

ШТУЧНІ НЕЙРОННІ МЕРЕЖІ ЯК АПРОКСИМАЦІЙНИЙ АПАРАТ В ЗАДАЧАХ ПРОЕКТУВАННЯ РАДІОТЕХНІЧНИХ ПРИСТРОЇВ

Адаменко В. О.¹, аспірант; Мірських Г. О.², к.т.н., доцент

¹ Національний технічний університет України

"Київський політехнічний інститут", м. Київ, Україна,

*² Національний університет біоресурсів і природокористування України,
м. Київ, Україна*

ARTIFICIAL NEURAL NETWORKS AS APPROXIMATE PROCEDURE IN WIRELESS DEVICES DESIGNING PROBLEMS.

V. Adamenko¹, postgraduate, G. Mirskikh², PhD, associate professor

¹ National Technical University of Ukraine "Kiev Polytechnic Institute", Kyiv, Ukraine

² National University of Bioresources and Nature Management of Ukraine, Kyiv, Ukraine

Вступ

На сьогодні сфера застосування нейронних мереж (НМ) стрімко зростає, адже нейронні мережі це перш за все потужний апарат математичного моделювання, який здатен відтворювати залежності будь якої складності [1]. При цьому часто найважливішим аргументом на користь використання НМ є її здатність до навчання на підставі експериментальних даних (а то і на підставі дослідження окремих варіантів фізичної реалізації відповідного об'єкту) дозволяє моделювати невизначені залежності, тобто такі, які не відтворюються прийнятою математичною моделлю. За будь-яких умов першим і найважливішим етапом використання НМ є визначення її оптимальної архітектури та методу навчання, адже саме це безпосередньо впливає як на час навчання, так і на величину похибки відтворення шуканої залежності (величину точності апроксимації). Як правило, вибір архітектури НМ здійснюється на підставі відповідних експериментів (чисельних та/або натурних) та має ситуативний характер, тобто повністю залежить від досліджуваного об'єкту та особливостей завдання, яке має бути розв'язане за допомогою НМ.

Наведена робота присвячена вибору оптимальної конфігурації та методу навчання НМ, як складового елементу архітектури модифікованого ансамблю НМ прийнятого авторами для реалізації алгоритмів проектування частотовибіркових НВЧ пристроїв (металодіелектричних смугопоскних фільтрів — МДФ) [2].

Визначення оптимальної конфігурації НМ

Для реалізації алгоритмів проектування МДФ навчання НМ проводиться з метою апроксимації розрахункових значень S -параметрів МДФ з кількістю резонаторів $n = 3 \div 6$ та нерівномірністю в смузі пропускання $L_a = 1$ дБ. На вхід мережі (рис. 1) подавалася інформація звичайна для реалізації алгоритмів синтезу смугопропускних НВЧ фільтрів [3]: кількість резонаторів, значення нерівномірності амплітудно-частотної характеристики у смузі пропускання (в дБ) та приведені значення частоти (див. рис. 1). На виході НМ формується інформація щодо S -параметрів. Навчальна вибірка складалася з 404 наборів значень.

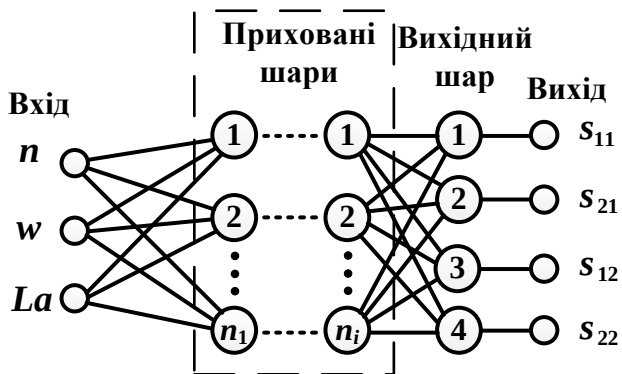


Рис. 1. Конфігурація нейронної мережі

Варіантів НМ введемо наступне позначення параметрів її архітектури: $vd \times n_1 \times n_2 \times n_i \times vhd$, де vd – кількість входів, vhd – кількість нейронів у вихідному шарі (для даного експерименту $vd = 3$, $vhd = 4$), n_i – кількість нейронів в i -тому шарі. Кількість нейронів шару для двохшарової мережі обиралася межах від 10 до 500 (10 різних конфігурацій), для решти мереж від 5 до 50 (по 5 різних конфігурацій), максимальна кількість нейронів шару визначалася експериментально по різкому зростанню часу навчання. Вибіркові результати навчання зведено до табл. 1.

Визначення оптимальної конфігурації НМ проводилася шляхом аналізу результатів тестових навчань групи МН з різною кількістю шарів та нейронів в цих шарах. При цьому у всіх нейронах використана сигмоїдальна функція активації, яка описується за допомогою гіперболічного тангенсу.

З метою ідентифікації

Таблиця 1.

Результати навчання нейронних мереж з різними параметрами конфігурації

№	Параметри архітектури	Приведений час навчання	Похибка навчання, 10^{-4}	Середньоквадратичне відхилення, 10^{-4}
1	2	3	4	5
2	3x50x4	0,4	1,9	8,1
3	3x100x4	0,8	1	5,3
4	3x250x4	6,5	510	210
5	3x10x10x4	0,3	7,3	20
6	3x30x30x4	2	2,1	4
7	3x50x50x4	12,5	5,6	2
8	3x15x15x15x4	1	$\approx 0,1$	2,65
9	3x30x30x30x4	9,8	$\approx 0,1$	2,64
10	3x35x35x35x4	14,5	$\approx 0,1$	1,75

Продовження таблиці 1

1	2	3	4	5
11	3x10x10x10x10x4	0,5	$\approx 0,1$	5,6
12	3x20x20x20x20x4	2,1	$\approx 0,1$	9,2
13	3x25x25x25x25x4	13,4	$\approx 0,1$	2,18

В якості алгоритму навчання обрано метод Левенберга-Маркварда (*Levenberg-Marquardt*) [4, 5]. Кожна НМ підлягала навчанню декілька разів (створено декілька мереж однієї конфігурації, але з різними початковими значеннями вагових коефіцієнтів), результат обирався зі схожих варіантів для усунення випадкового фактору, який притаманний методам зворотного поширення помилки.

Головними показниками визначення оптимальної конфігурації НМ є якість апроксимації та час навчання затрачений на досягнення потрібного результату.

Перевірка якості апроксимації під час навчання відбувається за допомогою контрольної вибірки, яка резервується з навчальної вибірки та не приймає участь у навчанні НМ. Після кожного етапу навчання на вхід НМ подавалася контрольна вибірка, і якщо помилка навчання та помилка результатів реалізації контрольної вибірки зменшуються, то вважалося, що навчання йде в потрібному напрямку (адже не має підстав для протилежного висновку). Якщо ж похибка навчання зменшувалася, а похибка реалізації контрольної вибірки залишалася незмінною або зростала, то приймалася гіпотеза про запам'ятовування НМ навчальної вибірки (результатом є більше значення середньоквадратичної похибки у порівнянні з похибкою навчання).

На якість апроксимації впливає кількість регульованих елементів введених у НМ. До цих елементів відносяться вагові коефіцієнти нейронних зв'язків та зміщення нейронів. Якщо кількість зв'язків не достатня, то НМ не здатна апроксимувати характеристику з заданою точністю. Збільшення ж вільних зв'язків може призвести до ефекту перенавчання НМ, коли запам'ятовуються виключно дані навчальної вибірки, а в проміжних точках значення, надані НМ, практично непередбачувані.

Оцінку якості апроксимації проведено за середньоквадратичним відхиленням між навчальними даними та даними, отриманими на виході НМ. В якості критерію зупинки навчання прийнята умова поведінки помилки на контрольній вибірці: якщо вона перестає зменшуватися або зростає, то мережа досягла оптимального режиму апроксимації і подальше навчання тільки погіршить його. Для зручності порівняння результатів час навчання НМ нормовано, адже він залежить від швидкодії обчислювальної техніки, на якій проводяться розрахунки. Нормування здійснено по відношенню до часу навчання мережі з конфігурацією №8

Аналіз експериментальних даних показує, що найменшу середньоквадратичну похибку апроксимації отримано з використанням НМ №№7, 8, 9, 10 та 13. Взявши до уваги час навчання, можна зробити висновок, що оптимальною архітектурою є № 8, адже решта варіантів має значно більший час навчання. Крім того, мережа № 13 має складнішу конфігурацію, що може сприяти виникненню традиційних проблем перенавчання та прокляття розмірності [6].

Збільшення кількості шарів у нейронній мережі дозволяє зменшити кількість нейронів в кожному шарі без погіршення (а то й з покращенням) апроксимуючих властивостей НМ. При цьому зменшується кількість регульованих елементів (вагових коефіцієнтів та коефіцієнтів зміщення), що в свою чергу захищає НМ від традиційних проблем перенавчання, потрапляння до локальних мінімумів та паралічу.

Вибір оптимального методу навчання

Для вибору оптимального методу навчання взято нейронну мережу з еонфігурацією, яка показала найкращу якість апроксимації в ході попереднього експерименту: $3 \times 15 \times 15 \times 15 \times 4$, тобто з трьома входами, трьома шарами по 15 нейронів в кожному та вихідним шаром з 4 нейронів. Всього проведено порівняння 7 поширених методів навчання нейронних мереж: методу Левенберга-Маркварда, навчання з використанням квазі-Ньютонівського методу BFGS [7], методу регуляризації Байєса [8], методу спряжених градієнтів Пауелла – Беале (*Powell-Beale*) [9], методу градієнтного спуску, методу градієнтного спуску з врахуванням моментів [10] та алгоритму зворотного пружного поширення [11]. Для кожного методу було проведено три експерименти, результати навчання за даними алгоритмами зведено до табл. 2. Нормування часу навчання проведено до класичного методу навчання мереж: методу градієнтного спуску.

Таблиця 1.

Результати порівняння різних методів навчання

№	Назва методу	Приведений час навчання	Похибка навчання, 10^{-4}	Середньоквадратичне відхилення, 10^{-4}
1	Метод Левенберга-Маркварда	0,1	0,78	2,36
2	Квазі-Ньютонівський метод BFGS	0,13	57	82
3	Метод регуляризації Байєса	0,3	0,014	0,32
4	Метод зв'язаних градієнтів Пауелла – Беале	0,05	40	48
5	Метод градієнтного спуску	1	160	170
6	Метод градієнтного спуску з врахуванням моментів	1,03	130	150
7	Алгоритм зворотного пружного поширення	0,08	52	57

Слід відзначити, що використання класичних методів навчання НМ (методу градієнтного спуску та градієнтного спуску з врахуванням моментів) не припустимо при значній складності залежностей, які повинна апроксимувати НМ. Адже в такому випадку поверхня помилки має дуже складний характер і не зважаючи на значно вищу швидкість обрахунку однієї епохи у порівнянні з іншими методами, час навчання значно зростає. Прийняттого результату не вдалося досягти і при збільшенні кількості епох навчання на декілька порядків.

На рис. 2 наведено процес навчання нейронної мережі різними методами (1 – Квазі-Ньютонівський, 2 – регуляризації Байєса, 3 – Левенберга-Маркварда, 4 – зворотнього пружного поширення, 5 – зв'язаних градієнтів Пауелла – Беале).

Метод градієнтного спуску та градієнтного спуску з моментами на графіку не наведено, так як табличного значення середньоквадратичної похибки вони досягають лише після декількох тисяч епох навчання. Збільшення кількості епох навчання для решти методів практично не впливає на зниження похибки апроксимації, а тільки призводить до перенавчання нейронної мережі.

Навчання НМ методом Левенберга-Маркварда показує найкращу динаміку зменшення похибки на початкових етапах навчання, і метод регуляризації Байєса наздоганяє перший тільки після 20 епох навчання. Квазі-Ньютонівський метод, алгоритм зворотного пружного поширення та метод зв'язаних градієнтів Пауелла-Беале демонструють практично однакову динаміку, причому вже після 40 епох практично припиняється зниження похибки апроксимації.

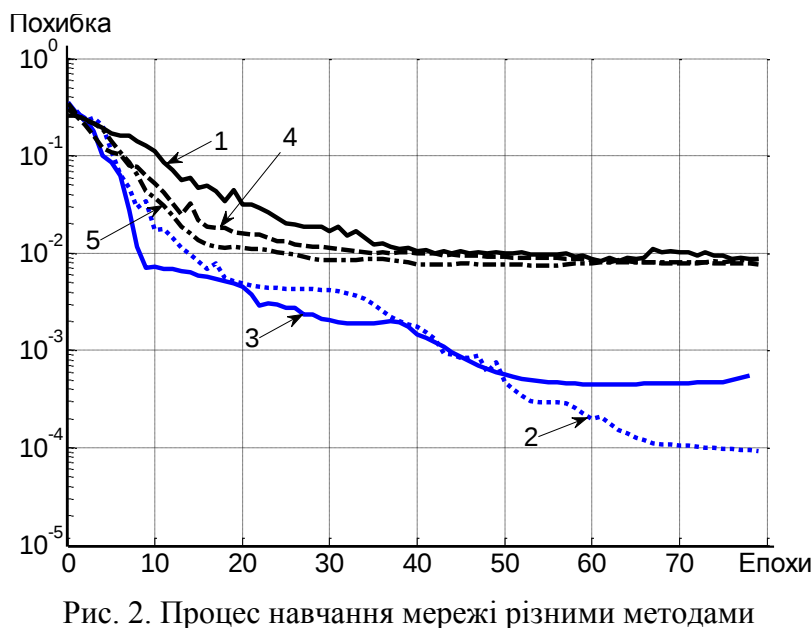


Рис. 2. Процес навчання мережі різними методами

Виходячи з експериментальних даних можна зробити висновок, що для поставленої задачі апроксимації S - параметрів мікрохвильових НВЧ пристроїв доцільно використовувати метод регуляризації Байєса, який після проходження 60 епох починає випереджати метод Левенберга-Маркварда по значен-

ню помилки навчання. Також на відміну від останнього, метод регуляризації Байєса показує значно кращу динаміку зниження перевіркової похибки, що доводить вищу якість апроксимації, рис 3.

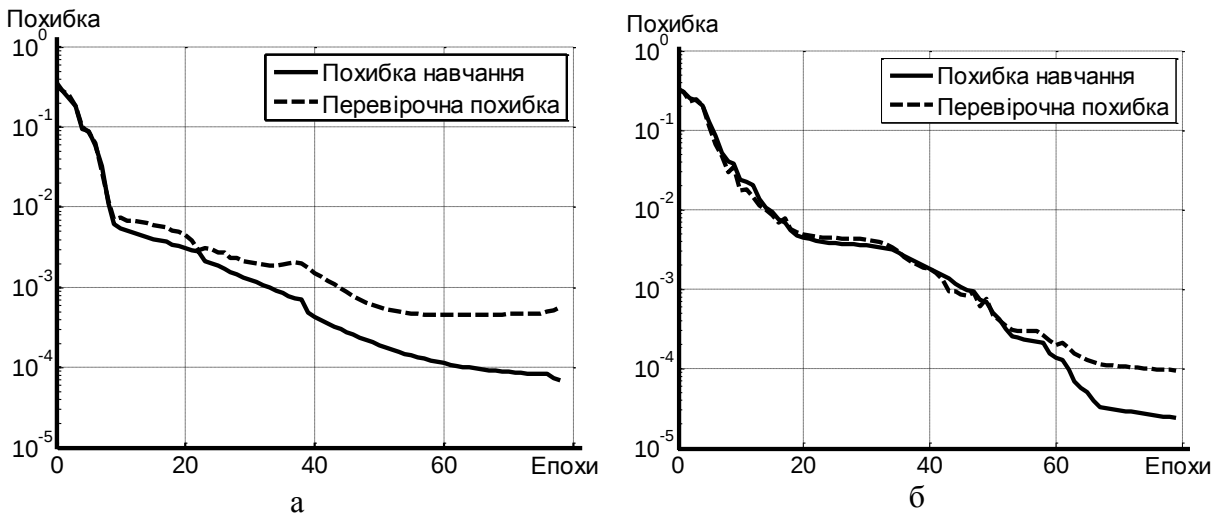


Рис. 3. Динаміка зміни похибки навчання та перевіркової похибки для методу Левенберга-Маркварда (а) та методу регуляризації Байєса (б)

Схожість поведінки помилки навчання для цих методів зумовлена тим, що метод регуляризації Байєса оснований на модифікації значень вагових коефіцієнтів та зміщень за алгоритмом Левенберга-Маркварда, але крок та напрям зміни визначається за байєсівською ймовірністю, яка дозволяє значно покращити узагальнюючі властивості мережі.

Зважаючи на швидкість роботи методу зв'язаних градієнтів Пауелла – Беале та алгоритму зворотного пружного поширення можна зробити висновки, що дані методи доцільно використовувати для апроксимації простих характеристик, при яких вони здатні забезпечити прийнятну точність апроксимації та максимальну швидкодію.

Висновки

Використання нейронних мереж дозволяє проводити апроксимацію складних характеристик, наприклад таких, як частотні залежності S – параметрів частотно-вибіркових НВЧ пристроїв. Точність такої апроксимації напряму залежить від належного вибору конфігурації та методу навчання нейронної мережі.

Збільшення шарів нейронної мережі призводить до значного покращення її апроксимаційних характеристик. Результати дослідження показують, що найбільш ефективним є використання 4 шарів, кількість нейронів в кожному шарі повинна знаходитися в межах від 10 до 20. Це дозволяє забезпечити необхідну кількість регульованих елементів, для якісної апроксимації при цьому не виникає явища перенавчання мережі.

Оптимальним методом навчання для складних характеристик є метод регуляризації Байєса, але для зменшення часу навчання можна використовувати і метод Левенберга-Маркварда, який також дозволяє отримати за-

довільну точність апроксимації. Решта з методів, серед яких проводилися дослідження, не підходять для апроксимації складних функцій, їм не вистачає гнучкості. Традиційні методи градієнтного спуску та градієнтного спуску з врахуванням моментів хоч і потребують мінімуму часу на проходження епохи, натомість при складних функціях необхідна кількість епох зростає до десятків тисяч.

Література

1. Горбань А. Н. Обобщенная аппроксимационная теорема / А.Н. Горбань // Сибирский журнал вычислительной математики, 1998. — Т.1 — № 1. — С.12 — 24. .
2. Адаменко В. О. Використання нейронних мереж для синтезу мікрохвильових пристроїв / В. О. Адаменко, Г. О. Мірських // Вісник НТУУ "КПІ". Серія — Радіотехніка. Радіоапаратобудування. — 2012. — №49. — С. 102—107.
3. Маттей Д. Л. Фильтры СВЧ, согласующие цепи и цепи связи / Д. Л. Маттей, Л. Янг, Е.М.Т. Джонс / Пер. с англ. / Под ред. Л.В. Алексеева, Ф.В. Кушнера. — М.: Связь, 1971. — т.1. — 248 с.
4. Hagan M. T. Training feedforward networks with the Marquardt algorithm / M. T. Hagan, M. B. Menhaj // Neural Networks, IEEE Transactions on Neural Networks — 1994. — vol. 5, No. 6. — P. 989 — 993.
5. Meng-Hock Fun. Levenberg-Marquardt training for modular networks / Meng-Hock Fun, O. K. Stillwater, M. T. Hagan // IEEE International Conference on Neural Networks. — 1996. — vol. 1 — P. 463 — 473.
6. Адаменко В. О. Штучні нейронні мережі в задачах реалізації матеріальних об'єктів. Частина 2. Особливості проектування та застосування / В. О. Адаменко, Г. О. Мірських // Вісник НТУУ "КПІ". Серія — Радіотехніка. Радіоапаратобудування. — 2012. — №48 — С. 213 — 221.
7. Dennis J. E. Numerical Methods for Unconstrained Optimization and Nonlinear Equations / J. E. Dennis, R. B. Schnabel — SIAM. — 1987. — 394 P. — ISBN: 978-0-89871-364-0
8. Foresee F. D. Gauss-Newton approximation to Bayesian regularization / F. D. Foresee, M. T. Hagan // Proceedings of the 1997 International Joint Conference on Neural Networks — 1997. — P. 1930 — 1935.
9. Медведев В. С. Нейронные сети. MATLAB 6 / В. С. Медведев, В. Г. Потемкин. — М. : ДИАЛОГ-МИФИ, 2002. — 496 с. — ISBN 5-86404-163-7.
10. Hagan M. T. Neural Network Design / M. T. Hagan, H. B. Demuth, M. H. Beale — China Machine Press, 1996. — 734 P.— ISBN 7-111-10841-8.
11. Riedmiller M. A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm / M. Riedmiller, H. Braun // IEEE International Conference On Neural Networks, 1993.

References

1. Horban A. N. Obobshchennaya approksymatsyonnaya teorema. Sybyrskyy zhurnal vychyslytelnoy matematyky, 1998, vol.1, no. 1, pp.12-24.
2. Adamenko V. O., Mirskykh G. O. Using neural networks for the synthesis of microwave devices. Visnyk NTUU "KPI". Seriya — Radiotekhnika. Radioaparatobuduvannya, 2012, no. 49, pp. 102-107. (in Ukrainian)
3. Matthey D.L., Yanh L., Dzhons E.M.T. Fyltry SVCh, sohlasuyushchye tsepy y tsepy svyazy. Moscow, Svyaz, 1971, vol. 1, 248 p.
- 4 Hagan M. T., Menhaj M. B. Training feedforward networks with the Marquardt algorithm. Neural Networks, IEEE Transactions on Neural Networks, 1994, vol. 5, no. 6. pp. 989-993.

5. Meng-Hock Fun, Stillwater O. K, Hagan M.T. Levenberg-Marquardt training for modular networks. IEEE International Conference on Neural Networks, 1996, vol. 1, pp. 463-473.
6. Adamenko V.O., Mirskykh G.O. Artificial neural networks in problems of material objects implementation. Part 2. Networking principles and Classification. Visnyk NTUU "KPI". Seriya — Radiotekhnika. Radioaparotobuduvannya, 2012, no. 48, pp. 213-221. (in Ukrainian)
7. Dennis J. E., Schnabel R. B. Numerical Methods for Unconstrained Optimization and Nonlinear Equations. SIAM, 1987, 394 p., ISBN: 978-0-89871-364-0
8. Foresee F. D., Hagan M. T. Gauss-Newton approximation to Bayesian regularization. Proceedings of the 1997 International Joint Conference on Neural Networks, 1997, pp. 1930-1935.
9. Medvedev V. S., Potemkin V. G. Nejronnye seti. MATLAB 6. Moscow, DIALOG-MIFI, 2002, 496 p., ISBN 5-86404-163-7.
10. Hagan M. T., Demuth H. B., Beale M. H. Neural Network Design. China Machine Press, 1996, 734 p., ISBN 7-111-10841-8.
11. Riedmiller M., Braun H. A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm. IEEE International Conference On Neural Networks, 1993.

Адаменко В. О., Мірських Г. О. Штучні нейронні мережі як апроксимаційний апарат в задачах проектування радіотехнічних пристроїв. В роботі розглянуто переваги застосування НМ в якості апроксимаційного апарата в задачах проектування частотовибіркових НВЧ пристроїв. Визначено та проаналізовано час та похибку навчання НМ різної конфігурації (кількість шарів та нейронів у кожному шарі) при апроксимації s-параметрів МДФ з різною кількістю резонаторів. Досліджено динаміку зміни похибки при використанні різних методів навчання НМ та проведено вибір оптимального методу навчання з врахуванням затраченого часу та похибки.

Ключові слова: штучна нейронна мережа; частотно-вибіркові мікрохвильові пристрої; апроксимація характеристик; оптимальна конфігурація НМ;

Адаменко В. А., Мирских Г. А. Искусственные нейронные сети как аппроксимирующий аппарат в задачах проектирования радиотехнических устройств. В работе рассмотрены преимущества применения нейронных сетей (НС) в качестве аппроксимирующего аппарата в задачах проектирования частотноизбирательных СВЧ устройств. Определено и проанализировано время и погрешность обучения НС разной конфигурации (количество слоев и нейронов в каждом слое) при аппроксимации s-параметров МДФ с разным количеством резонаторов. Исследовано динамику изменения погрешности во время использования разных методов обучения НС и проведен выбор оптимального метода обучения с учетом затраченного времени и погрешности.

Ключевые слова: искусственная нейронная сеть; частотноизбирательные микроволновые устройства; аппроксимация характеристик; оптимальная конфигурация НС;

Adamenko V., Mirskikh G. Artificial neural networks as approximate procedure in wireless devices designing problems.

Purpose. Present work is dedicated to the optimal configuration selection and training method of neural network (NN). This NN is architecture's element of modified NN ensemble accepted by authors for implementation of frequency-selective microwave devices design algorithms.

Optimal configuration determining of NN. Optimal configuration determining of NN was received by analyzing the results of test NN training with different number of layers and neurons in these layers. The main parameters optimal configuration determining of NN is the approximation quality and total learning time.

Choosing of optimal teaching method. NN training methods comparison was carried out for 7 popular training methods: Levenberg-Marquardt backpropagation, BFGS quasi-Newton backpropagation, Bayesian regulation backpropagation, Conjugate gradient backpropagation with Powell-Beale restarts, Gradient descent backpropagation, Gradient descent with momentum backpropagation and Resilient backpropagation.

Conclusions. NN using allows to approximate complex features of microwave devices, such as frequency dependencies of S – parameters etc. The approximation accuracy depends on configuration and method of NN training. Increasing the number of NN layers leads to improvement of approximate characteristics. According to our results the most effective is usage of 4 layers and the number of neurons in each layer should be over the range 10 to 20. Optimal training method for complex characteristics is Bayesian regulation backpropagation, for time training reduction can be used Levenberg-Marquardt backpropagation.

Keywords: artificial neural network, frequency-selective microwave devices; approximation characteristics, optimal configuration of neural network;