A Review of a Single Neuron Weight Optimization Model for Adaptive Beam Forming

K.S. Senthilkumar1, Lorothy Singkang2, P.R.P Hoole3, Norhuzaimin Julai3, S. Ang3, Shafrida Sahra3, Kismet Anak Hong Ping3, K. Pirapaharan4 and S.R.H. Hoole5
1Department of Computers and Technology, St. George’s University, Grenada, WI.
2Department of Mathematics, Science and Computer, Polytechnic Kuching, Sarawak, Malaysia.
3Department of Electrical and Electronic Engineering, Universiti Malaysia Sarawak, Sarawak, Malaysia.
4College of Engineering, Institute of Engineers Sri Lanka, Sri Lanka.
5Department of Electrical and Computer Engineering, Michigan State University, Michigan, USA.
prhoole@unimas.my

Abstract—In this paper, we review our recent, reported work on using artificial intelligence based software technique to control electronic sensor or wireless communication equipment in narrow and diverging paths such as in underground tunnels and at traffic junctions. In order to make the systems fast as well as needing minimal computational calculations and memory – thus to extend the battery life and minimize cost – we used the single layer Perceptron to successfully accomplish the formation of beams which may be changed according to the nature of the junctions and diverging paths the mobile or stationary system is to handle. Moreover, the beams that survey the scenario around (e.g. in case of guiding a driverless vehicle) or communicating along tunnels (e.g. underground mines) need to be kept narrow and focused to avoid reflections from buildings or rough surfaced walls which will tend to significantly degrade the reliability and accuracy of the sensor or communicator. These requirements were successfully achieved by the artificial intelligence system we developed and tested on software, awaiting prototype development in the near future.

Index Terms— Beam Forming; Neural Network; Single Layer Perceptron; Smart Antenna.

I. INTRODUCTION

In telecommunication, the applications of Adaptive Array Antenna have become popular due to its consistency in faster beam steering techniques that cannot be obtained by using a mechanical system or the switched beam array. Adaptive Array Antenna has the ability to detect and track other communication units and to generate narrow beams in a direction to align itself towards desired users. It can simultaneously minimize unwanted interferences or shadowing to achieve an optimized weight thus make it more flexible, smart and reliable. Here, a smart antenna called a Smart Beam Forming Antenna is proposed by combining an Adaptive Array Antenna and Artificial Neural Network (ANN) System. A Single Neuron Model is used to achieve weight optimization. A fast Neural Network Adaption is used for beam steering in order to align the adaptive beam towards the desired users while reducing or nulling interference from unwanted signals. The crucial part of designing a smart antenna is the ability of the antenna in handling intricate situations such as moving traffic patterns by providing flexible electrical tilt, beam width and azimuth control. Smart Antennas have been used at base stations (BS) as it provides a solution that is more versatile, cheaper, with low memory usage and fast beam steering technique. In parallel, the growth of fast cell site expansion, expanding the quantity of cell sectors and data transfer capacity (bandwidth), and better air interface abilities will be basic to move into introducing artificial intelligence (Perceptron) smart antennas to the best possible 5G frameworks.

A. Artificial Neural Network (ANN)

An ANN is a numerical structure which comprises interconnected artificial neurons, in a highly reduced scale works like the brain. An ANN can gain “experience” from information either in a managed or unsupervised way. In Figure 1 is shown how the ANN tries to copy the human brain that gets signals from sensors like the eye, ear and touch. These signals are then processed by the brain. In ANN the sensors might be real-time image sensors (camera), sound sensors (microphone) or capacitive touch sensors which are the inputs to the ANN. On account of smart antennas, the inputs are normally transmitted signals from transmitting antennas. In ANN the input signals are processed mathematically, such as by multiplying each input signal by a number (namely a weight, \(w_i\)) and phase shifting the signal (where complex weights are used, \(b_i\)). Subsequently, the weighted input signals are summed up and placed at the input of a transfer function (or activation function) block that will yield the final output signals. For the human brain, the final output signals might be activating signs to the muscles, for instance, to move the human body for physical activity. In smart antennas, the final output signals might be to redirect the beam towards the desired users. An ANN is structured (Figure 2) with a substantial number of highly interconnected processing elements called Artificial Neurons, which are organized in layers. The Weights \((w_i)\) and biases \((b_i)\) are known as Adjustable Scalar Parameters of the neuron. The parameters can be adjusted to meet the desired behavior as part of the network training process. The transfer function is expressed in one of the following forms: hard-limit (or step), the linear and the sigmoid (or logistic) function where the final output signals are mathematically expressed as

\[
Y = F(S) = F \left[ \sum_{k=1}^{N} x_k W_k + b \right] \tag{1}
\]
C. Single-Layer Perceptron (SLP)

The first pattern recognition machine for optical character recognition problem was designed in late 50’s known as Rosenblatt’s Perceptron. It comprises of binary activations and was trained to perceive linearly separable patterns in a limited number of steps. Linearly Separable is the problems with input patterns which can be classified using straight lines (or a single hyperplane) whereas Non-Linearly Separable the problems which can be classified but not by straight lines. A simple example of input patterns such as AND, OR operations are linearly separable and XOR operations are non-linearly separable. The equation for activation of the PE is:

\[ y(i) = f[w_1x_1 + w_2x_2 + b_1] \]  

The weights \( w_1, w_2 \) which indicate the gradient of the line and the bias \( b_1 \) indicate the offset. From the two-dimensional plane, the surface is:

\[ w_1x_1 + w_2x_2 - b_1 = 0 \]

Therefore,

\[ x_1 = \frac{b_1 - w_2}{w_1}x_2 \]  

It is a straight line. The input samples have no longer linearly controlled final outcome on SLP but rely on the activation function which clearly minimizes the output error. The learning rule makes a limited number of steps in SLP to achieve an ideal solution for linearly separable problems. The setup of the algorithm used in SLP is as follow:

1. Set weights with small random values for each connection.
2. For each training pair \( (x, d(i)) \): Calculate actual output \( y(i) \), Calculate error, \( \delta = (d(i) - y(i)) \) and use the error to adjust and update weights using the Equation (2).
3. Repeat step 2 until the error, \( \delta \) is minimized (closed to zero).

During the execution of SLP, the training will not stop until the problem is linearly separable and the fast adjustment to the weight values near to the end of training may influence the classification execution. Here, the Perceptron Learning Algorithm looks just for an acceptable answer, consequently the network may not perform well on data that is excluded in the training data. The quantity of outputs in an SLP is regularly controlled by the number of classes in the dataset. An SLP is observed to be helpful in characterizing a continuous-valued set of inputs into one of two classes as it were. When the problem that is not linearly separable it cannot be comprehended by SLP. Consequently, the SLP can be used only as a linear pattern recognition machine. However, ANNs are very powerful as it can represent the linear and non-linear relationships. The ANN has the ability to roughly model the input-out relationship by optimizing the weights using known input-output training pairs. Once the training is done, it is able to obtain the needed antenna radiation beam for a given set of inputs by adaptive signal processing [2, 3]. Many neural network architectures operate on real values but some applications may require the complex value inputs. Therefore, techniques such as Back-Propagation, Hopfield Model and Perceptron Learning Rules
are being used for complex value inputs. Their performances were tested using the pattern classification and time series experiments and its generalization capability was found to be satisfactory [4].

New types of complex-valued sigmoid activation function for multi-layered neural network was studied [5]. Their simulation results proved that their proposed network reduced 54% of testing time compared to a neural network that uses normal sigmoid activation function. In 1975, the complex Least Mean Squares (LMS) algorithm and Complex-valued neural network algorithm were published [6]. In the paper of Deville, Y [7], a complex activation function for digital Very Large Scale Integration (VLSI) neural networks was implemented. It was claimed that it required lesser hardware than the conventional real-valued neural network. In his theses, Prashant A [8], suggested that the input data should be scaled to some region in complex domain and to overcome the implementation problem, split sigmoid activation function could be trained for tracking the network. In a complex-valued neural network, inputs, output, threshold, and weights are complex values and selecting appropriate activation function is a challenging part. Here, a Smart Beam Forming Antenna using Single Neuron Model is presented to achieve the weight optimization of an Adaptive Array Antenna.

II. ADAPTIVE ARRAY MODEL

Adaptive Array Model is usually designed using a simple array of dipoles placed in a straight line. Array models have been tested with five and seven elements placed in the straight line. The array model equations are expressed in Equation (6) and (7) where \( f(\phi) \) is the desired beam function.

\[
w_1 e^{j\beta \cos \theta} + w_2 e^{j\beta \cos \theta} + w_3 + w_4 e^{-j\beta \cos \theta} + w_5 e^{-j2\beta \cos \theta} + w_6 e^{-j3\beta \cos \theta} = f(\phi) \tag{6}
\]

\[
w_1 e^{j2\beta \cos \theta} + w_2 e^{j2\beta \cos \theta} + w_3 + w_4 e^{-j2\beta \cos \theta} + w_5 e^{-j3\beta \cos \theta} + w_6 e^{-j\beta \cos \theta} + w_7 e^{-2j\beta \cos \theta} + w_8 e^{-3j\beta \cos \theta} = f(\phi) \tag{7}
\]

The dipoles arrangement can be made in any way since the current amplitude and the phase are adjustable to get the desired radiation patterns. Any arbitrary set of dipoles arranged in a straight line will produce a radiation pattern that is symmetrical on both sides of the plane where the dipoles are placed. Subsequently, the dipole placement must be chosen based on the desired radiation patterns. If the desired radiation patterns are symmetrical over a common axis then the dipoles can be placed in that common axis so that all of the current components are in phase. But with a different set of current amplitudes may be used to get respective radiation patterns. Otherwise, the dipole placement will not along a common axis and the current components will have different phases and amplitudes. Therefore, the in-phase and the different phase current components will result in real and complex optimized weight values, respectively. Hence, two types of activation functions are proposed to achieve the optimization of real and complex weights. To fulfill the objective of framing a resultant single beam, the optimization of the complex weights \( w_1, w_2 \), and \( w_3 \) should be done such that the resultant field must coordinate to a desired single beam function, \( f(\phi) \). Thus equation can be written as,

\[
w_1 e^{j(\beta x_1 \cos \phi + y_1 \sin \phi)} + w_2 e^{j(\beta x_2 \cos \phi + y_2 \sin \phi)} + \ldots + w_n e^{j(\beta x_n \cos \phi + y_n \sin \phi)} = f(\phi) \tag{8}
\]

Bodhe S.K et al. [9] have proposed a rectangular array structure to provide a solution for the condition when the desired radiation patterns are unsymmetrical on a common axis. However, arrays with a minimum of three elements that are not placed in common linear axis will produce the complex weight values.

III. SINGLE NEURON WEIGHT OPTIMIZATION MODEL (SNWOM)

The single neuron model is briefly discussed. It needs to optimize the weights which are used in adaptive beamforming [10-12]. The simple perceptron model has three parts and in the first part, the input signals \( x_1, x_2, \ldots, x_n \) are multiplied by the weights \( w_1, w_2 \ldots w_n \). The second part is the net function that sums all weighted inputs and bias as in (9). In the final part, the sum of weighted inputs and bias is passed through a transfer function to get the final output signal. A single neuron Perceptron is used and a nonlinear activation function \( \sigma \) is used to find out the output \( y \) as in (10) to simplify the calculation complexity and to reduce the processing delay. The deviation, \( \Delta \) is the error between the desired output, \( y_0 \) and the actual output, \( y \) as in (11). The weights are continuously adjusted in every iteration using the selected learning rate or coefficient, \( k_0 \) as shown in (12) until the Trained Means Error, TMR as in (13) is minimized below the predefined value \( TMR_{\text{r}} \) or until the defined maximum number \( N \) of iterations is reached. For training and during the testing process, different angles, \( \phi \) is used in the range of \( 0^\circ \) to \( 360^\circ \).

\[
z = b + \sum_{k=1}^{n} w_k x_k \tag{9}
\]

\[
y = \sigma(z) = \frac{1}{1 + e^{-z}} \tag{10}
\]

\[
\Delta = y_0 - y \tag{11}
\]

\[
w_j = w_j + (k_0 \times \Delta \times x_j) \tag{12}
\]

\[
TMR = \frac{\Delta}{y_0} \times 100 \tag{13}
\]

A. Complex form Activation Functions (AFs)

The crucial part of an ANN is the activation function used for limiting the amplitude of the output of a neuron called squashing functions [13]. It crushes the permissible amplitude range of the output signal to some finite value. This model is tested using different complex form activation functions including the Hyperbolic Tangent function, Bipolar sigmoid function and Squash and the Elliot Function.

IV. RESULT AND DISCUSSION

A. Real activation functions

Experiments are set up by placing the dipoles in a straight line. The desired beam function, \( f(\phi) \) is fixed as \( \cos 2\phi \) and the distance between two adjacent antenna elements is half wavelength. The weights are computed for five and seven element array antennas using Least Mean Square (LMS) optimization in order to compare the accuracy between the
weights optimized from SNWOM method [10] and the traditional LMS method. After the convergence is achieved, the radiation patterns for the optimized weights are obtained and compared with the desired beam for five element arrays as in Figure 3(a). From the result, it is clearly observed that the Perceptron output is almost matched with the desired beam. Figure 3(b) shows the radiated patterns after increasing the elements to seven. The perceptron output very closely matches the desired beam. Comparison of the beams shown in Figure 3 shows that SNWOM has a better output than the LMS method. The SNWOM beamforming methods are widely applied for communication system, for instance, at the junction of underground tunnels in mines or the streets with vehicles travelling along a four-corner junction with the smart antenna base station located at the junction. The error between the desired and the optimized beams corresponding to the angles are compared to study the accuracy differences between five and seven element smart antennas. The results are shown in Figures 4(a) and (b). Clearly, the increase in the number of elements has minimized the error range; but with high frequency of error oscillation. To test the precision of the SNWOM with a variety of activation functions, the function specified in (14) is selected:

\[
f(\phi) = \frac{1}{9}[3 + 4\cos(\pi\cos\phi) + 2\cos(2\pi\cos\phi)]
\]  

(14)

![Comparison of radiation patterns obtained from SNWOM and LMS methods](image)

The tests are run for both five and seven elements. The results generated by the perceptron and LMS are compared and shown in Figure 5(a) and (b). From the results, the SNWOM radiation beam is seen to be slightly inferior to the LMS beam and has larger side-lobes in the 0° and 180° directions as seen in Figure 5(b). It can be recognized that when the desired beam is narrow an increased number of dipole elements is required. Similar to results shown in Figure 6, comparison of error has shown that the range of error is reduced while the frequency of oscillation increases with the increase of the number of elements.

![Comparison of radiation patterns of SNWOM and LMS methods](image)

The work is extended to optimize the weights for three, four and six elements to study the actual output beam patterns using the SNWOM model with different initial weights, bias and learning rate and the appropriate activation function. From the results, it can be observed that as the number of elements increases the optimized beam patterns are better matched to the desired beam and the beam width is also reduced. A narrow beam would have a greater coverage while utilizing less power as compared to an Omnidirectional antenna.
B. Complex activation functions

Further tests on the precision of the SNWOM method were carried out with a variety of desired array factors, the desired array factor function selected is \( f(\phi) = \text{sinc}(\phi - \phi_0) \) in order to form a single desired beam, where \( \phi_0 \) is the desired angle. The beam steering towards desired angles of \( \frac{\pi}{3}, \frac{\pi}{2}, 2\frac{\pi}{3} \) and \( 5\frac{\pi}{3} \) using six elements were done with the bipolar sigmoid activation function. It is observed that the desired radiation patterns closely matched the Perceptron generated beam. The results in Figure 7 show the generated beams using three different complex activation functions and it is observed that the SNWOM works well in generating radiation beams that closely match the desired single beam. The experiments done using Single Perceptron shows that the technique is fast where the best set of antenna patterns can be provided within milliseconds. Although the generated beam precision depends on the dipole placement and the selected characteristics of the desired beam, it is a fast, efficient and simple method for the weight optimization and smart antenna beam generation. In the antenna based beamforming, the beamformer may handle both signal receiver and a single cluster of receivers in one geometrical location, or multiple clusters or antennas.

![Hyperbolic Tangent function](image1)

![Bipolar sigmoid function](image2)

Figure 6: The error between desired and optimized beam with the corresponding angle [11]

Figure 7: Radiation patterns for six elements using different activation functions [12]

V. CONCLUSION

The accuracy and the efficiency of adaptive beamforming problem have been tested for the new method proposed: the SNWON. The weights are optimized using appropriate real and complex activation functions. The results show that the performance is much better than LMS method where the radiation patterns obtained from optimized weights closely match the desired radiation patterns. The weight coefficients are complex values. From the error comparison, it can be concluded that the power loss in the main beam has been reduced by increasing the number of array elements and the beam-widths becoming closer to the desired beams. When the Perceptron is used to construct a broadside beam antenna, the increase in the number of elements gives a greater match with the two main beams on either side of the line along which the linear array is placed. Knowing that the Perceptron based linear antenna does not provide a single rotatable beam antenna and tends to produce back lobes as well, to obtain single beam Perceptron beamforming complex activation functions need to be used. The proposed Perceptron based beamformer may be implemented on any chip-based MIMO techniques, including transmitting beamforming, spatial multiplexing, space-time block coding and cyclic delay diversity.

ACKNOWLEDGEMENT

The work was supported by Faculty of Engineering, University Malaysia Sarawak and F02/FRGS/1488/2016 grant from the Ministry of Higher Education (MOHE), Malaysia.

REFERENCE


