Bulletin of Electrical Engineering and Informatics
Vol. 13, No. 6, December 2024, pp. 4079∼4091
ISSN: 2302-9285, DOI: 10.11591/eei.v13i6.7262

An ANN enabled joint power allocation and base station switching system for EE heterogeneous networks

Shornalatha Euttamarajah[1], Yin Hoe Ng[1], Chee Keong Tan[2]
[1]Faculty of Engineering, Multimedia University, Cyberjaya, Malaysia
[2]School of Information Technology, Monash University, Bandar Sunway, Malaysia

Article Info

Article history:

Received Jul 27, 2023
Revised Jun 14, 2024
Accepted Jun 26, 2024

Keywords:

Coordinated multi-point
Energy efficiency
HetNets
Hybrid-power
Neural network

ABSTRACT

In recent years, dynamic and complex development in wireless communication in network models or environments led to more tedious and complicated resource management issues (i.e., power allocation and base station switching (BSS)). Conventional solutions often suffer from delays and degraded network service quality. Due to the ability of machine learning in analyzing huge volumes of data and automatically adapt to environmental changes, it emerges as a highly sought-after technique. In this work, we propose a machine learning approach based on feed-forward neural network to predict the active BS sets and estimate the power allocation to each UE within the active BSs for energy-efficiency (EE) maximization of a coordinated multi-point (CoMP-enabled) cellular system with hybrid-powered transmitting nodes in a HetNet-based architecture. By training the neural network model efficiently using a regression-based supervised learning technique that employs various backpropagation algorithms, almost similar EE performance (less than 5% difference) can be achieved with significantly reduced computational complexity and delay compared to the traditional methods, such as the well-known dual decomposition and brute force techniques. The effects of various hyper parameters and back-propagation algorithms are also investigated. Our results demonstrate that the proposed framework is a promising solution for establishing a fully green and intelligent network.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Yin Hoe Ng
Faculty of Engineering, Multimedia University
Cyberjaya, Malaysia
Email: yhng@mmu.edu.my

1. INTRODUCTION

Lately it has been witnessed that the number of commercially deployed 5G systems is increasing with the densification of base stations (BS) by deploying heterogeneous networking systems, particularly in densely populated urban areas. HetNet system is an emerging new network architecture [1], which deploys various BSs endowed with diverse transmission power capacities to provide seamless coverage with varying cell sizes (e.g., pico-cell, micro-cell, femto-cell, and macro-cell). The unprecedented BS densification through HetNet system deployment has led to two main issues: i) the scarcity of spectrum and ii) the increased energy consumption. The overlapping areas due to different cells in the HetNet system have led to the prohibition of frequency reuse caused by the co-channel interference, resulting in spectrum scarcity. Furthermore, a recent report [2] also reveals that about 60-80% of the consumed power in a cellular system is primely contributed by the BS.

To enable a massive rollout of HetNet systems for 5G networks, making the HetNet system spectrally- and energy-efficient (EE) not only helps to attain long-term sustainability and profitability for service providers but also brings a positive environmental impact.

Firstly, the spectral issue in the HetNet system gets alleviated by adopting a coordinated multi-point joint-processing (CoMP-JP) approach [3]. The channel state information (CSI) of the BSs can be shared in a CoMP-JP enabled HetNet to ensure coordination and synchronization among different transmission points to reject inter-cell interference. Secondly, the energy consumption problem in the HetNet system can be alleviated by implementing a BS switching (BS-Sw) technique [4], which is designed to shut down the BSs with low spectrum utilization to save energy and allows the spectrum to be used by other cells. In this work, the spectral and EE problem will be considered in a HetNet system equipped with energy harvesters, in which the harvested energy (HE) is stored and distributed among different BSs via a smart grid [5]. With the deployment of the smart grid, BSs can transfer excess energy to those that are experiencing an energy shortage. As a measure to uphold the growth in wireless communication and efficiently solve the problem, this study adopts a trending and notable field of artificial intelligence (AI) [6].

Recently, machine learning (particularly neural networks) has started making inroads into HetNet systems to solve various problems such as power assignment, channel allocation, user association, interference rejection and other resource management problems [7], [8]. The application of artificial neural network (ANN) in wireless communications has garnered tremendous interest from researchers and industry players because ANN offloads the computational burden on the real-time decision-making process in a resource management problem by shifting the computation to the training phase of ANN. In other words, the ANN can gradually learn the relation between the data input and output and develops a mapping function or network that can immediately predict the future output based on the newly fed input [9]. Once an ANN model is well-trained with comprehensive data, it can provide an immediate decision based on the input to the model, which is eagerly desired by the current HetNet system and 5G networks to offer instantaneous and real-time services to the user equipment (UE).

The adoption of ANN in wireless communication has been prevalent in the last decade. The necessity of ANN in future communication networks is discussed in detail [10]. This work highlights the influential factors of ANN that will boost the wireless network performance, like intelligent data analysis and prediction, powerful and smart data-driven network optimization, improved physical layer functions and robust user-centric services. The intricate associations between the observed input and output data can be learned by an ANN. The study by Zappone *et al.* [11] points out the remarkable features of ANN that outshine other classic machine learning approaches and detailed discussion on the integration of ANN into future communication networks.

ANN is used to learn the optimal transmit power allocation computed through the weighted minimum mean square error (WMMSE) algorithm [12]. The proposed scheme is proven to achieve similar or higher performance than the complex conventional iterative approaches that incur higher computational time. The study conducted by Ahmed *et al.* [13], utilizes ANN to learn genetic algorithm (GA) based power allocation and sub-band solutions for throughput maximization problems. Results are almost accurate with the GA model used for training but with minimal computational complexity. However, the work presented in [12], [13] do not incorporate BS-Sw into their frameworks.

EE transmit power is predicted by feeding the communication channel gains as the input to the trained network [14]. Deep ANN technique is utilised for the network training and is proven to achieve similar performance as the traditional optimization optimization method but with less complexity. However, complexity may arise during the initial training phase for this model. The study by Liang *et al.* [15] proposes a deep neural network which aids direct sum rate maximization in the time of training process and obtain the power allocation vector based on a unique unsupervised learning approach. Once again, the use of deep ANN is proven to outperform the conventional method at reduced computational complexity. Nevertheless, neither the technique in [14] nor [15] takes into account of BS-Sw.

In this study, we establish a novel HetNet system which is able to switch off a set of BSs and efficiently allocate the transmit power to the UEs which leads to the EE maximization of the system. The power allocation and BS-Sw strategies depend highly on the count of users served by every cell, energy harvesting capability and the energy cooperation mechanism enforced by the smart grid. In this paper, the formulated EE maximization problem is a non-deterministic NP-hard problem where finding an optimal solution requires an exhaustive search (either using the Brute-force method [16] or the dual decomposition approach [17]) which is computationally complex. Despite their optimality, the conventional methods incur high computation complexity and delay especially when they are applied in a complex environment. Therefore, the ANN architecture is adapted to the HetNet environments to learn to predict the PA and BS-Sw outputs.

The EE maximization problem is first solved using the Brute-force method [16], where the optimal power allocation and BS-Sw outcomes that maximize the EE are obtained for different HetNet scenarios. The dataset acquired from the Brute-force method is utilised to train and subsequently test the proposed ANN-based PA and BS-Sw algorithm. In this paper, a feed-forward neural network-based machine learning approach is proposed to intelligently predict the BS on/off status and power allocation with the aim to maximize the EE of a HetNet based cellular system, where the ANN is used to predict the active BS sets and estimate the power allocation to each UE within the active BSs. The surplus HE of the BSs are distributed among each other according to individual energy demand through a centralized smart grid. To the best of our understanding, this is regarded as the first work to employ an AI based approach to maximize EE of a CoMP-enabled HetNet with hybrid-powered BSs by performing a joint BS-Sw and transmit power allocation. The prime contributions of this work are highlighted as follows:

- Develop a feed-forward ANN structure to learn the BS-Sw and UE PA techniques which are established to maximize the EE of the HetNet system;
- Design a low-complexity and less-delayed training model and optimize the hyper-parameters of the proposed ANN-based BS-Sw and PA scheme to obtain the best EE performance;
- Analyze and prove the worth of employing a neural network compared to the conventional methods for a complex HetNet in terms of computational complexity.

The structure of the paper is outlined as follows: section 2 delineates the modelling of a HetNet system and the formulation of the EE maximization problem. In section 3, the proposed ANN architecture is presented, and the algorithm of the ANN-based BS-Sw and PA scheme is developed. Next, extensive simulation is performed in section 4 to thoroughly evaluate the performance of the proposed framework and examine the effects of various hyperparameters and system parameters. To assess the practicality of the proposed scheme, the proposed ANN-based PA and BS-Sw scheme is compared to the Brute-force method with the optimal settings. Finally, section 5 summarizes the key findings, discusses their implications, and suggests directions for future research.

2. MATERIALS AND METHODS

Consider a downlink HetNet system with J small cells under laid on a macro-cell with a BS placed on the centre of each cell. The BS placed in the macro-cell is known as MBS serving M number of UEs while the BSs placed in small cells are identified as SBSs serving N number of UEs. The cell radii of macro-cell and the small cells is denoted as d_{macro} and d_{small}, respectively such that $d_{macro} > d_{small}$. Each BS is equipped with a renewable energy harvesting source with different harvesting capability. At the same time, the BSs are linked to a centralized smart grid which can supply grid power and exchange the excess HE through the power line. It is also assumed that the BSs are CoMP-enabled and they can autonomously coordinate among each other in terms of communication to avoid interfering the UEs in different cells. The HetNet system with multiple small cells underlaid by a single macro-cell is demonstrated in Figure 1.

For simplicity, the list of variables and parameters used in this work are tabulated in Table 1. In this paper, R_i and P^{Total} represents the data rate of user i and the total consumption of grid power, respectively. These terms are expressed as (1) and (2):

$$R_i = B_o \log_2 \left(1 + \frac{\sum_j a_{ji} b_j p_{ji} g_{ji}}{B_o N_o}\right) \tag{1}$$

$$P^{Total} = \sum_j \left[\frac{b_j}{\lambda_j} \sum_i a_{ji} p_{ji} + b_j P_j^{cir} - b_j \rho E_j\right] \tag{2}$$

where a_{ji} and b_j indicate Boolean values where a_{ji} is the user association indicator such that $a_{ji} = 1$ if user i is affiliated with BS j and vice versa while b_j is the BS-Sw indicator such that $b_j = 1$ if BS j is in on mode and vice versa. The total power consumed by a BS mainly depends on the power used for downlink transmission from the BS to all served UEs and the circuit power used by the BS. Naturally, the power consumed by a

BS increases linearly with the number of UEs that the BS serves and also depends on the locations of UEs within the cell. If the UEs are situated at the cell edge, the BS needs to allocate more power for the downlink transmission to attain a minimum quality of service (QoS) (or minimum data rate) for all UEs. On the other hand, from (2), it is seen that the total power consumed by the BS gets compensated by the HE either harvested by itself or channelled by other BSs via a smart grid.

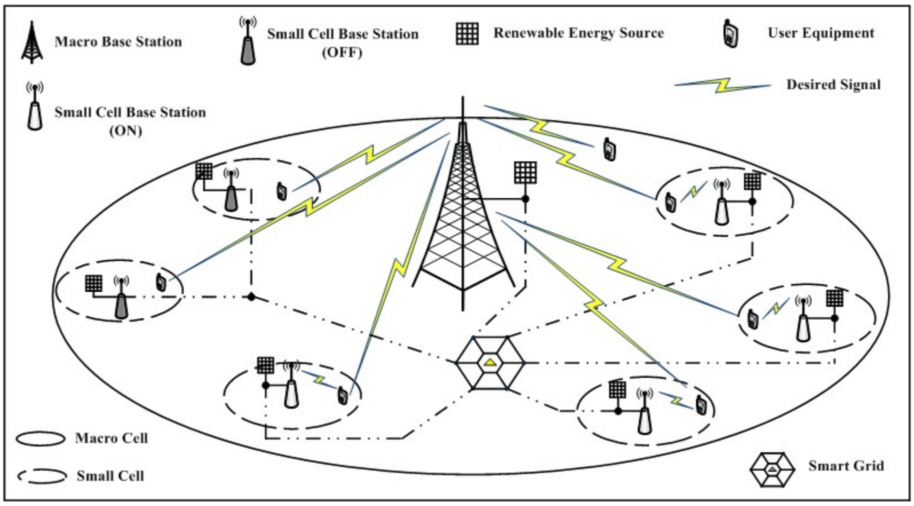


Figure 1. Model of CoMP-enabled HetNet structure with hybrid-powered MBS and SBSs

Table 1. Summary of major parameters and variables

Notations	Descriptions
R_i	Data rate
P^{Total}	Total on-grid power consumption
B_o	Bandwidth of a channel
p_{ji}	Transmit power from BS *j* to user *i*
N_o	Noise power
g_{ji}	Channel gain from BS *j* to user *i*
E_j	Harvested renewable energy at BS *j*
p_j^{cir}	Cicuit power of BS *j*
ρ	Efficiency of transferred energy
λ_j	Efficiency of power amplifier

In this work, the main objective is to develop a HetNet system that provides optimal PA and BS-Sw strategy that maximizes EE which is equated as the ratio of total system throughput to the total consumption of grid power. The maximization of the objective function (EE) can be defined as (3)-(3d):

$$\max_{\{p_{ji}\},\{b_{j\epsilon J\setminus\{1\}}\}} EE = \frac{\sum_i R_i}{P^{Total}} \tag{3}$$

$$s.t. \sum_i a_{ji} b_j p_{ji} \leq b_j P_{j,max}, \quad \forall j, \tag{3a}$$

$$\frac{\sum_j a_{ji} b_j p_{ji} g_{ji}}{B_o N_o} \geq SNR_{i,thr}, \quad \forall i, \tag{3b}$$

$$\sum_j a_{ji} b_j \geq 1, \quad \forall i, \tag{3c}$$

$$\frac{b_j \rho E_j}{\sum_i a_{ji} b_j p_{ji}} \geq E^R_{j,thr}, \quad b_j = 1 \tag{3d}$$

where $b_1 = 1$, which indicates that the MBS is always "on" all the time. In the maximization of the EE, each BS autonomously allocates their transmission power to each user if the BS chooses to be "on"; otherwise, the BS will allocate zero powers to all users within its cell if the BS decides to switch off. The power transmitted by the BS is limited to $P_{j,max}$ by constraint 3(a). Next, in order to guarantee each user to achieve a minimum SNR, constraint 3(b) sets a threshold $SNR_{i,thr}$. Followed by constraint 3(c) which ensures that at least one BS serves each UE. Furthermore, BS-Sw condition is enforced by constraint 3(d) by determining the ratio of HE to the power transmitted by BS. Only if the ratio is atleast the threshold, $E^R_{j,thr}$ value, the SBS will be permitted to be turned on [18]. This way, the less harvesting SBS will be switched off.

3. ARTIFICIAL NEURAL NETWORK BASED ALGORITHM

EE maximization is a non-deterministic NP-hard problem and the time consumed by NP-hard optimization is quite high. However, the environment of wireless communication (i.e., number of BS, number of UE, and pathloss) is also becoming more complex and any operational delays caused by the complexity is intolerable. Therefore, utilizing a smart AI based system such as ANN quickly adapts to the change in the environment without sacrificing the EE.

3.1. Feed-forward neural network

Feed-forward neural network (FFNN) is a simple and prominent ANN method used if the achievable results are already known. As the output of the joint BS-Sw and UE power allocation technique is already established through the conventional brute force and dual decomposition approaches in earlier work [18], the simulation results for various scenarios (i.e., different number of SBSs and UEs distribution) can be collected and utilised for the training of the FFNN to match the performance at reduced complexity. In FFNN, firstly, the inputs for training are fed forward. Secondly, the error is calculated based on the output and backpropagated. Finally, the weights are adjusted accordingly. Several factors must be considered while choosing the model of the FFNN. The prime focus is to achieve higher accuracy for a new set of data (different from the training data). Next, the computational complexity of the training should be low while maintaining a reasonable MSE.

Notably, the training computational complexity of FFNN increases as the hidden layers and/or hidden neurons are increased. Apart from that, using a complex backpropagation algorithm and reducing the learning rate increases the computation complexity of the network. Therefore, the values of these factors have to be analyzed and set efficiently to achieve lower computational complexity.

Figure 2 illustrates the process of FFNN, where p_j and w_j denote the inputs and weights fed into the network, respectively. The bias of the network is depicted by z while and a_1 indicate the activation function and output, respectively. Weight is a parameter associated with the input to determine the emphasis of each input to the output a_1. Bias is used to adjusting the estimation accuracy of the network based on the training error. The weighted inputs and bias are summed, as shown in (4a) before passing the activation function. Next, the output can be generated by activating the summed value n using a linear or non-linear activation function, depending on the relationship between the input and output. In this case, due to the non-linear relationship between the inputs and outputs, a non-linear logistic sigmoid (log-sigmoid) activation function is used, as expressed in (4b).

$$n = \sum_{j=1}^{J} w_j p_j + z \tag{4a}$$

$$a_1 = F(n) = \frac{1}{1 + e^{-n}} \tag{4b}$$

The output is expressed as (5):

$$a_1^{m+1} = F^{m+1}(w^{m+1} a_1^m + z^{m+1}) \tag{5}$$

where a_1^m and $a_1^{(m+1)}$ are the outputs of mth and $(m+1)th$ layers for $m = 0, 1, ..., (M$ - 1) where M is number of layers. For instance, if there are two layers in the network, 1 hidden layer and 1 output layer, the output is represented by a_1^2 as shown in (6). $w^{(m+1)}$ denotes the weight at the $(m+1)th$ layer.

$$a_1^2 = F^2\Big(\sum_{i=1}^{S} w_{1,i}^2 F^1\Big(\sum_{j=1}^{J} w_{i,j}^1 p_j + z^1\Big) + z^2\Big) \tag{6}$$

where S represents the number of neurons, ^{1} and ^{2} represents the activation functions of the hidden layer and output layer, respectively, and z^1 and z^2 refers to the bias of the neuron in the hidden layer and the bias of the neuron in the output layer, respectively. On the other hand, the weight that connects ith neuron to jth input is represented by $w^1_{(i,j)}$. Whereas, the weight that connects the ith source of hidden layer to the output layer is represented by $w^2_{(1,i)}$.

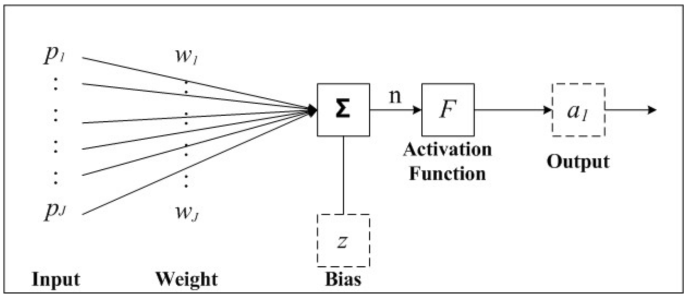


Figure 2. Interpretation of output acquisition through FFNN [19]

The mean square error (MSE) is the performance index applied by the back propagation algorithm, which is reduced by tuning network parameters using the following expression:

$$f(x) = E[e^T e] = E[(t - a_1)^T (t - a_1)]$$

where t is the target output, x is the vector matrix of network weights and biases and e is the error, $t - a_1$. In other words, the bias and weights of the network are tuned based on the discrepancy between the targetted output and output to minimize the error.

3.2. Generating dataset

As mentioned in section 3.1, the dataset is generated from work in [18]. For different scenarios of BS position and UE distribution within a macro cell, the BS on/off status and UE power allocation are collected in array form. This array is then used to train the FFNN, as explained in the next section. It is noteworthy that the threshold of the constraints is set while collecting the data itself, and during the training, the constraints fall within the set threshold values. The collected set of arrays is trained to determine the BS on/off status and the user power allocation.

3.3. Proposed algorithm

Figure 3 illustrates the FFNN used in this work, while the inputs and outputs of the FFNN are tabulated in Table 2. If g_{ji} (for $j > 1$) of the network is 0, eventually, the user is served only by MBS. Hence, p_{ji} (for $j > 1$) of network for this case would be 0 as well, indicating that there is no power allocation from SBS. Whereas the value of g_{1i} of the network will never be 0 as MBS will always be switched on. It is noteworthy that the network is a regression-based neural network model. Furthermore, the threshold which determines the binary output of BSS is set to 0.5, where any value less than 0.5 sets the BSS status as 0 and vice versa.

The number of neurons, S and the number of hidden layers, M are set according to preference during the training phase. Furthermore, J and I represent the total count of BSs and users in the system, respectively. Figure 3 can be formulated as (7) and (8):

$$For M = 1 : Y_k = F^2\Big(\sum_{i=1}^{S} w^2_{1,i} F^1\Big(\sum_{j=1}^{(J-1)+2I} w^1_{i,j} X_k + z^1\Big) + z^2\Big), k = 1, 2, ..., (J-1) + 2I \quad (7)$$

$$For M > 1 : Y_k^{m+1} = F^{m+1}(w^{m+1} Y_k^m + z^{m+1}, k = 1, 2, ..., (J-1) + 2I \quad (8)$$

It is worth noting that the weight w and bias z of each output varies accordingly during the training phase to produce a best-matching outcome. X_k and Y_k represents the inputs and outputs of the FFNN network respectively as explained in Table 2. The gain between the transmitters and receivers is taken into consideration as the inputs because these factors play a vital role in determining the BS transmit power [20] which eventually evaluates the throughput, power consumption and EE of a system as can be seen in section 2. Additionally, the distance between the MBS and SBS is used as part of the inputs for the neural network to enhance the determination of the BS on/off status.

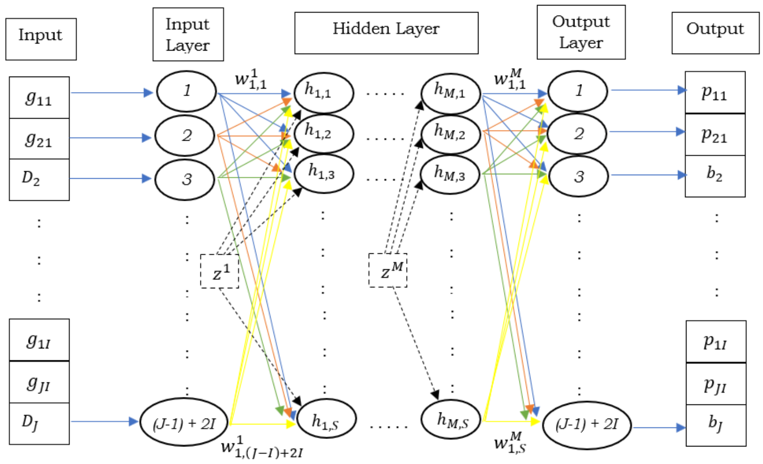


Figure 3. Feedforward neural network utilized to learn BS status and UE power allocation of the CoMP-enabled HetNet

Table 2. List of inputs and outputs of a network of the feed-forward neural network

Inputs/outputs	Notation	Description
Inputs $(X_k), k = 1, 2, ..., (J-1) + 2I$	g_{1i}	Gain between MBS and user $i(\forall i)$
	g_{ji}	Gain between SBS j and user $i(\forall i, j > 1)$
	D_j	Distance between MBS and SBS $j(j > 1)$
Outputs $(Y_k), k = 1, 2, ..., (J-1) + 2I$	p_{1i}	Power from MBS to user $i(\forall i)$
	p_{ji}	Power from SBS j to user $i(\forall i, j > 1)$
	b_j	BS j on/off status $(j > 1)$

First of all, the relation of the input and output parameters is derived through training phase of the neural network model. For the FFNN, the gain between the BS and users, and the distance between the MBS and SBS are fed as inputs and the power between power between the BS and the users, and the BS on/off status are fed as outputs to train the model accordingly as explained in section 3.1. A back propagation algorithm is used in the training process to minimize the MSE as explained in section 3.1. Levenberg-Marquardt (LM), Bayesian regularization and scaled conjugate gradient are some of the prominent and effective back propagation algorithms for FFNN. The error minimization of LM follows the well-known Gauss-Newton algorithm except for the Hessian matrix computation which is replaced with a very less complex Jacobian matrix [21]. The weights and bias values of Bayesian regularization (BR) are updated similar to the LM algorithm, except for the weights distribution which is related to conditional and marginal probabilities [1]. In common back propagation algorithm, weights adjustment performed based on steepest descent direction whereas in scaled conjugate gradient, conjugate gradient direction is employed to achieve much faster convergence [2]. The neural network based Algorithm 1 can be divided into two phases, training and implementation:

Phase A: training

– Learn BS on/off status and UE power allocation through FFNN.
  a. As mentioned above, the weight and bias of each equation act as variables which are adjusted accordingly to match the input and output.
  b. After many trials and errors, the best weight and bias are fixed for the FFNN equation.

Phase B: implementation

– Evaluate the environment and obtain all required parameters of the system model as listed in the input section of Table 2.

   $a.\ g_{1i}, g_{ji} \& D_j$

– Retrieve the SBS on/off status and UE power allocation based on environment (i.e., learned from FFNN).

   $a.\ p_{1i}, p_{ji} \& b_j$

Algorithm 1 . ANN-PA-BSSw

Initialize weights and biases
Initialize the hyper-parameters: learning rate, number of epochs, hidden neuron and hidden layers
Set minimum acceptable MSE, e
while $MSE > e$ **do**
  for $\forall$ *training data* **do**
    Define the inputs
    Compute the outputs
    Adjust the weights and biases using backpropagation
  end for
  Evaluate MSE
end while
Retrieve the final value of weights and biases
Apply for new inputs
Evaluate new outputs

4. RESULTS AND DISCUSSION

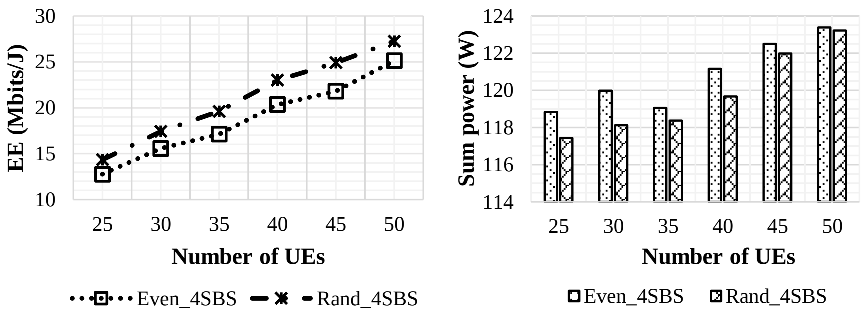
First, the training samples are collected using a combinatorial algorithm of conventional sub-gradient and a brute force method. Eventually, this combinatorial method is used as a benchmark to evaluate the bias of the designed feed-forward ANN. In a neural network, optimizing the hyper-parameters of the proposed ANN-based BS-Sw and PA scheme (i.e., the number of epochs, hidden layers, hidden neurons, and training data) is considered a crucial step to determine the performance of the network. Therefore, intensive simulation has been conducted to find the best set of hyper-parameter values for the proposed scheme.

The trained ANN model is evaluated by testing it with a new set of data simulated through MATLAB, and further, the outputs are used to determine the EE of the network. Further, the evaluation is repeated for a different set of hyper-parameters, as mentioned earlier, and the obtained EE is also compared with the benchmark model, which is modelled through the conventional method. In this work, different numbers of UEs are simulated for different cells within which UEs' locations are randomly deployed. Therefore, an analysis is done to prove the importance of random user distribution across the cells in EE evaluation of the system. The analysis is carried out with the total count of users in the system is either evenly (divided equally) or randomly (random distribution) distributed among the cells to experiment with the power consumption and EE of both situations. The simulation setting used in the training phase of this work is tabulated in Table 3.

Table 3. Configuration of the simulation parameters

Parameter	Configuration
System bandwidth	10 MHz
Macrocell radius	500 m
Small cells radius	40 m
Noise power density	-174 dBm/Hz
Channel fading model	Exponentially distributed Rayleigh Zero mean and unit variance
Path loss models	MBS: 128.1 + 37.6 log(d_{km}) dB SBS: 140.7 + 36.7 log(d_{km}) dB
Shadowing model	iid log-normal Zero mean and 10 dB standard deviation
Static power	MBS: 130 W SBS: 6.8 W
Power amplifier efficiency	MBS: 39% SBS: 7%

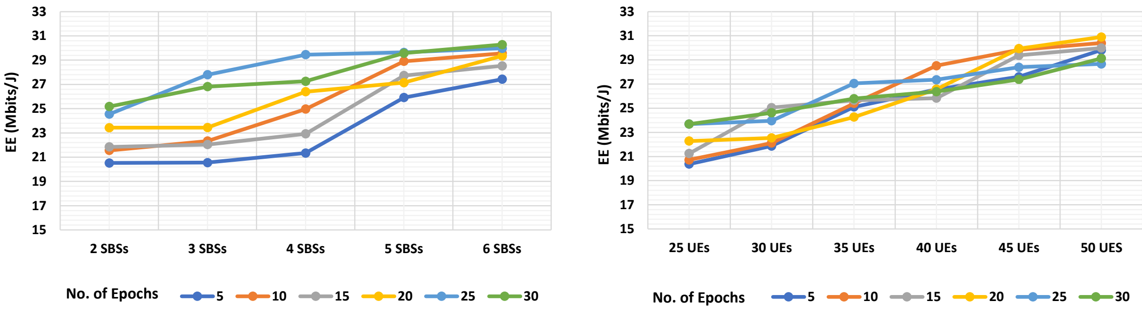
Figure 4 shows the impact of user distribution or traffic patterns in the cellular network on the EE of the system employed with four SBSs. It is worth noting that this analysis is carried out using the conventional method to prove the significance of user distribution. Practically, the count of users in each cell varies randomly and this condition surely contributes differently to the power consumption of the system compared to distribution of same number of users in every cell. As illustrated in Figure 4(a), the EE of both types of user distribution rises as the total count of UEs are increased, but the overall EE of random distribution is around 10% higher than that of even distribution. This can be explained in Figure 4(b), where the power consumed by the former is lesser than the latter because, underutilized SBSs (i.e., during low peak traffic conditions) are switched off to save more power. Whereas, if the traffic pattern is the same for all SBSs, this rule does not apply.



(a) (b)

Figure 4. Comparison of; (a) EE and (b) total power consumption for even and random user distribution in a CoMP-enabled HetNet with 4 small cells

Figure 5 illustrates the influence caused by the number of epochs on the EE of the system for different numbers of SBSs Figure 5(a) and users Figure 5(b). Generally, as the count of SBSs or users grows, the value of EE also increases as it is supposed to be as in the conventional method [18]. In addition, increasing the number of epochs contributes to a further increase of EE, but this is not true for all conditions, as can be observed in Figure 5(b). This phenomenon is caused by overfitting of the network where the trained network fails to generalize to the new set of data due to over-adaptation to the training data. By over tuning the parameters to learn the training data, the new results are unable to be derived accurately by the network using the new set of data as can be seen obviously in the graph trend of 30 epochs in Figure 5(b). This analysis demonstrates that the performance of the network does not necessarily get improved by increasing the number of epochs.



(a) (b)

Figure 5. EE of a neural network trained with a different number of epochs with one hidden layer for different number of; (a) SBSs and (b) UEs

Figure 6 depicts the effect of employing three different types of backpropagation algorithms to the neural network, namely, LM [22], BR [23], and scaled conjugate gradient (SCG) [24]. Along with it, the influence of the number of neurons and learning rate on the EE is also analyzed. It is the rate at which the network model learns or adapts to the mapping of input-output of the training data. As the theory suggests, the EE of the network increases as the learning rate becomes smaller, but it is worth noting that the computational time dramatically increases along with it. Moreover, the EE of the network trained at learning rate of 0.001 is just around 6% greater than that of the 0.01. Due to its fast convergence characteristics, LM is one of the most recommended back propagation algorithms. Other algorithms like BR induces higher computational time due to the distribution of the weights as explained in section 3.2 and SCG is more preferably used for a very high number of training data for faster convergence at the cost of performance decay. On the other hand, for LM and BR, the network trained using 20 neurons shows around 4% and 5% higher EE than 10 neurons, respectively. Whereas for SCG, the network trained using 10 neurons seems to perform almost equally with 20 neurons (the difference is around 1%).

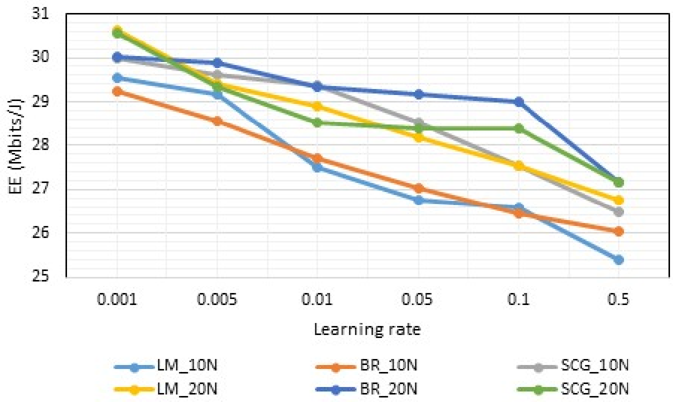


Figure 6. EE of a neural network trained using different types of backpropagation algorithm, LM, BR, and SCG with one hidden layer

Figure 7 illustrates the impact of the number of hidden layers on the trained ANN. At first glance, it can be interpreted that increasing the number of hidden layers contributes to improved performance. But, in fact, the EE of a network with five hidden layers is just nearly 7% higher than that of one hidden layer. On the other hand, it is proven that increasing the count of hidden neurons does not significantly improve the network's performance. It is noteworthy that increasing the count of hidden layers and/or neurons greatly increases the complexity and computational time of the neural network training at an insignificant improvement of EE. Moreover, theoretically, one hidden layer is capable of handling most of the complex functions, and it is also proven in this case.

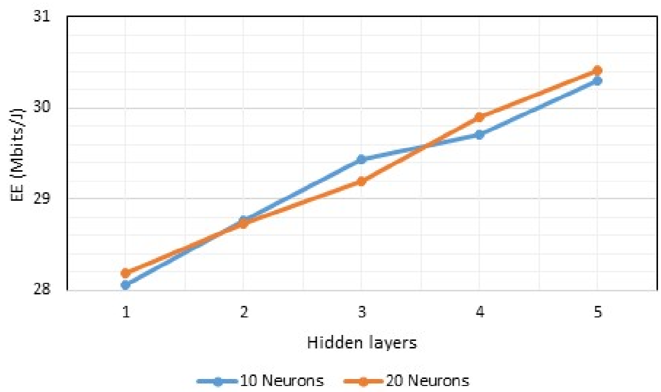


Figure 7. EE for neural network trained with different number of hidden layers using 10 and 20 hidden neurons

Figure 8 shows the bias of the overall feed-forward ANN trained with different types of backpropagation algorithms and the number of SBSs. Bias, in this case, is the measure of the difference between the expected value which was obtained from [18] (used as benchmark) and output from trained ANN applied to a new data set. Generally, the bias of networks trained using LM, BR and SCG algorithms achieves around 2%, 5%, and 3%, respectively, where the differences are trivially small. The network trained using LM algorithm achieves the highest bias overall (also as highlighted in [25]) and further LM's complexity and computational time is considerably low compared to other backpropagation algorithm. It is noteworthy that in this case, the values of network parameters which can reduce the complexity and computational time can be considered while designing the training network model.

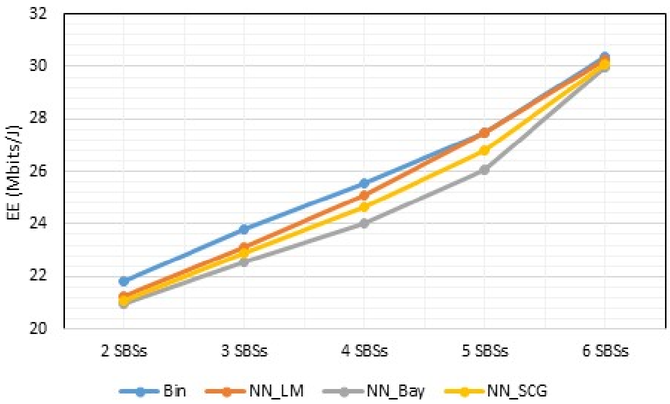


Figure 8. Bias of a neural network trained using different types of back propagation algorithm with one hidden layer

5. CONCLUSION

This work proposes an ANN-based joint PA and BS-Sw scheme for a hybrid-powered CoMP-enabled HetNet system. By using a feed-forward neural network to train the model to determine UE power allocation and BS on/off statuses which contributes to EE maximization of the HetNet system, the computational delay and complexity of the conventional methods (i.e., brute force search and dual decomposition) are greatly reduced with almost similar EE performance (around 2% difference) achieved by the trained model in contrast to the conventional methods. It is noteworthy that the network parameters for the neural network training has to be set in a way to decrease the complexity and computational time of the network in order to prove the significance of the technique. The effectiveness of the proposed joint PA and BS-Sw scheme using three distinct backpropagation algorithms i.e., LM, BR, and SCG, is also examined. Results reveal that the LM algorithm demonstrates superior computational efficiency and reduced processing time, albeit with a marginally higher bias compared to the BR and SCG counterparts. The insights derived from this work are highly valuable for both the research community and telecommunication industry. They empower designers and operators of telecommunication systems to effectively tackle the issues of excessive energy consumption, rising energy costs, and rising CO_2 emissions. Future research may delve into development of self-optimizing networks that capable of analyzing more complex environmental changes and performing self optimization to reduce human intervention.

ACKNOWLEDGEMENT

The authors wish to express their appreciation to the reviewers for their helpful suggestions which greatly improved the presentation of this paper. This research and the APC was funded by Multimedia University under grant number MMUI180002.02.

REFERENCES

[1] N.-N. Dao, M. Park, J. Kim, J. Paek, and S. Cho, "Resource-aware relay selection for inter-cell interference avoidance in 5g hetero- geneous network for internet of things systems," *Future Generation Computer Systems*, vol. 93, pp. 877–887, 2019, doi: 10.1016/j.future.2018.03.037.

[2] O. Alamu, A. Gbenga-Ilori, M. Adelabu, A. Imoize, and O. Ladipo, "Energy efficiency techniques in ultra-dense wireless hetero-

geneous networks: An overview and outlook," *Engineering Science and Technology, an International Journal*, vol. 23, no. 6, pp. 1308–1326, 2020, doi:10.1016/j.jestch.2020.05.001.

[3] V. I. Tatsis, M. Karavolos, D. N. Skoutas, N. Nomikos, D. Vouyioukas, and C. Skianis, "Energy-aware clustering of comp-dps transmission points," *Computer Communications*, vol. 135, pp. 28–39, 2019, doi: 10.1016/j.comcom.2018.11.009.

[4] F. R. Arvaje and B. S. Ghahfarokhi, "User re-association and resource/power re-allocation of cell switched-off user equipments in green cellular networks," *Physical Communication*, vol. 37, p. 100873, 2019, doi: 10.1016/j.phycom.2019.100873.

[5] D. G., "A survey on smart grid technologies and applications," *Renewable energy*, vol. 146, pp. 2589–2625, 2020, doi: 10.1016/j.renene.2019.08.092.

[6] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3072-3108, Fourthquarter 2019, doi: 10.1109/COMST.2019.2924243.

[7] G. Zhu, D. Liu, Y. Du, C. You, J. Zhang, and K. Huang, "Toward an Intelligent Edge: Wireless Communication Meets Machine Learning," in *IEEE Communications Magazine*, vol. 58, no. 1, pp. 19-25, Jan. 2020, doi: 10.1109/MCOM.001.1900103.

[8] Y. Liu, S. Bi, Z. Shi, and L. Hanzo, "When Machine Learning Meets Big Data: A Wireless Communication Perspective," *IEEE Vehicular Technology Magazine*, vol. 15, no. 1, pp. 63-72, Mar. 2020, doi: 10.1109/MVT.2019.2953857.

[9] V. M. Joy, J. John, and S. Krishnakumar, "Backpropagation neural network based adaptive load scheduling method in an isolated power system," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 5, pp. 3000–3007, 2023, doi: 10.11591/eei.v12i5.4511.

[10] M. Chen, U. Challita, W. Saad, C. Yin and M. Debbah, "Artificial Neural Networks-Based Machine Learning for Wireless Networks: A Tutorial," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3039-3071, 2019, doi: 10.1109/COMST.2019.2926625.

[11] A. Zappone, M. Di Renzo and M. Debbah, "Wireless Networks Design in the Era of Deep Learning: Model-Based, AI-Based, or Both?," *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 7331-7376, Oct. 2019, doi: 10.1109/TCOMM.2019.2924010.

[12] W. Lee, M. Kim and D. -H. Cho, "Deep Power Control: Transmit Power Control Scheme Based on Convolutional Neural Network," *IEEE Communications Letters*, vol. 22, no. 6, pp. 1276-1279, Jun. 2018, doi: 10.1109/LCOMM.2018.2825444.

[13] K. I. Ahmed, H. Tabassum and E. Hossain, "Deep Learning for Radio Resource Allocation in Multi-Cell Networks," *IEEE Network*, vol. 33, no. 6, pp. 188-195, Nov.-Dec. 2019, doi: 10.1109/MNET.2019.1900029.

[14] A. Zappone, M. Debbah and Z. Altman, "Online Energy-Efficient Power Control in Wireless Networks by Deep Neural Networks," *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Kalamata, Greece, 2018, pp. 1-5, doi: 10.1109/SPAWC.2018.8445857.

[15] F. Liang, C. Shen, W. Yu and F. Wu, "Towards Optimal Power Control via Ensembling Deep Neural Networks," in *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1760-1776, Mar. 2020, doi: 10.1109/TCOMM.2019.2957482.

[16] N. Piovesan, A. F. Gambin, M. Miozzo, M. Rossi, and P. Dini, "Energy sustainable paradigms and methods for future mobile networks: A survey," textslComputer Communications, vol. 119, pp. 101–117, 2018, doi: 10.1016/j.comcom.2018.01.005.

[17] Q. Han, B. Yang, N. Song, Y. Li, and P. Wei, "Green resource allocation and energy management in heterogeneous small cell networks powered by hybrid energy," *Computer Communications*, vol. 160, pp. 204–214, 2020, doi: 10.1016/j.comcom.2020.06.002.

[18] S. Euttamarajah, Y. H. Ng, and C. K. Tan, "Energy-efficient joint base station switching and power allocation for smart grid based hybrid-powered comp-enabled hetnet," *Future Internet*, vol. 13, no. 8, p. 213, 2021, doi: 10.3390/fi13080213.

[19] C. Lv *et al.*, "Levenberg–Marquardt Backpropagation Training of Multilayer Neural Networks for State Estimation of a Safety-Critical Cyber-Physical System," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3436-3446, Aug. 2018, doi: 10.1109/TII.2017.2777460.

[20] J. Borah and J. Bora, "Energy-efficient ici mitigation with dynamic and location-based power allocation in mobility-based 5g HetCN," *Wireless Personal Communications*, vol. 117, pp. 1441–1457, 2021, doi: 10.1007/s11277-020-07930-x.

[21] O. Baghirli, "Comparison of lavenberg-marquardt, scaled conjugate gradient and bayesian regularization backpropagation algorithms for multistep ahead wind speed forecasting using multilayer perceptron feedforward neural network," Ph.D. dissertation, Dept. of Earth Sciences, Uppsala University, Uppsala, Sweden, 2015.

[22] R. Ullah *et al.*, "A machine learning approach for 5g SINR prediction," *Electronics*, vol. 9, no. 10, p. 1660, 2020, doi: 10.3390/electronics9101660.

[23] Q.-V. Pham *et al.*, "Deep learning for intelligent demand response and smart grids: A comprehensive survey," *arXiv*, 2021, doi: 10.48550/arXiv.2101.08013.

[24] H. T. Thien, P.-V. Tuan, and I. Koo, "Deep learning-based approach to fast power allocation in siso swipt systems with a power-splitting scheme," *Applied Science*, vol. 10, no. 10, p. 3634, 2020, doi: 10.3390/app10103634.

[25] A. A. Ajao *et al.*, "Cellular network bandwidth improvement using subscribers' classification and wi-fi offloading," *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 1, 2017, vol. 11, no. 2, pp. 917–925, 2022, doi: 10.11591/eei.v11i2.2575.

BIOGRAPHIES OF AUTHORS

Shornalatha Euttamarajah received the B.Eng. degree and the M.Eng. degree in telecommunications engineering from Multimedia University, Malaysia, in 2011 and 2016, respectively. From 2011 to 2017, she was with the Engineering Department of Mettube Sdn. Bhd., Malaysia as an Electronic Engineer. She is currently pursuing the Ph.D. degree from Multimedia University, Malaysia. Her research interest includes green energy communications, HetNets, energy efficient and cooperative techniques, and optimization theory. She can be contacted at email: shornalatha16@gmail.com.

Yin Hoe Ng received the B.Eng. (Hons.) degree in electronics engineering, Master of Engineering Science, and Ph.D. from Multimedia University, Malaysia, in 2004, 2008, and 2013, respectively. Since 2004, he has been with the Faculty of Engineering, Malaysia. His current research interests include advanced signal processing techniques for digital communication systems, machine learning, and indoor positioning. He can be contacted at email: yhng@mmu.edu.my.

Chee Keong Tan holds the degrees of B.Eng. (Electronics majoring in Telecommunication), M.Eng.Sc (Information and Communications Technology) and Ph.D. (Information and Communications Technology) from Multimedia University, Malaysia. Currently, he serves as a senior lecturer in School of Information Technology, Monash University Malaysia. As the lecturer in Monash University, he has involved in lecture for many undergraduate and postgraduate courses such as computer systems, computer networks, computer security, and research methods. His current research interests include machine learning, AI, unmanned aerial vehicle (UAV), indoor positioning, IoT, and beyond 5G communications. He is a computer networking expert and consultant with over 10 years of experience in networking research, lectures and practice. He is also an experienced speaker and teacher, having taught many highly-rated university level courses, industry short courses, tutorials at conferences, to a wide range of students from industry, academia and government. Specializing in wireless system design, he has carried out projects for networking companies and cellular service providers which led to the development of two patents on wireless algorithm and protocol design. Furthermore, he is the main contributing authors to more than 15 international journal papers and conference papers, particularly on the radio resource management for various network systems. He can be contacted at email: tan.cheekeong@monash.edu.