevada, CD-ROM.
- [A38] Sláma L., Balátě M., Krejsa J., Slavík J. Identification of a Nonlinear Vibration System using Simulated Annealing, Genetic Algorithm and Programming, *17th Biennial Conference on Mechanical Vibration and Noise*, 1999, Alexis Park Resort and Spa, Las Vegas, Nevada, CD-ROM.

- [A39] Doležal R., Krejsa J., Březina T. Identification of Nonlinear Dynamic Systems Using Neuro-Fuzzy Networks, *Mechatronics and Robotics 99*, 1999, Brno, pp.111-116.
- [A40] Krejsa J., Woodbury K. A., Ratliff J. D., Raudenský M. Assessment of Strategies and Potential for Neural Networks in the IHCP, *int. journal Inverse Problems in Engineering*, 1999, Vol 7, pp. 197-213.
- [A41] Březina T., Krejsa J., Kratochvíl C. Neurocontrol Using Continuous CMAC, *Engineering Mechanics 2000*, 2000, Vol III, pp. 163-168.
- [A42] Březina T., Krejsa J., Sláma L. Setting Parameters for Quadratic PID-controller for Magnetic Bearing by Genetic Algorithm, *Mendel 2000*, ČR 2000, pp 33-38.
- [A43] Březina T., Krejsa J., Doležal R. Recurrent Versus Feed-forward Neural Networks used for Identification of Non-linear Systems, *Mechatronika 2000*, Warsaw, Poland, 2000, přijato k publikaci.
- [A44] Kozánek J., Březina T., Krejsa J., Kratochvíl C. Some Notes to Modelling of the Magnetic Bearing, *Eighth Conference on Nonlinear Vibrations, Stability and Dynamics of Structures*, Blacksburg, Virginia USA, 2000, přijato k publikaci.
- [A45] Březina T., Krejsa J., Kratochvíl C., Kozánek J., Pavelka J., Šafr M. Reinforcement Learning - Control of the Magnetic Bearing, *Eighth Conference on Nonlinear Vibrations, Stability and Dynamics of Structures*, Blacksburg, Virginia USA, 2000, přijato k publikaci.

SUMMARY

Presented PhD thesis is focused on the use of artificial neural networks on the dynamic systems modeling task. Thank to the fast development in past few years, the neural networks seem to be a suitable method for such a task due to various advantageous properties, mainly the fact that neural networks do not require mathematical model of modeled system. Unfortunately there are also certain disadvantageous properties, mainly the difficult verification of the networks results, high demands during the network learning process and necessity of the neural network expert, who determines which type of the network use on the base of its experiences. Given work analyses the influence of several parameters essential when neural networks are used for given task and as the main results presents the recommending principles that should be helpful for prospective researchers.

The work contains theoretical overview of used neural network paradigms and the set of parameters essential for the network behavior on given task. The set is further examined. The both classical and nontraditional paradigms of both feed-forward and recurrent neural networks are used (various versions of back-propagation algorithm, cascade correlation, radial basis functions, Elman model, Jordan model, recurrent cascade correlation, recurrent RBF networks, etc.). The influence of chosen

parameters is analyzed on the data describing the behavior of simple system from various points of view, mainly the precision, generalization and computational requirements.

The influence of initial random setting of networks weights and biases was found to be essential. The influence is processed using statistical methods on the set of 400 networks of the same class with identical topology and parameter settings, differing only in initial weights setting, both for feed-forward and recurrent networks. The work also discusses the possibility of automatic generation of neural networks using genetic programming, which seems to be promising method for the future, but so far it is disproportionately demanding.

The recommending principles for the use of neural networks in mentioned class of tasks are the main result of the work and principles are verified on more complex systems - system with hysterezy stiffness and MIMO system. The networks designed on the base of recommending principles showed good results for both tasks.

ŽIVOTOPIS

Narozen 20.6.1971 v Boskovicích, národnost česká, svobodný, trvalé bydliště A. Slavíčka 19, Svitavy.

Vzdělání

- Gymnázium v Boskovicích, specializace výpočetní technika, ukončeno maturitou 1989.
- Vysoké učení technické v Brně, Fakulta strojní, obor Inženýrská mechanika, ukončeno v roce 1994 státní závěrečnou zkouškou a obhajobou diplomové práce na téma: "Řešení inverzního problému rovinné úlohy pružnosti pomocí neuronových sítí".
- Od roku 1994 doktorandské studium v oboru Inženýrská mechanika na Ústavu mechaniky těles FS VUT Brno.

Praxe:

1996 - dosud: vědecký pracovník Ústavu mechaniky těles, FS VUT Brno

Vědecká práce:

Výzkum v oblasti neuronových sítí použitých v inženýrských úlohách. Více než 40 publikací ve vědeckých časopisech a na tuzemských i zahraničních konferencích.

Výuka:

Základní kurzy - kinematika, dynamika

Specializace - neuronové modelování

Jazykové znalosti: plyná angličtina, základy francouzštiny.