A Neural Network Based Approach for Call Admission Control in Heterogeneous Networks

Bushra Naeem¹,²*, Razali Ngah¹, Siti Zaiton Mohd Hashim³, Wajahat Maqbool¹,², Muhaddisa Barat Ali²,⁴

¹ Faculty of Electrical Engineering (FKE), University Teknologi Malaysia (UTM), Skudai, Malaysia
² Faculty of Information and Communication Technology, BUITEMS, Quetta, Pakistan
³ Faculty of Computing Universiti Teknologi Malaysia (UTM), Skudai, Malaysia
⁴ Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey

bushra.naeem@ymail.com

Abstract: The next generation wireless networks will be based on infrastructure with the support of heterogeneous networks. In such a scenario, the users will be mobile between different networks; therefore the number of handovers that a user has to make will become greater. Thus, at a given instant, there will be great chance that a certain cell does not have capacity to sustain the need of users. This may result in great loss of calls and lead to poor quality of service. Moreover, in the future generation of wireless networks, end users will be able to connect any suitable network amongst available set of heterogeneous networks. This ability of an end user being connected to the network of their choice may also affect network load of various base stations. This necessitates for a suitable call admission control scheme for the implementation of heterogeneous networks in the future. Since the behavior of users arriving at any cell in heterogeneous network is unpredictable, we utilize neural network to model our heterogeneous network to admit network load, therefore the learned neural network is able to estimate when call should be admitted in a new situation. Results obtained indicate that neural network approach solves the problem of call admission control unforeseen real-time scenario. The neural network shows reduced error for the increased values of learning rate and momentum constant.

Keywords: Load balancing; Heterogeneous Network (HetNet); picocell; femtocell; macrocell

1. Introduction

The worldwide cellular communication has developed in previous years connecting more than the half of the world's population. Owing to the demanding user needs, the cellular network is moving towards a paradigm where the number of Base Stations (BSs) increases rapidly every year [1]. Furthermore, the picocells and femtocells are being added to the existing wireless communication architecture [2-4]. According to[5], there will be 50 million BSs by 2015. It has been also predicted by [6] that in next 10-15 years , there will essentially be greater number of BSs than number of cell phone subscribers. According to 1 this will result in a cloud-like "data shower" where mobile station will be capable of connecting to multiple BSs at any given instant. A heterogeneous network scenario can be visualized as a cluster consisting of heterogeneous cells containing macro, pico and femto cells. This architecture calls for a need of appropriate load balancing in a HetNet scenario. Load balancing consists of accepting or rejecting a new user request and forcing users connected to a heavily loaded BS to handover to a lightly loaded one [8]. According to current practice and research, the network traffic has been associated with Access Point (AP) of highest Received Signal Strength (RSS). This practice of using RSS for traffic steering results in unbalanced load and unfair distribution of resources in the network .

Various load balancing schemes have been presented in literature. H. Son and S. Lee[7] present a load balancing mechanism named as “soft” load balancing. In this mechanism, the Internet Protocol (IP) packets in the downlink are delivered via two heterogeneous access networks, i.e. they are divided into subflows accordingly to their air interfaces. They obtain an optimal Load Balancing Ratio (LBR) to determine the volume of traffic delivered to each network in an overlaid multi-cell environment. However this technique is only used for OFDMA and CDMA systems, and results evaluated are in terms of outage probability, instead of showing direct effect of their technique on network traffic. The work also does not cover Call Admission Control (CAC). Vuong et al[8] put forward a novel approach for load balancing in heterogeneous wireless packet networks. They introduced an approach for computing network load metric based. The algorithm consists of two main processed: Admission control and Handover Enforcement. Their load metric hides the heterogeneity of radio resources among different

http://www.lifesciencesite.com

238

lifesciencej@gmail.com
integrated access technologies. The load metric is defined as the ratio between required resources and total resources. Authors in [9] provide a rate distribution of bandwidth into sub-flows for voice but did not mention CAC mechanism assuming that it is already present. In [10], the authors provide a Radio Access Technology (RAT) section scheme based on the properties of incoming call but again do not explain how load distribution will take place. In a recent research, the gray relational grade between an ideal network and network at some instant is calculated and optimal network is chosen [11]. The scheme presented in [12] also monitors load at each gateway and initiates handover in case of heavily loaded gateway status. In [13], the authors only estimate network traffic using active and hold states but do not give a mechanism how load between various HetNet cell be distributed. They perform handover on the basis of estimated load. In recent times, numerous load balancing schemes have been presented, [7-13].

In this work, use three layer neural network, i.e input layer, hidden layers and output layer. We propose neural network output based algorithm to make decision of call admission in a heterogeneous. We consider a HetNet cluster having macrocell overlaid on femtocells and picocells. The results show that the heterogeneous cells learned the proposed scheme and learning error was recorded very less for greater values of learning rate and momentum constants.

2. Mathematical Modeling of Users Heterogeneous Cluster

Let $i$ represent any network in the area under consideration, where $i$ is an element from the set of heterogeneous wireless networks, i.e. $i \in \{1,2,...,N\}$. We consider a footprint of globe having $i$ heterogeneous cells and we call this area as a cluster, denoted by $C_i$. Figure 1 shows a single macrocell overlaid on femtocells and picocells. The mobile terminals in the range of both picocell(or femtocell) may be connected to the macro or picocell (or femtocell) respectively.

The traffic load on each heterogeneous cell can be represented by the occupied user slots at a given instant of time. We consider a three tier case, having exactly each femto, pico and macro cell in a HetNet cluster.

We define the current load in a network $i$ as $LC_i$ where $i=f,p,m$ for femtocell, picocell and macrocell respectively. We further define the total capacity of network $i$ by $T_i$, the total load utilization in network $i$ at a given instant is then denoted as $W_i$ and calculated as:

$$W_i = \frac{LC_i}{T_i}$$

(1)

The total load utilization ($L_u$) inside heterogeneous network cluster at any instant can be obtained by following algorithm:

1. Define $S=\sum W_i$ for all $i$ inside $C_T$
2. Calculate $L_u = S/N$

Let $U_j$ represent number of users in a network at any instant, where $j \in \{0,1,2,...,M\}$. We model the probability of users in each heterogeneous cell $i$ as a discrete uniform random variable $X_j$. Since the probability of any user occupying a user slot in $C_i$ is equally likely, we obtain the following Probability Mass Function(PMF) for $U_j$ in any $C_i$.

$$P_X(j) = \begin{cases} \frac{1}{M+v+1} & \text{if } j = v, v+1, \ldots, M \\ 0 & \text{otherwise} \end{cases}$$

(2)

Where $v$ is the minimum value $j$ can take.

The event that a cell $i$ is overloaded occurs when $W_i = 1$. Similarly the event of the whole heterogeneous network cluster is overloaded occurs when $L_u=1$.

The event that $W_i = 1$ implies the following expression

$$\sum_{j=1}^{N} P_X(j) = 1$$

(3)

Similarly, the event that $L_u=1$ implies

$$\frac{1}{N} \sum_{l=1}^{N} \sum_{j=0}^{M} P_X(j) = 1$$

(4)

It follows from above equations that the network load must be distributed in a way that the following
conditions from (3) and (4) be maintained in the heterogeneous network.

\[
\sum_{x=1}^{N} P_x (j) < 1
\]

\[
\frac{1}{N} \sum_{i=1}^{N} \sum_{j=0}^{M} P_x (j) < 1
\]

3. Neural Network Modeling

Neural networks belong to the class of data-driven approaches where the analysis depends on available data, with little reasoning about possible interactions. Neural networks have been successfully applied for modelling non-linear phenomena including load balancing in multi-computer systems [14].

However, to the best of my knowledge, it has not yet been applied for controlling the call admission in a heterogeneous network environment. In this work a three-layer neural network approach is used, where three layers are input layer, hidden layers, and output layer. Figure 2 shows the adapted neural network system design.

Figure 2. Three layer neural network

This work uses the error back-propagation algorithm of Rumelhart et al. [14], based on gradient-descent, to train the networks, with the goal of minimizing the error, \( e \) between the desired target values and network outputs, averaged over all the training inputs. The training phase can be described as follows.

In each step in the training phase, a matrix, \( x \) of inputs is input to the network. The network is asked to predict the output value. The error between the target value and desired value is then measured and propagated backwards along the connections. The weights, \( w \) of links between units are continuously changed, to minimize the error nodes and links. The weight update is done using momentum constant \( \alpha \) which is multiplied to the change in current and previous weight. A single 'epoch' i.e., cycle of training inputs comprises of applying all input patterns once and modifying the weights after each step. We use sigmoid activation function as defined subsequently [15]. The parameters of the back-propagation algorithm are the 'learning rate' and 'momentum', which approximately describe the relative importance given to the current and past error values in modifying connection strengths. Here, \( m \) is the time index, \( H_{lk} \) is the weight from unit \( k \) to unit \( l \), \( q \) is an index over the cases (input samples), \( \delta_{ql} \) and \( O_{ql} \) is the propagated error signal seen at unit \( j \) in case \( q \), and \( O_{ql} \) is the output of the corresponding unit. For the sigmoid activation function, where

\[
O_{ql} = \frac{1}{1 + \exp(-\sum_k H_{lk}O_k - \theta_l)}
\]

4. Experimental Design

The future HetNets will be prone to the unbalanced load due to the fact that cell capacity in each different base station varies from other to a great extent. For example, a picocell may get over loaded with number of users as less at 10, while on the other hand, a macro cell may still not be overloaded with number of users equal to 50. Thus a dynamic load balancing technique is required to reflect this heterogeneity while distributing the network traffic. Thus objective of this load balancing technique is to make sure that in the HetNet, the load in heterogeneous cells is not unbalanced. We define a network load threshold for all networks as \( \rho \). The value of \( \rho \) is found by subtracting the small portion of capacity called guard band, \( \psi \) so that this guard can accommodate the handover calls in different cells. This value in determined by different operators according to the capacity of each cell. Thus it follows that

\[
\rho = 1 - \psi
\]

We design three neural networks, based on similar architecture as shown in fig.2, each for the load of femtocell, picocell and macrocell respectively using back propagation algorithm, and train them to accommodate calls in each cell until the load utilization, \( W_i \), as calculated by equation 1 in each network reaches to the network load threshold \( \rho = 0.8 \). For users that send request when \( W_i = 0.8 \), they will be admitted to other available cell \( i \) that has \( W_i \leq 0.8 \). The neural networks are trained for each of heterogeneous cell so that only calls below threshold level will be admitted to the cell under consideration. If the threshold limit is exceeded, this algorithm is explained in section A below. The neural network output is classified based on two classes, 0 or 1. If \( W_i \leq 0.8 \), call will not be accepted in current network. This is denoted by class 0. Conversely, if \( W_i > 0.8 \), the call will be admitted in current network. The output from each neural network is used by Algorithm A to perform call admission to the respective cell.
A. Proposed Neural Network Algorithm for Call Admission Control

Case 1: Call request to femtocell
If $W_f \leq \rho$, admit call to the femtocell
else if $W_f > \rho$ then compare $W_m \leq \rho$
else reject the call

Case 2: Call request to picocell
If $W_p \leq \rho$, admit call to the picocell
else if $W_p > \rho$ then compare $W_m \leq \rho$
else reject the call

Case 3: Call request to macrocell
Check availability of femtocell and picocell
If femtocell is in range AND $W_f \leq \rho$, admit call to the femtocell
else if picocell is in range AND $W_p \leq \rho$, admit call to the picocell
else compare $W_m \leq \rho$
else reject the call

5. Results and Discussion
The proposed neural network based approach is simulated using MATLAB. We simulate the network load $W_i$ for three different learning rate, $c$ as shown in the figure 3 below: The results show that the neural network error reduces as we increase the learning rate of network. Moreover, as the number of epochs increase, the error is reduced exponentially. The simulation parameters used are given in table 1 below.

![Figure 3. Change in error value for different values of learning rate](image)

Table 1. Simulation parameters for designing call admission neural network for different learning rates

<table>
<thead>
<tr>
<th>Momentum constant ($\alpha$)</th>
<th>Epoch min</th>
<th>Epoch max</th>
<th>Activation Function</th>
<th>Learning rate ($c$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>0</td>
<td>40000</td>
<td>Sigmoid</td>
<td>0.1, 0.4, 0.7</td>
</tr>
</tbody>
</table>

We again perform simulation to see the effect of varying momentum constants, as the impact of weight on neurons of neural network is also important. The simulation parameters used are shown in table 2. The results show that the impact of increasing the momentum constant of neuron weights is somewhat similar to that of learning rate. However, the reduction in error is not as sharp as that experienced for the case of learning rate. This is because, the weight of neurons have to be updated many times in order for it to learn required output.

![Figure 4. Change in error value for different values of momentum constant](image)

Table 2. Simulation parameters for designing call admission neural network for different momentum constants

<table>
<thead>
<tr>
<th>Momentum constant ($\alpha$)</th>
<th>Epoch min</th>
<th>Epoch max</th>
<th>Activation Function</th>
<th>Learning rate ($c$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1, 0.4, 0.8</td>
<td>0</td>
<td>40000</td>
<td>Sigmoid</td>
<td>0.1</td>
</tr>
</tbody>
</table>

6. Conclusions
In this paper, we have proposed a neural network based call admission control scheme for heterogeneous networks using neural network. To the best of our knowledge, no work has been done to achieve call admission control in heterogeneous network using trained heterogeneous cell approach. The neural networks are used in conjunction with algorithm A to allow learned call acceptance in unseen circumstances. The neural networks are evaluated for three heterogeneous cells, macro, pico and femto cells respectively, for varying learning rates and momentum constants. The results show that increasing learning rates or momentum constants reduce the error in neural network.

Acknowledgements:
The authors thank the Ministry of Science and Technology (MOSTI), Malaysia for providing financial support for this work through the eScienceFund Grant (R.J130000.7923.4S066). The grant is managed by Research Management Centre (RMC), Universiti Teknologi Malaysia (UTM).

Corresponding Author:
Bushra Naeem
Wireless Communication Centre,  
Faculty of Electrical Engineering (FKE)  
University Teknologi Malaysia (UTM)  
Skudai, Malaysia.  
E-mail: bushra.naeem@ymail.com

References