RBF ANN OPTIMIZATION OF SYSTEMS REPRESENTED BY SMALL FDTD MODELING DATA SETS

Ethan K. Murphy[1] and Vadim V. Yakovlev[2]
[1] Colorado State University, Fort Collins, CO, USA
[2] Worcester Polytechnic Institute, Worcester, MA, USA

ABSTRACT
The paper describes a radial network procedure of numerical optimization backed by full-wave numerical analysis. To reduce computational effort, 3D FDTD data are added to the database as long as they are needed to achieve appropriate accuracy. The method involves optimization of the regularization parameter and radial function's radius. Performance of the procedure is illustrated by optimizing the efficiency of a double waveguide window and a microwave oven with a load on a shelf.
KEYWORDS: neural network, numerical optimization, radial basis function, regularization, S-parameters.

INTRODUCTION

In our previous paper [1] we have developed a novel technique for efficiency optimization of microwave systems. The method based on the radial basis function (RBF) artificial neural networks (ANN) scheme is designed for complicated, electrically large systems. The system efficiency is interpreted as energy coupling which is determined by the magnitude of the reflection coefficient ($|S_{11}|$) in the range around the operating frequency f_0. The network is trained by geometrical parameters and the frequency responses of $|S_{11}|$ generated by 3D finite-difference time-domain (FDTD) simulation. The computational procedure finds the system geometry characterized by the coupling which is not worse than the corresponding constraint.

In contrast to the traditional techniques of ANN microwave optimization [2], the analysis part of our approach is not backed by a method employing equivalent lumped and transmission line element networks, resulting in quick generation of data for the optimized system, because those methods would not be sufficiently adequate in representing complicated microwave structures. Rather, we rely on full-wave electromagnetic modeling, which appears to be the only option here. This, however, makes our optimization computationally extensive.

In this situation, a primary goal of an efficient strategy of optimization could be associated with either a motivated choice of fewer design variables, or using fewer points for their representation. The first approach has been operated in the optimization of practical devices [1, 3, 4], but the required computational resources turned out to be dependent on the type of problem, quality of data, etc.

This paper describes the procedure solving the optimization problem with less FDTD data. Instead of first building the entire database (DB) (whose sufficient size should be guessed and which may become unnecessarily large) and then training the network, in our scheme, FDTD data are put in the database as long as they are needed. Other functions include optimization of the RBF's radius and the regularization parameter controlling smoothness of data, and a choice of type of an RBF function.

RBF ANN OPTIMIZATION

We construct an RBF ANN (Fig. 1) that works with input vectors $\mathbf{X}_i = [x_1 \ ... \ x_n]$, and output vectors $\mathbf{S}_i = [|S(f_1)| \ ... \ |S(f_K)|]$, $i = 1, ..., P$, where $x_1, \ldots, x_n$ are system parameters (design variables), $S(f_j)$ is the value of an S_{mn} parameter at the jth frequency, and $j = 1, \ldots, K$. In our analysis, a certain form of a frequency response of a particular S-parameter is considered an objective function of the optimal design.

With the use of FDTD simulations we generate P samples of input-output pairs such that the data set is made of the matrices $\mathbf{X}$ and $\mathbf{S}$ containing $\mathbf{X}_i$ and $\mathbf{S}_i$. Our RBF network is coupled with a linear model $\mathbf{S} = \mathbf{X}\mathbf{W}_1 + \hat{\mathbf{\Phi}}(\mathbf{X})\mathbf{W}_2$, where $\mathbf{W}_1$, $\mathbf{W}_2$ are some weight matrices, and $\hat{\mathbf{\Phi}}$ is a matrix function containing RBFs. We consider the mapping $\mathcal{F}$: $\mathbf{X} \rightarrow \mathbf{S}$ and deal with the equation $\widetilde{\mathbf{S}} = \mathbf{\Phi}(\mathbf{X})\mathbf{W}$, where $\mathbf{W} = [\mathbf{W}_1 \;\; \mathbf{W}_2]^T$, $\mathbf{\Phi}$ is the matrix incorporating RBF with the linear model, and $\widetilde{\mathbf{S}}$ is the mapped data. We consider the use of the local Gaussian and the global cubic radial basis functions.

In order to fully describe $\mathcal{F}$ and appropriately choose the correct number of RBF centers, we partition the data set into a training set, $\{\mathbf{X}^{(1)}, \mathbf{S}^{(1)}\}$ of P_1 elements, and a testing set, $\{\mathbf{X}^{(2)}, \mathbf{S}^{(2)}\}$ of P_2 elements ($P = P_1 + P_2$). When training, we minimize the error function $E_j(\mathbf{w}_j) = \sum_{i=1}^{P_1} | \mathbf{S}_{ij}^{(1)} - (\mathbf{\Phi}(\mathbf{X}^{(1)})\mathbf{w}_j)_i |^2$, $j = 1, \ldots, K$; this can be done by the weight matrix $\mathbf{w}_j = (\mathbf{\Phi}^T\mathbf{\Phi})^{-1}\mathbf{\Phi}^T\mathbf{S}_j^{(1)}$. We assume that the number of RBFs is the same as the number of training points and pick the centers exactly on the points of $\mathbf{X}_i$.

Then the model is tested to see how well the network generalizes and approximates data that was not learned in training. When using the Gaussian RBF, we test the network by the data in the matrix $\widetilde{\mathbf{S}}^{(2)} = \mathbf{\Phi}(r, \mathbf{X}^{(2)})\mathbf{w}_j$, and write an error function as the mean square error $E(r) = \sum_{j=1}^{K}\sum_{i=1}^{P_2} \frac{| \widetilde{\mathbf{S}}_{ij}^{(2)} - (\mathbf{\Phi}(r, \mathbf{X}^{(2)})\mathbf{w}_j)_i |^2}{P_2}$, and for each individual RBF, we find the optimal radius r^* which makes it minimal.

Since there are no restrictions on the dimensionality of the input and output domains (n and K), the weight matrix can be of arbitrary size. If it is too large, then there are too many RBFs, and this leads to noisy data. To resolve the problem, we use ridge regression which adds a penalty term to the sum-squared-error. In this case, $E_j(\mathbf{w}_j)$ can be minimized by the weight matrix $\mathbf{w}_j = (\mathbf{\Phi}^T\mathbf{\Phi} + \lambda\mathbf{I}_\eta)^{-1}\mathbf{\Phi}^T\mathbf{S}_j^{(1)}$, where λ is the regularization parameter controlling smoothness of data. If we use the Gaussian function for the RBF, we perform a 2-parameter optimization with respect to λ and r.

Finally, we find the optimal value $\mathbf{S}^*$ with respect to the optimal $\mathbf{X}^*$ by solving the problem $\mathbf{S}^* = \min_{\mathbf{X}^*}/\max_{\mathbf{X}^*}\left[\mathbf{\Phi}(r^*, \mathbf{X}^*)\mathbf{w}^*\right]$, where $\mathbf{w}^*$ is determined from $\{\mathbf{X}^{(1)}, \mathbf{S}^{(1)}\}$. The search for optimal design depends on available data and a desirable profile of the goal function. When $\mathbf{S}_i$ represents a rather smooth function, then a regular minimization is performed. For better results with highly nonlinear data (in particular, in the presence of strong resonances), we transform $\mathbf{S}$ to a special matrix $\widetilde{\mathbf{S}}$ which allows for minimizing the average value of $\mathbf{S}_i$, maximizing the slopes, and minimizing the area below the curve. We train and optimize the model with $\widetilde{\mathbf{S}}$ as the new output matrix.

OPTIMIZATION WITH MINIMUM FDTD DATA

The algorithm outlined in the previous section works with the database which is supposed to be available. In the meantime, for complex and electrically large systems, generation of a DB is the most time-consuming part of the whole process. We introduce here an original yet simple procedure which keeps it under control and enforces dynamic generation of as many samples P as necessary for a required accuracy. A corresponding algorithm is presented by the chart in Fig. 2.

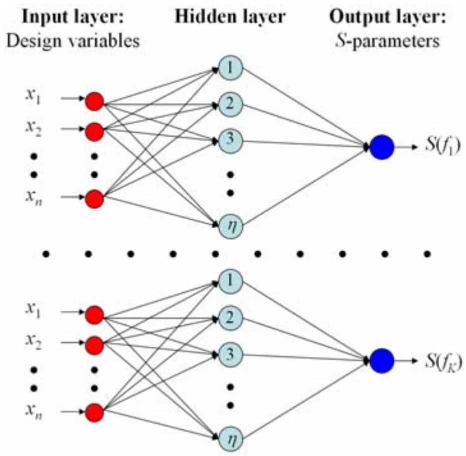

Fig. 1. Architecture of an RBF ANN

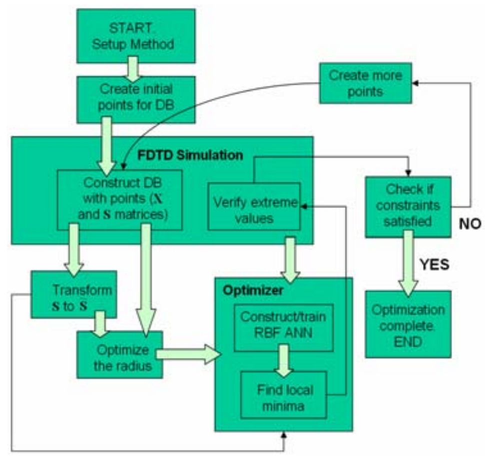

Fig. 2. Scheme of generating minimum FDTD data

An initial database is usually small. The user chooses d_k divisions of the design variables' ranges, i.e. creates a set of $P_{ini} = \prod_{k=1}^{n} d_k$ points equally spaced across the n-dimensional input domain. Then the procedure runs FDTD simulations to get the data matrices, creates the RBF network, transforms **S** in accordance with a goal function, optimizes the radius (if applicable), and uses the nonlinear least squares method to find a minimum solution. If a minimum solution passes the constraints, the procedure takes this minimum as the final solution and stops; otherwise, the algorithm proceeds to create more data points. The user also sets up a subsequent division g_k that specifies $P_{sub} = \prod_{k=1}^{n} g_k$ new points chosen in a quasi-random fashion. Cutting up each dimension into g_k pieces, the procedure makes P_{sub} boxes of dimension n and picks up new DB points as uniformly distributed random ones inside each box. After running the FDTD simulation with all these points, the database is of the size $P = P_{ini} + P_{sub}$. The algorithm continues in this fashion until the constraints are satisfied.

ILLUSTRATIONS AND DISCUSSION

The optimization procedure described above has been implemented in a MATLAB code in which minimization of the errors is performed with the use of the nonlinear least squares method (procedure `lsqnonlin` in the MATLAB Optimization Toolbox). Training/testing data for the network are generated by the full-wave 3D conformal FDTD simulator *QuickWave-3D* (http:// www.qwed.com.pl).

In the first example, we optimize a geometry of a double waveguide window (WW) consisting of a section of a rectangular waveguide WR340 and two rectangular dielectric plates (Fig. 3) [4]. The goal is to minimize $|S_{11}|$ (to make it less than 0.3) in the frequency range from 2.4 to 2.5 GHz by varying up to seven design variables. The FDTD model contains from 34,000 to 119,000 cells depending on the configuration of the device; the cell sizes are 4 mm (air) and 1.5 mm (Quartz). For 2-, 3-, and 4-parameter optimizations performed with the design variables (t_1, t_2), (t_1, t_2, s), and (t_1, t_2, s, $b_w = b_1 = b_2$) respectively, the optimal characteristic with the lowest average value in the zone of optimality is obtained with optimized r and λ rather than with non-optimized r and with the cubic RBF. When setting $d_k = 3$ and $g_k = 2$ in the 4-parameter optimization, the DB consists of only 92 points. Dealing with 4 or more design variables, the procedure generates optimal characteristics fully satisfying the constraints and guaranteeing 90% efficiency (Fig. 4).

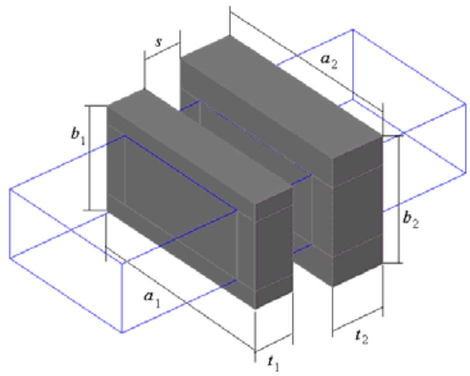

Fig. 3. Geometrical parameters of the WW model

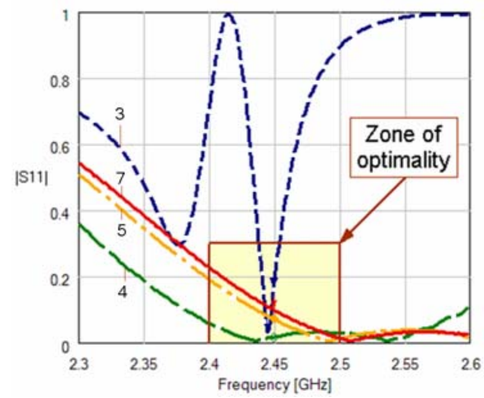

Fig. 4. Optimal WW characteristics obtained for different (3 to 7) number of design variables

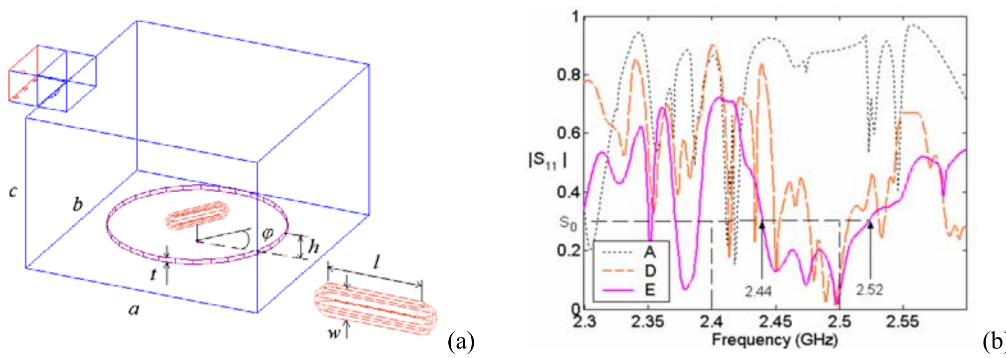

Fig. 5. Geometrical parameters of microwave the model of a microwave oven with a centered load on a glass shelf (a) and its characteristics (b): non-optimized (A) and resulted from three 2/3-parameter optimizations (D) [1] and one 5-parameter optimization (E).

In the second example, we optimize the efficiency of a microwave oven with a load on a shelf (Fig. 5, a) [1]. Since an adequate FDTD model requires from ~200,000 to 350,000 cells and at least 40,000 iterations in order to find the optimal solution within a reasonable time period, a special strategy was suggested in [1]: the problem was split into three subsequently solved 2- and 3-parameter optimizations. The characteristic obtained with this strategy does not satisfy the frequency constraints (curve D in Fig. 5, b). With a dynamically built DB, several 5-parameter optimizations have been successfully performed. The number of points does not exceed 500. One of the optimal characteristics is shown in Fig. 6 (curve D). It is seen that the curve's portion between 2.42 and 2.54 GHz lies below the constraint $S_0 = 0.3$.

These examples confirm that the proposed features (dynamically created DB, optimized RBF's radius, smoothing the data, etc.) of the ANN optimization backed by a full-wave numerical (FDTD) analysis make it computationally viable and efficient.

REFERENCES

1. E.K. Murphy and V.V. Yakovlev, *Proc. 9th AMPERE Conference on Microwave & HF Processing (Loughborough, U.K., September 2003)*, pp. 197-200.
2. Q.-J. Zhang, K.C. Gupta, and V.K. Devabhaktuni, *IEEE Trans. Microwave Theory Tech.*, vol. 51, No 4, pp. 1339-1350, 2003.
3. V.A. Mechenova, E.K. Murphy, and V.V. Yakovlev, *Proc. 38th IMPI Microwave Power Symp. (Toronto, Canada, July 2004)*, pp. 87-91.
4. V.V. Yakovlev, *Proc. 4th World Congress on Microwave & RF Processing (Austin, TX, Nov 2004*), pp. 172-174.