B15

A FUZZY-NEURAL APPROACH FOR THE CHARACTERISATION OF THE ACTIVE MICROWAVE DEVICES

Bekir KARLIK[1], Hamid TORPI[2], Musa ALCI[3]
[1]University of Bahrain, Dept. of Computer Eng., Kingdom of Bahrain
E-mail: bkarlik@itc.uob.bh
[2]Yildiz Technical University, Dept. of Electronics &Communication Eng., Turkey
E-mail: torpi@yildiz.edu.tr
[3]Ege University, Dept. of Electrical & Electronics Eng., Turkey
E-mail: alci@bornova.ege.edu.tr

Abstract — Artificial Neural Networks are emerging as a powerful technology for RF and microwave characterization, modeling, and design. Neural modeler helps us to immediately start developing neural models for RF/Microwave components and circuits and helps to provide neural models for our simulators. In this study, a novel fuzzy neural network structure is used for behavior of an active microwave device. Here, the device is modeled by a black box whose small signal and noise parameters are evaluated through a fuzzy clustering neural network based upon the fitting of both of these parameters.

I. Introduction

Recently, artificial neural networks are introduced to the microwave community, opening the door for an unconventional approach to microwave computer-aided design [1—11].

The problem of approximating measured device parameters or device response has been formulated as an optimization problem with respect to the equivalent circuit of a proposed model. The traditional approach in modeling is almost entirely directed at achieving the best possible match between measured and calculated parameters. This approach has serious shortcomings in two frequently encountered cases. The first case is when the equivalent circuit parameters are not unique with responses selected and the second is when the non ideal effects are not modeled adequately, the latter causing an imperfect match even if small measurement errors are allowed for. In both cases, a family of solutions for circuit model parameters may exist which produce reasonable and similar matches between measured and calculated responses. Besides, published literature is concerned with the equivalent circuit for the single-bias which are only either small-signal models or the noise behavior descriptions based on existing signal equivalent circuits that have nothing to do with the device noise characteristics. In [12] and [13] these two behaviors are combined in an unified classical circuit model for only a single-bias. A recent work [7] combines the signal and noise parameters in a neural network model over the fairly large operation band at a single bias point.

The aim of this work is to model a microwave transistor by a black-box for which small-signal and noise parameters are evaluated through a fuzzy-neural network, based upon the fitting of both of these parameters to the corresponding measured data over the whole operational range from DC to more than 10 GHz for multiple bias of various types of configuration.

II. Artificial Neural Networks

Artificial Neural Networks are information processing systems inspired by the ability of human brain to learn from observations and to generalize by abstraction. The fact that neural networks can learn totally different things, led to their use in diverse fields such as pattern recognition, speech processing, control, medical applications and more. Recently, microwave researchers developed techniques in which neural networks are trained from microwave data, and then used to enhance microwave design [5-8]. In particular, researchers have successfully used ANNs to model micro-strip via [9], packaging and interconnects [10], spiral inductors [11], MESFET devices [12], and CPW circuit components [13], to name just a few.

The ANN architecture used in this work is a feed-forward, three-layer perceptron structure (MLP3) consisting of an input layer, a hidden layer, and an output layer, as shown in Figure 1.

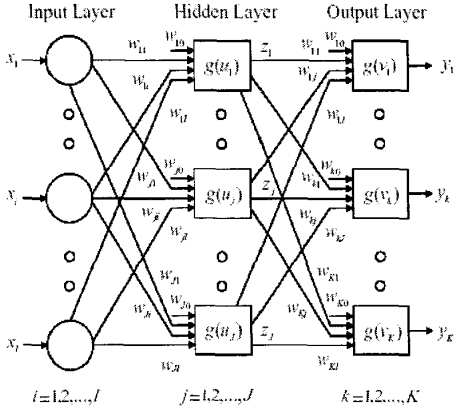

Figure 1. Artificial neural network architecture

The back propagation training through generalized delta rule of learning is an iterative gradient algorithm designed to minimize the root mean square error between the actual output of the multi-layer feed-forward NN and the desired output. Each layer is fully connected to the previous layer, and has no other connection. The cost function of the MLP is given as

$$\varepsilon(n) = \frac{1}{2}\sum_{k=1}^{N_o} e_k^2(n) \qquad (1)$$

where, $\varepsilon(n)$ is the instantaneous cost function at iteration n, $e_k(n)$ is the error from output node k at iteration n and N_o is the number of output nodes.

The error from each output node can be defined as

$$e_k(n) = d_k(n) - y_k(n) \qquad (2)$$

Where, $d_k(n)$ is the desired response of output node k at iteration n and $y_k(n)$ is the output of the output node k at iteration n.

Haykin (1994) gives a summary of the back-propagation algorithm as follows:

2002 12th Int. Crimean Conference "Microwave & Telecommunication Technology" (CriMiCo'2002). 9-13 September, Sevastopol, Crimea, Ukraine
© 2002: CriMiCo'2002 Organizing Committee; Weber Co. ISBN: 966-7968-12-X. IEEE Catalog Number: 02EX570

Initialization. Set all the weights and threshold levels of the network to small random numbers that are uniformly distributed.

Forward computation: Let a training example be denoted by [x (n), d (n)], with the input vector x (n) applied to the input layer and the desired response vector d (n) presented to the output layer. The net internal activity level $v_j^{(l)}(n)$ for neuron j in layer l is given by

$$v_j^{(l)}(n) = \sum_{i=0}^{p} w_{ji}^{(l)}(n) y_i^{(l-1)}(n) \qquad (3)$$

where $y_i^{(l-1)}(n)$ is the signal from neuron i in the previous layer l-1 at iteration n and $w_{ji}^{(l)}(n)$ is the weight of neuron j in layer l that is connected to neuron i in layer l-1 at iteration n. For i = 0

$$y_0^{(l-1)} = -1 \qquad (4)$$

and

$$w_{j0}^{(l)}(n) = \theta_j^{(l)}(n) \qquad (5)$$

Where $\theta_j^{(l)}(n)$ is the threshold applied to neuron j in layer l. With the use of a logistic function for the sigmoid non-linearity, the output of neuron j in layer l is given by

$$y_j^{(l)}(n) = \frac{1}{1+\exp(-v_j^{(l)}(n))} \qquad (6)$$

If neuron j is in the first hidden layer (i. e., l = 1), set

$$y_j^{(0)}(n) = x_j(n) \qquad (7)$$

Where $x_j(n)$ is the jth element of input vector x (n). If neuron j is in the output layer (i. e., l = L), set

$$y_j^{(L)}(n) = o_j(n) \qquad (8)$$

The error can be computed as

$$e_j(n) = d_j(n) - o_j(n) \qquad (9)$$

Where $d_j(n)$ is the jth element of the desired response vector d (n).

Backward computation: Compute the local gradients (δ) of the network by progressing backward, layer-by-layer. For neuron j in output layer L, the local gradient is given by

$$\delta_j^{(L)}(n) = e_j^{(L)}(n) o_j(n)\left[1 - o_j(n)\right] \qquad (10)$$

For neuron j in hidden layer l, the local gradient is given by

$$\delta_j^{(l)}(n) = y_j^{(l)}(n)\left[1 - y^{(l)}(n)\right]\sum_k \delta_k^{(l+1)}(n) w_{kj}^{(l+1)}(n) \qquad (11)$$

The weight of the network in layer l can be adjusted according to the generalized delta rule as,

$$w_{ji}^{(l)}(n+1) = w_{ji}^{(l)}(n) + \alpha\left[w_{ji}^{(l)}(n) - w_{ji}^{(l)}(n-1)\right] + \eta\delta_j^{(l)}(n) y_i^{(l-1)}(n) \qquad (12)$$

Where, η is the learning rate parameter, and α is the momentum constant. In this study, the momentum constants were chosen to be 0.95 and 0.85 for FCNN3 and FCNN4 respectively. Moreover, the same learning rate parameter (η = 1) was obtained after different combinations of training epochs. The number of hidden nodes is determined experimentally. It was found as 12 for both neural network structures.

III. Fuzzy C-Means Clustering

Fuzzy c-means (FCM) is a data clustering technique wherein each data point belongs to a cluster to some degree that is specified by a membership grade. This technique was originally introduced by Jim Bezdek in 1981 [14] as an improvement on earlier clustering methods. It provides a method of how to group data points that populate some multidimensional space into a specific number of different clusters?

Structure identification of fuzzy systems is possible by constructing enough rules with the appropriate input and output membership functions. The identified model can accordingly be used to describe the behavior of the target system as well as for prediction and control purposes. The idea of fuzzy clustering is to divide the data into fuzzy partitions, which overlap with each other. Therefore, the containment of each data to each cluster is defined by a membership grade in [0,1]. In formal words, clustering in unlabeled data $X=\{x_1, x_2, \ldots, x_N\} \subset \Re^h$, where N is the number of data networks and h is the dimension of each data vector, is the assignment of c number of partition labels to the vectors in X. c-Partition of X are sets of (c. N) membership values $\{u_{ik}\}$ that can be conveniently arrayed as a (cxN) matrix $U=[u_{ik}]$. The problem of fuzzy clustering is to find the optimum membership matrix U. The most widely used objective function for fuzzy clustering in X is the weighted within-groups sum of squared errors objective function J_m, which is used to define the following constrained optimization problem

$$\min_{U,V}\left\{ J_m(U, V; X) = \sum_{k=1}^{N}\sum_{i=1}^{c} (u_{ik})^m \left\| x_k - v_i \right\|_A^2 \right\} \qquad (13)$$

where,

$$U \in M_{fcn} = \left\{ U \in \Re^{cN} \middle| \begin{matrix} 0 \le u_{ik} \le 1 \quad \forall ik \;\&\; \forall k, u_{ik} > 0 \exists i \\ 0 < \sum_{k=1}^{N} u_{ik} < n \forall_i \;\&\; \sum_{i=1}^{c} u_{ik} = 1 \;\forall k \end{matrix} \right\}$$

$V=\{v_1, v_2, \ldots, v_c\}$ is a vector of (unknown) cluster centers, and $\|x\|_A = \sqrt{x^T A x}$ is an inner product norm. A is an hxh positive definite matrix, which specifies the shape of the clusters. The matrix A commonly selected as the identity matrix, leading to the definition of Euclidean distance, and consequently to spherical clusters. Fuzzy partitions are carried out by the fuzzy C-means (FCM) algorithm through an iterative optimization of equation (13) according to the following steps [15]:

Step-1: Choose the number of clusters (c), weighting exponent (m), iteration limit (iter), termination criterion ($\varepsilon > 0$), and the norm for error $= \|V_t - V_{t-1}\|$.

Step-2: Guess initial position of the cluster centers: $V_0 = \{v_{1,0}, v_{2,0}, \ldots, v_{c,0}\} \subset \Re^{ch}$.

Step-3: Iterate for t = 1 iteration; Calculate

$$u_{ik,t} = \left[\sum_{j=1}^{c} \left(\frac{\left\| X_k - V_{i,t-1} \right\|_A}{\left\| X_k - V_{j,t-1} \right\|_A} \right)^{\frac{2}{m-1}} \right]^{-1} \qquad (14)$$

© 2002: CriMiCo'2002 Organizing Committee; Weber Co. ISBN: 966-7968-12-X. IEEE Catalog Number: 02EX570

$$\text{and } V_{i,t} = \frac{\sum_{k=1}^{N} (u_{ik,t})^m x_k}{\sum_{k=1}^{N} (u_{ik,t})^m} \qquad (15)$$

IF error = $\|V_t - V_{t-1}\| \le \varepsilon$, THEN stop, and put ($U_f$, V_f) = (U_t, V_t) NEXT t.

IV. Method and Results

The aim of this work is to model a microwave transistor by a black-box for which small-signal and noise parameters are evaluated through a fuzzy-neural network, based upon the fitting of both of these parameters to the corresponding measured data over the whole operational range from DC to more than 10 GHz for multiple bias of various types of configuration. So the stages of the work can be ordered as follows:

(a) Cluster all data using fuzzy C means clustering method,
(b) Establish a multi layered neural network of feed forward type with a single hidden layer,
(c) Use back-propagation and sigmoid activation functions, train the network for multiple bias and multiple configuration of any type of microwave transistor.
(d) Establish performance measure of the model,

From the classical point of view, unified small-signal-noise equivalent circuit for a microwave transistor can be divided in to two parts: Extrinsic circuits and intrinsic circuit. The intrinsic circuit characterizes the active region under the gate (or base) whose parameters are functions of bias conditions and device technological parameters, whereas the extrinsic parameters depend, at least to a first approximation, only on the technological parameters. If a unified circuit for a MESFET is considered, the most important extrinsic parameters are the gate, source and drain inductances due to the bond wires and the gate, source and drain resistances. The four main intrinsic parameters are the input capacitance C_{GS}, the trans conductance g_m, the output conductance g_d, and the feedback capacitance C_{GD}. Moreover, the electrical behavior of the intrinsic device requires the introduction of two more parameters: The intrinsic resistance R_i, which can be related to distributed nature of the RC input network, and the delay τ, which is introduced in the expression for the current generator and corresponds to the time needed for the carriers to travel under the gate. Briefly due to the intrinsic and extrinsic device properties, both the signal and noise parameters are the functions of the bias conditions, frequency, and configuration types. The way to approximate these functions in the literature so far is considered in the forthcoming subsection.

In this study, First we used experimental data. To evaluate the quality of the fit to measured data the following error terms are found to be convenient:

$$E_{S_{ij}} = \frac{1}{n}\sum_{k=1}^{n} \frac{\left|S^k_{ij_{meas.}} - S^k_{ij_{meas.}}\right|}{\left|S^k_{ij_{meas.}}\right|} \qquad (16)$$

$$E_{N_i} = \frac{1}{n}\sum_{i=1}^{n} \frac{\left|N^k_{i_{meas}} - N^k_{i_{predict.}}\right|}{\left|N^k_{i_{meas.}}\right|} \qquad (17)$$

Where S_{ij} and N_i are, respectively the signal and noise parameters, and n is the number of discrete frequencies used. Total average error can be defined as the average of the signal and noise errors:

$$E_T = \frac{1}{4}\sum_{i=1}^{4} E_{i_{signal}} + \frac{1}{3}\sum_{i=1}^{3} E_{i_{noise}} \qquad (18)$$

Then, a novel fuzzy clustering neural network structure was used for the training of these data. This neural network structure has two stage. At the first stage; inputs and outputs values of feed-forward type neural network are found using Fuzzy C-means Algorithm. Then at the second stage, these clustering data is applied as desired values of Multi-layered perceptron (MLP), which has only one hidden layer. (see Figure 2.)

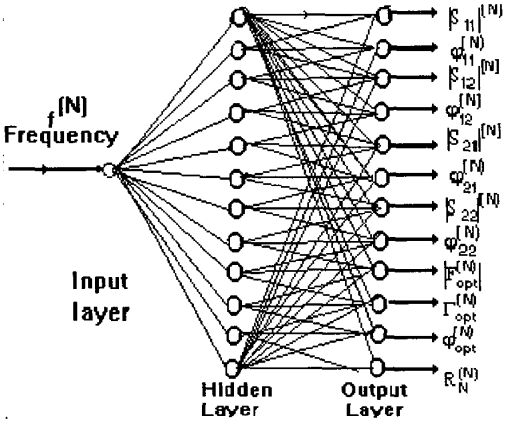

Figure 2. The structure of used neural network

After 10000 iterations, simulation results for NE02135 type transistor with respect to bias point (V_{CE}=10V I_C=2-25 mA) are given for various frequencies in Figure 3, which shows Distribution of errors with frequency for the whole band generalization at multiple bias points for the common emitter configuration. In this figure, FCNN3 means consisting of 3 clustering, in similar manner FCNN4, is consisting of 4 clustering. It can be seen this figure, after 20000 iterations, consisting of both clustering MSE errors are very close.

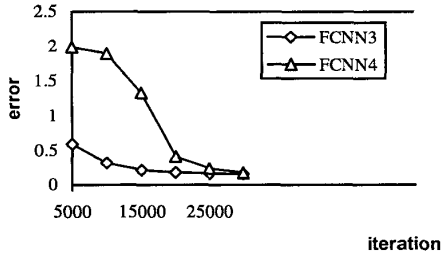

Figure 3. MSE errors for two different neural networks

2002 12th Int. Crimean Conference "Microwave & Telecommunication Technology" (CriMiCo'2002). 9-13 September, Sevastopol, Crimea, Ukraine
© 2002: CriMiCo'2002 Organizing Committee; Weber Co. ISBN: 966-7968-12-X. IEEE Catalog Number: 02EX570

Table 1 shows both results of FCNN and experimental values for consisting of 4 clustering.

III. Discussion

Models must represent components of circuits and systems as a pre-requisite of any computer-aided design. The availability, accuracy, flexibility of models, and the speed of model computation affect largely the effectiveness at all stages of computer-aided design. The concurrent modeling procedure doesn't require solving device physics equations repeatedly during optimization. Compared to the existing device modeling techniques, the proposed approach has the capability to make high-dimensional models for highly nonlinear devices. Using fuzzy clustering we can reduce data of variations of characterization of active microwave devices. It is very useful for the active microwave device users. Moreover, we don't need use the wrong experimental data because of the fuzzy clustering.

IV. References

[1] J. A. Jargon and K. C. Gupta «Artificial Neural Network Modeling For Improved On-Wafer Line-Reflect-Match calibrations», Proc. of 31st European Microwave Conference, vol. 2, pp. 229-232, London, ENGLAND, Sep. 2001.

[2] V. K. Devabhaktuni, M. C. E. Yagoub, Y. H. Fang, J. J. Xu, Q. J. Zhang, «Neural networks for microwave modeling: model development issues and nonlinear modeling techniques,» Int. J. RF and Microwave CAE, Vol. 11, 4-21, 2001.

[3] F. Gunes, F. Gurgen, H. Torpi, «Efficient Model Parameter Extraction Using Neural Networks for Active Microwave Devices», Proc. PIERS'96, 1996

[4] F. Gunes, F. Gugen, H. Torpi, «Signal-noise neural network model for active microwave devices», IEE Proc. Circuits Devices and Syst., Vol.143, No.1, February 1996, pp. 1-8.

[5] F. Wang and Q. J. Zhang, «Knowledge based neural models for microwave design,» IEEE Trans. Microwave Theory and Techniques, vol. 45, pp. 2333-2343, December 1997.

[6] A. Patnaik and R. K. Mishra, «ANN techniques in microwave engineering», IEEE Microwave Magazine 1 (2000), pp. 55-60.

[7] K. C. Gupta, «EM-ANN models for microwave and millimeter-wave components», IEEE Int Microwave Symp, Workshop on Applications of ANNs to Microwave Design, Jun. 1997, pp. 17-47.

[8] F. Wang, V. K. Devabhaktuni, C. Xi, and Q. J. Zhang, «Neural network structures and training algorithms for RF and microwave applications», Int. J RF and Microwave CAE 9 (1999), pp. 216-240.

[9] P. M. Watson and K. C. Gupta, «EM-ANN models for microstrip via and interconnects in dataset circuits», IEEE Trans Microwave Theory Tech 44 (1996), pp. 2495-2503.

[10] A. Veluswami, M. S. Nakhla, and Q. J. Zhang, «The application of neural networks to EM-based simulation and optimization of interconnects in high-speed VLSI circuits», IEEE Trans Microwave Theory Tech 45 (1997), pp. 712-723.

[11] G. L. Creech, B. J. Paul, C. D. Lesniak, T. J. Jenkins, and M. C. Calcatera, «Artificial neural networks for fast and accurate EM-CAD of microwave circuits», IEEE Trans Microwave Theory Tech 45 (1997), pp. 794-802.

[12] HU Z. R., YANG Z. M., FUSCO V. F., J. A. C. STEWART, «Unified small-signal-noise model for active microwave device», IEE Proceedings-G, Vol.140, (1), 1993, pp. 55-60.

[13] ROUX, J. P., ESCOTTE L., PLANA R., GRAFFEUIL, DELAGE S. L, »Small Signal and Noise Model Extraction Technique for hetero-junction Bipolar Transistor at Microwave Frequencies, «IEEE Trans. 1995-MTT-43 (2), pp. 293-298.

[14] J. C. Bezdek, «Pattern Recognition with Fuzzy Objective Function Algorithms», Plenum Press, New York, 1993.

[15] M R Emami, I B Turksen, A A Goldenberg, «An Improved Fuzzy Modeling Algorithm, Part II: System Identification», Proceeding of NAFIPS, pp.294-298, 1996.

Table 1. The values of consisting of 4 clustering

Type	$[S_{11}]^N$	$[\varphi_{11}]^N$	$[S_{21}]^N$	$[\varphi_{21}]^N$	$[S_{12}]^N$	$[\varphi_{12}]^N$	$[S_{22}]^N$	$[\varphi_{22}]^N$	$[F_{opt}]^N$	$[T_{opt}]^N$	$[\varphi_{opt}]^N$	$[R_N]^N$
FCNN	0.6698	-142.21	5.597	93.21	0.0832	31.44	0.4203	-58.97	1.365	0.3339	96.20	0.1324
EXPE	0.6699	-144.61	5.588	93.41	0.0850	30.97	0.4244	-59.55	1.351	0.3348	96.70	0.1299
FCNN	0.6491	156.67	2.731	63.96	0.0952	29.57	0.2789	-74.22	2.064	0.4980	181.81	0.0812
EXPE	0.6500	177.91	2.751	64.01	0.1000	27.00	0.3100	-74.00	1.999	0.5000	163.97	0.0500
FCNN	0.6540	151.86	1.961	46.26	0.1218	29.65	0.2962	-87.83	2.526	0.4730	168.64	0.0894
EXPE	0.6544	157.67	1.915	46.26	0.1244	28.24	0.3101	-88.30	2.490	0.4755	-168.78	0.0777
FCNN	0.6626	133.01	1.546	24.74	0.1526	32.45	0.2752	-111.01	3.803	0.6652	-139.84	0.2413
EXPE	0.6650	135.05	1.352	22.12	0.1500	26.00	0.3449	-110.90	3.638	0.6932	-140.21	0.1748
FCNN	0. 6625	133.30	1.549	25.20	0.1519	32.13	0.2773	-129.11	3.759	0.6598	-140.71	0.2362
EXPE	0.6600	130.27	1.760	27.84	0.1543	38.96	0.2043	-130.32	3.933	0.6308	-139.72	0.3061
FCNN	0.6506	179.64	5.089	75.08	0.0625	47.85	0.1839	-97.00	1.860	0.3310	157.78	0.1152
EXPE	0.6500	175.99	5.070	75.00	0.0600	49.00	0.1600	-95.00	1.900	0.3300	169.00	0.1300
FCNN	0.6468	172.71	2.935	56.41	0.1035	47.92	0.1926	-109.43	2.630	0.5029	-162.89	0.0592
EXPE	0.6460	157.90	2.964	56.36	0.0973	50.03	0.1566	-107.00	2.686	0.4999	-171.76	0.0886
FCNN	0.6400	-162.76	9.785	93.07	0.0419	45.92	0.2604	-79.40	1.799	0. 1606	150.72	0.1503
EXPE	0.6400	-163.00	9.880	93.00	0.0400	46.00	0.2600	-79.00	1.800	0.1600	149.00	0.1500

© 2002: CriMiCo'2002 Organizing Committee; Weber Co. ISBN: 966-7968-12-X. IEEE Catalog Number: 02EX570