

Optimized Artificial Neural Network Using Differential Evolution for Prediction of RF Power in VHF/UHF TV and GSM 900 Bands for Cognitive Radio Networks

Sunday Iliya, Eric Goodyer, Mario Gongora and Jethro Shell

Centre for Computational Intelligence,
School of Computer Science and Informatics,
De Montfort University, The Gateway,
Leicester LE1 9BH, England,
United Kingdom

Email: sundayiliyagoteng@yahoo.com, eg@dmu.ac.uk,
mgongora@dmu.ac.uk, jethros@dmu.ac.uk

John Gow

School of Engineering,
Media and Sustainable Development,
De Montfort University, The Gateway,
Leicester LE1 9BH, England,
United Kingdom

Email: jgow@dmu.ac.uk

Abstract—Cognitive radio (CR) technology has emerged as a promising solution to many wireless communication problems including spectrum scarcity and underutilization. The knowledge of Radio Frequency (RF) power (primary signals and/ or interfering signals plus noise) in the channels to be exploited by CR is of paramount importance, not just the existence or absence of primary users. If a channel is known to be noisy, even in the absence of primary users, using such channels will demand large quantities of radio resources (transmission power, bandwidth, etc) in order to deliver an acceptable quality of service to users. Computational Intelligence (CI) techniques can be applied to these scenarios to predict the required RF power in the available channels to achieve optimum Quality of Service (QoS). While most of the prediction schemes are based on the determination of spectrum holes, those designed for power prediction use known radio parameters such as signal to noise ratio (SNR), bandwidth, and bit error rate. Some of these parameters may not be available or known to cognitive users. In this paper, we developed a time domain based optimized Artificial Neural Network (ANN) model for the prediction of real world RF power within the GSM 900, Very High Frequency (VHF) and Ultra High Frequency (UHF) TV bands. The application of the models produced was found to increase the robustness of CR applications, specifically where the CR had no prior knowledge of the RF power related parameters. The models used implemented a novel and innovative initial weight optimization of the ANN's through the use of differential evolutionary algorithms. This was found to enhance the accuracy and generalization of the approach

I. INTRODUCTION

Most of the licensed radio spectral are grossly underutilized while the unlicensed spectral are overcrowded. The present spectrum scarcity is the direct consequence of spectrum allocation policy and not the fundamental lack of spectrum. The first bands to be approved for CR communication by the US Federal Communication Commission (FCC) because of their gross underutilization in time, frequency and spatial domain are the very high frequency and ultra-high frequency

(VHF/UHF) TV bands [1] [2]. This paper is focused on the study of real world RF power distribution in specific selected channels (54MHz to 110MHz, 470MHz to 670MHz, 890MHz to 908.3MHz GSM up-link, 935MHz to 953.3MHz GSM down-link) within the VHF/UHF bands and the GSM 900 band. In order to circumvent the problem of spectrum scarcity and underutilization, a new paradigm of wireless communication needs to be adopted. Advanced Cognitive Radio (CR) or Adaptive Spectrum Sharing (ASS) is one of the ways to optimize our wireless communications technologies for high data rates whilst maintaining user desired quality of service (QoS) requirements. CR is a radio equipped with the capability of awareness, perception, adaptation and learning of its radio frequency (RF) environment [3]; through which many of the digital signal processing that were traditionally done in hardware are implemented via software. Irrespective of the definition of CR, it has the followings basic characteristics: observation, adaptability and intelligence. CR is the key enabling technology for dynamic spectrum access and a promising solution for the present problem of spectrum scarcity and underutilization. Cognitive radio network consists of two users i.e. the primary users (PU) who are the incumbent licensed owners of the spectrum and the cognitive radio commonly called the secondary users (SU) who intelligently and opportunistically access the unused licensed spectrum on a no interference base i.e. when the licensed owners are not using the channel (the overlay) or SU can coexist with the PU as long as the interference to the PU is below a given threshold which will not be harmful to the PU nor degrade the QoS requirements of the PU (the underlay CR network scheme) [4] [1]. In this paper we combine the power of ANN and differential evolution (DE), and developed an optimized ANN model for the prediction of real world radio frequency (RF) power. The proposed algorithm used a priori data to enable the system to avoid noisy channels. The previous knowledge of the RF power allowed the cognitive radio to predictively select channels with the least noise among those that were unused

or free. This would allow for a reduced utilization of radio resources including transmitted power and bandwidth, in turn maximizing the usage of the limited spectrum resources. The data used in this study was obtained by capturing real world RF data for two months using Universal Software Radio Peripheral 1 (USRP 1). The digital signal processing and capturing of the data were done using gnuradio which is a combination of Python and C++ while the ANN models and prediction were done in Matlab. The experiment was conducted at the Centre for Computational Intelligence, De Montfort University, UK, located very close to Leicester city centre.

The rest of this paper is constructed of the following sections. Section II outlines previously presented research in this field focusing on the implementation of CI methods. This will be followed by Section III, section IV and V, that will present the methodology used within this research, the optimization algorithm and experiments carried out. The paper is concluded with Section VI, which discusses the results of the experiments, Section VII presents a summary of the work.

II. RELATED WORK

There are different types of Computational Intelligence (CI) and machine learning that can be used in CR such as genetic algorithms, fuzzy logic system (FLS), neural network, hidden Markov model, game theory, linear regression and linear predictors, Bayesian inference based predictors, swarm intelligence, etc. Some for learning and prediction, some for optimization of certain transmission parameters while others for decision making [5][6][7]. Spectrum hole prediction using Elman recurrent artificial neural network (ERANN) was proposed in [8]. It uses the cyclostationary features of modulated signals to determine the presence or absence of primary signals while the input of the ERANN consists of time instances. The inputs and the target output used in the training of the ERANN and prediction were modeled using ideal multivariate time series equations and assumptions which are often different in real life where PU signals may be embedded in noise and/ or interfering signals. The fact that spectrum holes were detected does not depict any information about the expected noise and/ or level of interference based on the channels history which is vital in selecting the channels to be used among the unused channels. Traffic pattern prediction using seasonal autoregressive integrated moving-average (SARIMA) was proposed for reduction of SU hopping rate and interference effects on PU while maintaining a fare blocking rate [9]. The model (SARIMA) does not depict any information about the expected noise power.

Transmission rate prediction using adaptive neural fuzzy inference system (ANFIS) was proposed in [10]. This model is used to predict the data rate (6, 12, 24, 36, 48 and 54 Mbps) that can be achieved in WLAN using a 802.11a/g configuration as a function of time. The training data was obtained by generating a random data rate with an assigned probability of occurrence at a given time instance, thereby forming a time series. In this paper, no real world data was used. Additionally, the research did not take into account the noise or interference level which can affect the predicted rates. Semi Markov model (SMM) and continuous-time Markov chain (CTMC) models have also been used for the prediction of

packet transmission rates [11]. This avoids packet collisions by periodic spectrum sensing and prediction of temporal wireless local area network (WLAN) activities combined with hopping to a temporary inactive channel. However, no reference is made to the expected noise level among the inactive channels that are selected.

Fuzzy logic (FL) is a CI method that can capture and represent uncertainty. As a result it has been used in CR research for decision making processes. An FL based decision system was modeled for spectrum hand-off decision-making in a context characterized by uncertain and heterogeneous information [12] and fuzzy logic transmit power control for cognitive radio. The proposed system was used for the minimization of interference to PU's while ensuring the transmission rate and quality of service requirements of secondary users [13]. The researcher did not, however, include any learning from past experience or historical data.

An exponential moving average (EMA) spectrum sensing using energy prediction was implemented in [14]. The EMA achieved a prediction average error of 0.2436 with the assumption that the channel utilization follow exponential distribution with rate parameter =0.2 and signal to noise (SNR) of 10dB; RF real world data was not used in their study.

Within this paper we demonstrate the use of an ANN for prediction of real world RF power of selected channels within the GSM band, VHF and UHF bands. An optimized ANN model was produced by combining the global search ability of differential evolution (DE) algorithms and the local search advantages of back-propagation algorithms (BPA). The initial weights of the ANN were evolved using DE after which the ANN was trained more accurately using back-propagation algorithms. This methodology demonstrates the application of previously acquired real world data to enhance the prediction of RF power to assist the implementation of CR applications.

III. METHODOLOGY

Artificial Neural Networks (ANN) are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems [15]. As a result of the nonlinearity often associated with RF traffic, coupled with erratic interfering signals or noise resulting from both natural and artificial sources, a fully connected multilayer perceptron (MLP) ANN with two hidden layers was used in this research. The input layer was cast into a high dimensional first hidden layer for effective features selection. A nonlinear hyperbolic tangent function was used for the activation functions of the two hidden layers while a linear symmetric straight line is used for the output activation function. Other activation functions were also used, but this combination gave the best results. These functions introduced a nonlinear transformation into the network. The hidden layers serve as a feature detector i.e. during the training they learn salient features that characterize the training samples. The ANN is trained using differential evolution (DE) and backpropagation algorithms (BPA). The synaptic weights and biases were updated every epoch to minimize the Mean Square Error (MSE). A supervised batch training method was adopted of which 60% of the data were used for training the ANN, 20% for validation and 20% for testing the trained ANN. Since the back propagation algorithm

is used as a local searcher, the learning rate was kept low at 0.01 for the model with constant learning rate while for the model with varying learning rate, the initial learning rate is 0.8 which changes with change in MSE. The inputs of the ANN consist of time domain data of varying rates of change i.e. second, minute, hour, week day (1 to 7), date day (1 to at most 31), week in a month (1 to 5), and month whilst the output gives the power in Decibels (dB). Each input enables the ANN to learn the trend of RF activities (RF power variation) as a function of that particular input. In this paper we designed three ANN models and evaluated their efficiency in terms of memory demand, training time, computational complexity and prediction accuracy/ generalization. The first model is shown in Fig 1; it consists of only one output neuron and is dedicated for power prediction of only one channel which implies that each channel will have its own dedicated ANN. The second model is depicted in Fig 2, it has only one output neuron but with one additional input representing the channels. The third model has output neurons equal to the number of channels to be considered, this is shown in Fig 3. The data used in this paper were obtained by capturing real world RF signals in the GSM 900, VHF and UHF TV bands and FM bands for a period of two months. In all the models, no RF power related parameters such as signal to noise ratio (SNR), bandwidth, and modulation type, are used as the input of the ANN. This makes the models robust for cognitive radio application where the CR has no prior knowledge of these RF power related parameters. Four ANN topologies were used: feed forward (FF), feed forward with output feedback (FFB), cascaded feed forward (CFF) and layered recurrent (LC) ANN.

The accuracy and level of generalization of ANN depend largely on the initial weights and biases, learning rate, momentum constant, network topology, and also the training data. In this paper, momentum was kept constant at 0.008 while the initial weights and biases were evolved using differential evolution variants i.e. DE/Rand/1/Exp and DE/Rand/1/bin. The initial weights and biases were randomly generated and constrained within the decision space of -2 to 2 using a toroidal bound. A population of 30 was maintained over 1000 generations. 30 parents and offspring were evolved every generation as described in the pseudocode, Section IV-B. The MSE obtained using the test data was used as our fitness function. The use of the test data (data not known by the ANN nor used to train it) for computation of the fitness function (MSE) does not only result in a more accurate network but also a more robust and generalized ANN model. The ANN weights and biases were initialized using each parent. The network is trained and the fitness of each parent is computed. This process is repeated using the offspring. If an offspring outperformed its parent, the particular parent is replaced by the offspring otherwise the parent will continue in the next generation. We use both binomial and exponential cross over for generating the offspring [16], with a cross over probability CR of 0.2. The mutation factor F was kept constant at 0.3. After 1000 generations, the most fit candidate, the global best (this is defined as the candidate with the least MSE among the population) was selected. The weights and biases of the ANN was reinitialized with the global best and retrained (fine tune) using backpropagation algorithms to produce the final model.

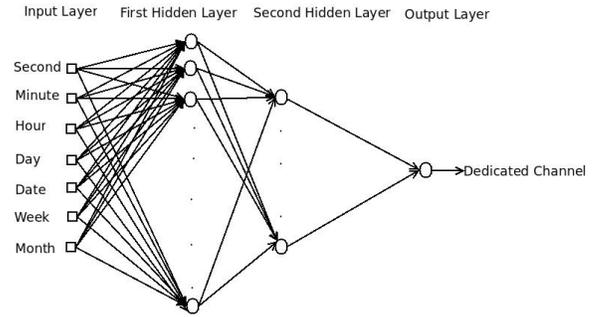


Fig. 1. Dedicated ANN model for one channel

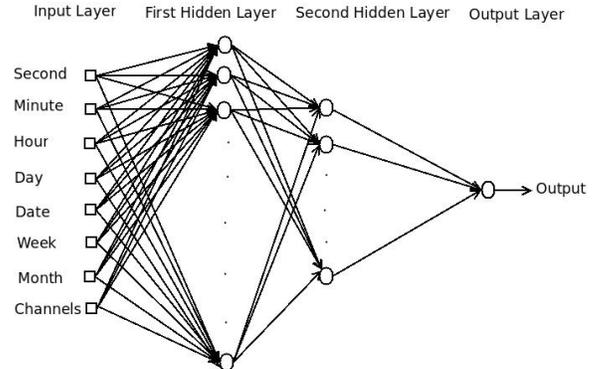


Fig. 2. Multiple channels, single output ANN model

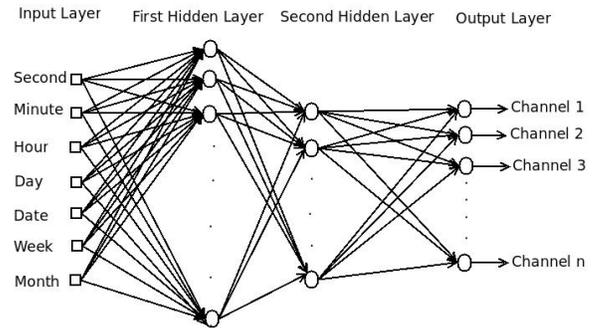


Fig. 3. Multiple channels, parallel outputs ANN model

IV. OPTIMIZATION

This section gives a brief description of the optimization algorithms implemented. We combine the global search capability of deferential evolutionary and algorithms with the local search advantages of BPA to evolve the weights and biases of the optimized ANN model as described in selection process, IV-B.

A. Differential Evolution (DE)

DE are population based direct search algorithms used to solve continuous optimization problems [17][18]. DE aims at evolving NP population of D dimensional vectors which encodes the G generation candidate solutions $X_{i,G} = \{X_{i,G}^1, \dots, X_{i,G}^D\}$ towards the global optimum, where $i = 1, \dots, NP$. The initial candidate solutions at $G = 0$ are

evolved in such a way as to cover the decision space as much as possible by uniformly randomizing the candidates within the search domain using (1).

$$X_{i,G} = X_{min} + rand(1, 0) \cdot (X_{max} - X_{min}) \quad (1)$$

Where $i = 1, \dots, NP$, $X_{min} = \{X_{min}^1 \dots X_{min}^D\}$, $X_{max} = \{X_{max}^1 \dots X_{max}^D\}$ and $rand(1, 0)$ is a uniformly distributed random number between 0 and 1, X_{min} and X_{max} are the minimum and maximum of X respectively.

1) *Mutation*: For every individuals (target vectors) $X_{i,G}$ at generation G , a mutant vector $V_{i,G}$ called the provisional or trial offspring is generated via certain mutation schemes. The mutation strategy implemented in this study is DE/rand/1 [2] [18].

$$V_{i,G} = X_{r_1,G} + F \cdot (X_{r_2,G} - X_{r_3,G}) \quad (2)$$

Where the indexes r_1 , r_2 , and r_3 are mutually exclusive positive integers and distinct from i while F is the mutation constant.

B. Pseudocode

```

Generate an initial population  $X_{G=0}$  of  $Np$  individuals.
Initialize ANN weights and biases with each solution in the
population  $Np$  and Evaluate fitness (MSE) of each using
the test data.
while termination condition is not met (Generation) do
  for  $i = 1$  to  $Np$  do
    Initialize ANN weights and biases with  $X_{i,G}$  (parent)
    and evaluate its fitness  $FitParent_i$  using the test data.
    Generate trial offspring  $V_{i,G}$  by mutation using (2).
    Generate offspring  $U_{i,G}$  by either binomial crossover
    for DE/rand/1/bin or exponential crossover for DE/rand/1/exp.
    Initialize ANN weights and biases with  $U_{i,G}$  (offspring)
    and evaluate its fitness  $FitOffspring_i$  using the test
    data.
  end for
  for  $i = 1$  to  $Np$  do
    Selection Process: Select the best between parent and
    offspring.
    if  $FitOffspring_i \leq FitParent_i$  then
       $X_{i,G+1} = U_{i,G}$ 
    else
       $X_{i,G+1} = X_{i,G}$ 
    end if
  end for
end while
Select the global best  $X_{gb}$  (i.e. individual with the best
fitness at the end of the budget (generation)), initialize the
ANN weights and biases with  $X_{gb}$  and train it (fine tune the
weights) using BPA constrain within a maximum of just 200
epoch and 6 validation fails. This yield the final optimized
ANN model.

```

V. EXPERIMENT AND SIMULATION DATA

The datasets used in this study were obtained by capturing real world RF signals using universal software radio peripheral 1 (USRP 1) for a period of two months. The USRP are computer hosted software radios with one motherboard and interchangeable daughter board modules for various ranges of

frequencies, these serve as the RF front end. Two daughter boards, SBX and Tuner 4937 DI5 3X7901, having a frequency ranges of 4MHz to 4.4GHz and 50 MHz to 860 MHz respectively, were used in this study. The daughterboard perform the function of down-conversion, filtering, analog to digital conversion (ADC) and other signal conditioning while the motherboard decimate the signal to a lower sampling rate that can easily be transmitted to the host computer through a high-speed USB cable where the signal is processed by software. For TV channels with channels bandwidth of 5MHz, we divided the channels into subchannels each consisting of 500 KHz bandwidth. To ensure that no spectral information was lost, we used a sample frequency of 1MHz and obtained 1024 samples for each sample time. For GSM 900 and FM band with a bandwidth of 200 KHz, we used 1MHz sample frequency and 512 samples for each sample time. The power was obtained using both the time and frequency domain data. For the frequency domain, after passing the signal through the channel filter, the signal was windowed using a hamming window in order to reduce spectral leakage. The stream of the data was converted to a vector and decimated. This is then converted to the frequency domain and the magnitudes of the bins were passed to a probe sink. The choice of probe sink is essential because it can only hold the current data and does not increase thereby preventing stack overflow or a segmentation fault. This allows Python to grab the data at run time for further analysis or to be saved. The power was not computed using gnuradio block rather the magnitude of the bins were saved so that it can be used for further analysis. The interval of time between consecutive sample data was selected at a random value between 5 seconds and 30 seconds. The choice of this range is based on the assumption that for any TV programme, FM broadcast or GSM call will last for between 5 to 30 seconds. In order to capture all possible trends, the time between consecutive sample data is selected at random within the given range instead of using regular intervals. For the VHF and FM band we captured RF signals from 54MHz to 110MHz and 470 to 670MHz for the VHF TV bands. For the GSM band, 62 down-link channels (935MHz to 953.3MHz) and 62 uplink channels (890MHz to 908.3MHz) were captured. The real world RF data was divided into three subsets, randomly selected with 60% used for training the ANN, 20% for validation and 20% for testing. The training or estimation data were the only known data sources used in training the ANN. The test data set was unknown to the network i.e. they are not used in training the network but are used in testing the trained ANN as a measure of the generalization performance of the ANN model. The ANN design and the simulation were carried out using Matlab while the capturing of the data and the signal processing were carried out using gnu-radio which is a combination of Python and C++.

VI. RESULTS

From the results depicted, the three ANN models really adapt to the trends of variation of the RF power with good generalization. Table II shows the average results obtained by testing each ANN models on different channels selected at random within the studied bands, this demonstrates that the dedicated ANN model designed for the prediction of RF power in only one channel has the least MSE and is 82.55% more

accurate than the EMA. The EMA was tested on data obtained from mathematical models which are often different from real world data. In this study, real world RF data was captured and used for the training and prediction yet it outperformed the EMA. The results depicted in Table II, Fig 4, Fig 5 and Fig 6 where obtained using the test data i.e. data that were not used in training the ANN nor known by the ANN, this implies that the models are well generalized. One of the drawbacks of the dedicated model is that each channel will have a separate ANN. To circumvent this problem, we developed two ANN models that can handle multiple channels, one with a single output while the other with parallel outputs. The single output gives the RF power of many channels but one at a time, it has an additional input for identification of the channels and only one output neuron. The one with parallel output gives the RF power of all the channels it has been trained to predict simultaneously, given a time domain input. The number of channels is equal to the number of output neurons. The multi-channel models were trained to adapt to RF power of 20 channels and the results is as shown in Table II. In terms of accuracy, the parallel output is 19.69% more accurate than the single output but it demands more computer resources and computational complexity as it has many neurons than the single output model. The average MSE of the single output model is 0.0706 and that of the parallel model is 0.0567 for the 20 channels. The single output model and the parallel model are 71.01% and 76.72% more accurate than EMA respectively. The results discussed so far are based on the optimized ANN. Table II shows the improvement of the optimized ANN over the unoptimized model of the same topology. To validate the fact that the combination of these variants of DE optimization algorithms with the famous BPA to evolves the weights and biases of ANN will produce a more accurate, robust and generalized model than using only BPA; we use the same topology but train with only BPA at constant learning rate of 0.01 and another one with varying learning rate starting from 0.8 and keep on changing with change in MSE using inbuilt Matlab training function traingda. For both models trained with only BPA, each was run 30 times, each run was constrained within a maximum of 1200 epoch and 6 validation fails, the average MSE (AMSE), standard deviation (STD) of the MSE, and the Wilcoxon test results are as shown in Table I. For the hybridized training i.e combining DE with BPA, the DE were run for 1000 generations and the final best solution was used to reinitialized the ANN weights and further train using BPA constrained within just 200 epoch and 6 validation fails. In all cases, the hybridize training outperform the training with only BPA. From the results, the combination DE/rand/1/bin and BPA yield a more promising results compare to DE/rand/1/exp for this problem, hence it is used as our reference for the Wilcoxon test [19], where + mean that the reference outperformed the other algorithm it is compared with as depicted in Table I. The Wilcoxon test shown in Table I was obtained using the ANN model depicted in Fig 1. When the ANN was trained with only DE/rand/1/bin, the AMSE is 0.2685, after it was fine tune with BPA, the AMSE reduces to 0.0338 which is better than 0.0441 obtained using only BPA. This improvement demonstrate that the prediction accuracy and generalization of ANNs largely depend on the initial weights and biases. This also implies that, the generalization and accuracy of ANN can be improved by combining the global search advantages of DE and the local

search advantages of BPA.

From our study, we discovered that the FM broadcast band is grossly underutilized. Only two channels of the entire band seem to be used in the geographic location of our study except for possible low power devices that may exist within the band. Drawing from our findings, we can conclude that the FM band is a potential candidate band for CR communication. The noise floor of the FM band seems to be high. Alternatively, the GSM band is heavily used. 124 GSM channels were considered with 62 channels each for uplink and downlink. The improvement of the optimized ANN model over the unoptimize model is computed using (3).

$$Improvement = \frac{AMSE_U - AMSE_O}{AMSE_U} \cdot 100 \quad (3)$$

Where: $AMSE_U$ is the average mean square error of the unoptimized ANN model and $AMSE_O$ is the average mean square error of the optimized ANN model

TABLE I. WILCOXON TEST RESULTS WITH DE/RAND/1/BIN AS REFERENCE

Algorithms	Algorithms		Algorithms + BPA	
	AMSE	STD	AMSE	STD
DE/rand/1/bin	0.2685	0.0213	0.0338	(0.0024)
DE/rand/1/exp	0.3296	0.0453	0.0345	(0.0032+)
BPA (constant learning rate)			0.0441	(0.0055+)
BPA (varying learning rate)			0.2508	(0.0332+)

TABLE II. PERFORMANCE OF THE ANN MODELS

	ANN Models		
	Dedicated one channel	Multiple channels, single output	Multiple channels, parallel output
First Hidden Neurons	10	30	30
Second Hidden Neurons	5	25	25
Output Neurons	1	1	20
Number of Channels	1	20	20
Optimized Average MSE	0.0425	0.0706	0.0567
Unoptimized average MSE	0.0607	0.1067	0.0688
IOUA %	70.01	33.84	28.43
IOE %	82.55	71.01	76.72

Where IOUA is Improvement Over Unoptimized ANN and IOE is Improvement Over EMA

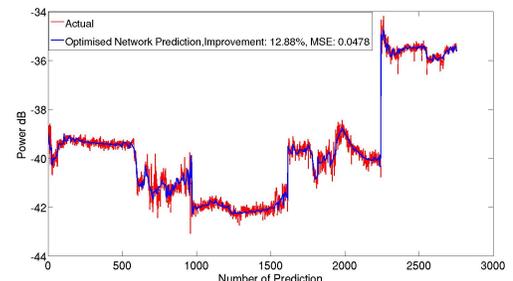


Fig. 4. Feed forward dedicated model prediction

VII. CONCLUSION

This paper demonstrate the application of an enhanced ANN through the use of DE and BPA in comparison to using

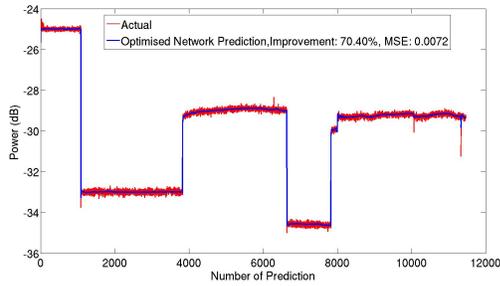


Fig. 5. Cascaded feed forward, parallel output model prediction

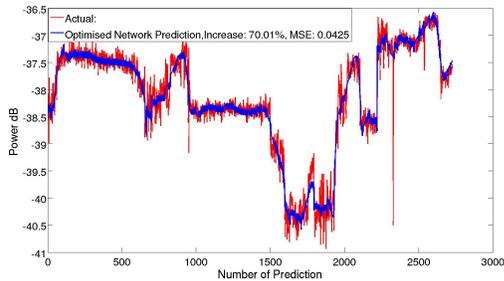


Fig. 6. Multiple channels, single output model prediction

only BPA, and EMA; for the prediction of real world RF power. The paper also authenticated the fact that the accuracy and generalization of ANNs largely depends on the initial weights and biases. The prior knowledge of the RF power can enhance the ability to avoid noisy channels. This can lead to an optimum satisfaction of users through higher QoS with less radio resources. Also, the prior knowledge of the aggregate RF power will lead to optimum and efficient spectrum sharing and spectrum mobility in both overlay and underlay CR schemes.

ACKNOWLEDGMENT

This work is partly supported by Petroleum Technology Development Fund (PTDF) Scholarship, Nigeria and De Montfort University, United Kingdom.

REFERENCES

- [1] FCC, "Federal communication commission notice of inquiry and notice of proposed rule making, in the matter of establishment of an interference temperature metric to quantify and manage interference and to expand available unlicensed operation in certain fixed, mobile and satellite frequency bands," no. 03-237, November 13, 2003.
- [2] S. Haykin, D. J. Thomson, and J. H. Reed, "Spectrum sensing for cognitive radio," in *IEEE Transactions on Cognitive Radio*, May 2009.
- [3] J. Oh and W. Choi, "A hybrid cognitive radio system: A combination of underlay and overlay approaches," in *IEEE Transactions on Cognitive Radio*, 2009.
- [4] C. Stevenson, G. Chouinard, Z. Lei, W. Hu, J. Stephen, and W. Caldwell, "The first cognitive radio wireless regional area network standard," in *IEEE 802.22*, 2009.
- [5] R. Azmi, "Support vector machine based white space predictors for cognitive radio," Master's thesis, 2011.
- [6] J. Proakis and D. Manolakis, *Digital Signal Processing Principles, Algorithms, and Applications*, 4th ed. Upper Saddle River, New Jersey 07458: Pearson Prentice Hall, 2007.

- [7] J. Proakis and M. Salehi, *Digital Communications*, 5th ed. Mcgraw-Hill, 2008.
- [8] M. I. Taj and M. Akil, "Cognitive radio spectrum evolution prediction using artificial neural networks based multivariate time series modelling," in *European Wireless, Vienna Austria*, April 2011.
- [9] X. Li and S. A. Zekavat, "Traffic pattern prediction and performance investigation for cognitive radio systems," in *IEEE Communication Society, WCNC Proceedings*, 2008.
- [10] S. Hiremath and S. K. Patra, "Transmission rate prediction for cognitive radio using adaptive neural fuzzy inference system," in *IEEE 5th International Conference on Industrial and Information Systems (ICIIS), India*, Aug 2010.
- [11] S. Geirhofer, J. Z. Sun, L. Tong, and B. M. Sadler, "Cognitive frequency hopping based on interference prediction: Theory and experimental results," vol. 13, no. 2, march 17, 2009.
- [12] L. Giupponi and A. Perez, "Fuzzy-based spectrum handoff in cognitive radio networks," 2008.
- [13] Z. Tabakovic, S. Grgic, and M. Grgic, "Fuzzy logic power control in cognitive radio," in *IEEE transactions*, 2009.
- [14] Z. Lin, X. Jian, L. Huang, and Y. Yao, "Energy prediction based spectrum sensing approach for cognitive radio network," in *978-1-4244-3693-4/09/\$25.00 ©2009 IEEE*, 2009.
- [15] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed., 2008.
- [16] D. Zaharie, "A comparative analysis of crossover variants in differential evolution," in *Proceedings of the International Multiconference on Computer Science and Information Technology*, 2007, pp. 171–181.
- [17] K. V. Price, R. Storn, and J. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*. Springer, 2005.
- [18] A. K. Qin, V. L. Huang, and P. N. Suganthan, "Differential evolution algorithm with strategy adaptation for global numerical optimization," in *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 2, April 2009.
- [19] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945.