Pakistan Journal of Science, 71 (4 Suppl.): 2019, Page 258-265
ISSN: 2411-0930

MACHINE LEARNING BASED CAPACITY ENHANCEMENT OF FEMTOCELLS FOR 5G HETEROGENEOUS NETWORKS

M.U. Iqbal*, E.A. Ansari and S. Akhtar

Department of Electrical and Computer Engineering, COMSATS University Islamabad (CUI), Lahore Campus, Pakistan
*Corresponding author's E-mail: usmaniqbal@cuilahore.edu.pk

ABSTRACT: Small cells based ultradense heterogeneous networks (HetNets) are being considered as the one the promising solution for increased coverage and capacity in the 5G cellular networks. However, in the multi-tiered architecture, co-tier and cross tier interference are a performance-limiting factor. The interference can be effectively handled through efficient resource allocation techniques in either a cooperative or distributive manner. However, the complexity of such resource allocation schemes linearly increases with the density of the HetNets due to unplanned deployment and dynamic behavior of small cells. The HetNets can be implemented only through an adaptive and self-organizing algorithm that can adapt to the dynamic conditions. In this research paper, a machine learning (ML) based adaptive resource allocation scheme is proposed for the femtocell based dense HetNets. The Q-Learning based scheme consider each femtocell base station (FBS) as the agent of the network and model the HetNets as multi-agent network to allocate optimal power to the FBS to maximize the capacity of the femtocell user equipment (FUEs) an macrocell user equipment (MUEs) while considering the quality of service (QoS) requirements. The proposed cooperative Q-Learning scheme increases the sum capacity of the FUEs by seven-folds and always ensures the minimum QoS requirements as compared to the prior work. Furthermore, the proposed solution also increased the number of supported femtocells by two-fold in comparison to the state of the art solution.

Keywords: Machine Learning (ML), Heterogeneous networks (HetNets), Q-Learning, Femtocells

INTRODUCTION

The exponential growth in mobile users (UEs) in the last few years cannot be handled with traditional cellular networks. It is expected that by 2020, 50 billion users will be connected to the cellular network. The next generation of the cellular network i.e. 5th Generation (5G) is expected to meet the capacity and other quality of service (QoS) requirements like higher data rate, throughput and zero latency (Agiwal *et al.,* 2016; Akpakwu *et al.,* 2018; Parvez *et al.,* 2018). However, meeting the fundamental requirements of the 5G is not an easy task. Many solutions have been proposed in the literature to enhance the capacity and other QoS requirements which include massive multiple inputs and multiple outputs (MIMO), millimeter-wave (mmW) communication, non-orthogonal multiple access (NOMA) and heterogeneous networks (HetNets) (Gupta and Jha, 2015). An overview of the 5G cellular network is presented in Fig-1. The HetNets are considered as the viable solution to meet the coverage and capacity requirements in 5G cellular networks, however, the deployment of the small cells in a multi-tiered architecture to form HetNets results in co-tier and cross tier inference. To effectively deploy the HetNets to increase capacity and coverage, the interference caused by the unplanned and dense deployment of the small cells has to be resolved (Gupta and Jha, 2015; Zhang *et al.*, 2015). In this paper, a cooperative power control scheme is proposed for the interference mitigation and enhancing and sum capacity of the femtocells. The proposed algorithm is based on the reinforcement learning (RL) scheme.

HetNets are one of the solutions proposed in the literature to increase the coverage and capacity of the existing and future cellular networks. The idea is to deploy low-powered and small-ranged access points or base stations connected to the backhaul through optic fiber. These small ranged cells comprised of the low powered access point or base station are called small cells. Based on the transmission power and deployment scenario, small cells can be classified as the picocells, microcells, and femtocells. The detailed classification of small cells is discussed by (Gupta and Jha, 2015). The coverage radius of femtocells is approximately 200m which is very less as compared to macrocell which have coverage radius of 10-40Km (Gupta and Jha, 2015; Tseng *et al.*, 2015). The comparison of different parameters of the femtocells and macrocell is presented in Table-1. Femtocells are connected to service providers through optic fiber. The macrocell and femtocells operate in the same frequency band and therefore macrocell users

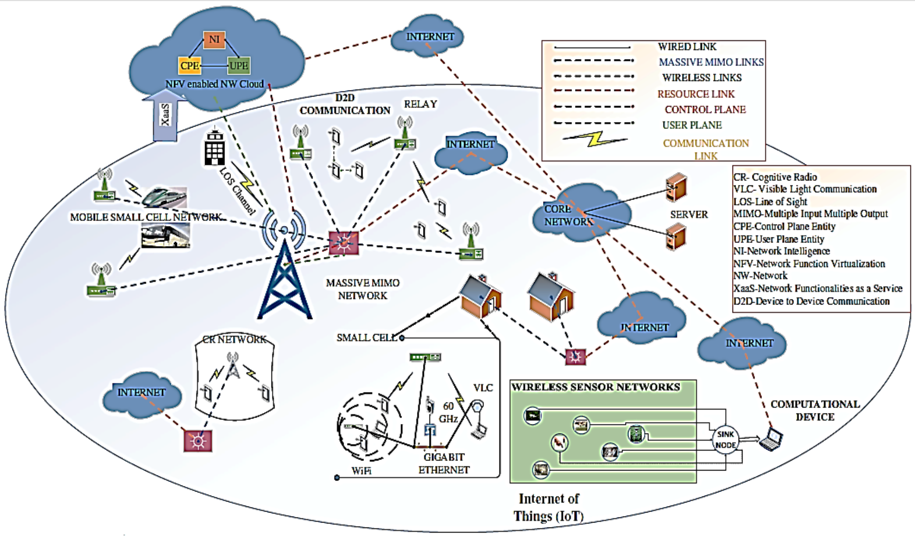

Fig-1: Overview of 5G Cellular Networks (Gupta and Jha, 2015)

Table-1: Comparison of Femtocell and Macro Cell in 5G Cellular HetNets (Gupta & Jha, 2015).

	Femtocell	Macrocell
Bandwidth (MHz)	20	60-75
Output power (W)	0.250	40-100
Cell Radius (m)	200	10-40km
No. of Users	32-100	1000+
Indoor/ Outdoor	Indoor/ Outdoor	Indoor/ Outdoor
Backhaul	Microwave/ Fiber	Microwave/ Fiber

(MUEs) become primary users and femtocell users (FUEs) are considered as the secondary users which sense the spectrum holes and transmit intelligently (Sanchez *et al.*, 2016). However, in the densely populated areas like an urban environment, the QoS is effected by the co-tier and cross tier interference. Interference can be mitigated in several ways however an adaptive power control is considered as an effective approach (Kurda *et al.*, 2014; Yang *et al.*, 2018). In the last few years, many interference mitigation and QoS improvement schemes have been proposed in the literature, however, most of such schemes do not consider the uncertainty for the deployment of femtocells in the macrocell and a self-organizing the capability of the femtocells to adapt to the dynamic conditions. Machine learning (ML), which has recently got the attention of the researcher and scientists to provide cognition to the different types of systems, has also found many applications in cellular networks. RL, which is a type of ML is developed to optimize the unknown system by interacting with it. RL is a perfect solution in situations where systems are unknown and conditions are dynamic. Unlike the many other ML techniques, RL does not require training prior to the application (Whitehead, 1991). Due to this property, RL is being applied in the communication networks especially in the areas like resource allocation problems, dynamic spectrum access and other distributed nature problems (Bin *et al.*, 2014; Feng *et al.*, 2009; Galindo-Serrano and Giupponi, 2010; Tefft and Kirsch, 2013). In this paper, the RL based scheme is proposed for power allocation to the femtocells to increase the sum capacity of the femtocells.

Related Work: To solve the optimization problem of resource allocation in HetNets, a proper reward function in Q-learning is essential. The reward function can be selected or proposed based on the underlying optimization problem. To optimize the power allocation problem in HetNets, several reward functions have been proposed in the literature (Amiri *et al.*, 2018; Bin *et al.*, 2014; Saad *et al.*, 2012; Tefft and Kirsch, 2013). The power allocation problem can be solved either cooperatively (Bin *et al.*, 2014; Galindo-Serrano and Giupponi, 2010) or in the distributive manner (Saad *et al.*, 2012; Tefft and Kirsch, 2013). Most of these schemes improve either QoS or Capacity of MUEs or FUEs and neglects the other parameters. The Round Robin approach has been used by (Bin et al., 2014), in order to improve the throughput of cell edge users while maintaining the fairness between FUEs and MUEs. (Saad *et al.*, 2012), presented the work based on the cooperative Q-learning technique to maximize the sum capacity and maintain the threshold for the FUEs and MUEs respectively. However, QoS for the FUEs are ignored by both Bin et al., 2014 and Saad et al., 2012. Furthermore, the reward functions presented in the literature are not designed to handle the ultradense and unplanned networks. (Galindo-Serrano and Giupponi, 2010; Saad *et al.*, 2012; Bin *et al.*, 2014). However, (Tefft and Kirsch, 2013), proposed a Q-Learning based solution to fairly allocate power to FBS by considering the proximity of FUEs and MUEs. However, this proposed solution could not provide the required QoS to the FUEs. Furthermore, the details of the cooperation among the FUEs are also kept hidden. In this paper, we have proposed a Q-Learning based solution to allocate optimal power to the FBS to increase the sum capacity of the FUEs while ensuring the minimum QoS requirements. The proposed solution successfully handled the limitations of the prior works.

MATERIALS AND METHODS

A. System Model: In this research, a single cell of a 5G cellular HetNets comprised of a single macro base station (MBS) and M femtocell base stations (FBSs) is considered. Initially, to keep the system model simple, only one femtocell user equipment (FUE) is assumed for each FBS. The objective of the research is to allocate optimal power in the downlink of a dense HetNets where co-tier and cross tier interference are the performance-limiting factor. Although only one MUE and FUE for each FBS is assumed in this paper, the proposed solution can be implemented for the higher number of MUEs and FUEs. The signal received by the MUE receiver include cross tier interference and some thermal noise. Therefore, the signal-to-interference-noise-ratio (SINR) at the MUE and FUE is calculated as follows:

$$SINR_{MUE} = \frac{P_{MBS}\, h_{MM}}{\sum_{i=1}^{M} P_{FBS_i}\, h_{FM} + \sigma^2} \qquad \#(1)$$

$where$
$P_{MBS} = Transmit\ Power\ of\ the\ MBS\ (Watts)$
$P_{FBS_i} = Transmit\ Power\ of\ the\ i^{th}\ FBS\ (Watts)$
$h_{MM} = Channel\ gain\ from\ MBS\ to\ MUE$
$h_{FM} = Channel\ gain\ from\ FBS\ to\ the\ MUE$
$\sigma^2 = Variance\ of\ AWGN$

$$\boldsymbol{SINR_{FUE_i} = \frac{P_{FBS_i}\, h_{FF}}{P_{MBS}\, h_{MF} + \sum_{j=1, j\neq i}^{M} P_{FBS_j}\, h_{FF_{ij}} + \sigma^2}} \#(2)$$

$where$
$P_{FBS_i} = Transmit\ Power\ of\ the\ i^{th}\ FBS\ (Watts)$
$h_{FF} = Channel\ gain\ from\ i^{th}\ FBS\ to\ i^{th}\ FUE$
$P_{MBS} = Transmit\ Power\ of\ the\ MBS\ (Watts)$
$h_{MF} = Channel\ gain\ from\ MBS\ to\ i^{th}\ FUE$
$h_{FF_{ij}} = Channel\ gain\ from\ i^{th}\ FUE\ to\ j^{th}\ FUE$

Following the assumptions made by (Feng *et al.*, 2009; Tefft and Kirsch, 2013), we also assumed that channel parameters are known by the FBS. As the FBSs are connected to the backhaul through optic fiber, therefore channel parameters can be easily shared.

Finally, the normalized capacity of MUE and FUEs is calculated as follows:

$$C_{MUE} = log_2(1 + SINR_{MUE}) \#(3)$$

$$\boldsymbol{C_{FUE_i} = log_2\left(1 + SINR_{FUE_i}\right)} \#(4)$$

$Where$
$i = 1,2,3, \dots \dots, M$

In the subsequent subsections, optimization problem and the proposed solution is presented in the detail.

B. Optimization Problem

The objective of the optimization problem to allocate optimal power to the FBSs which maximize the sum capacity of the FUEs, while ensuring the minimum defined QoS requirements for MUE and FUEs. The optimization problem is defined as follows:

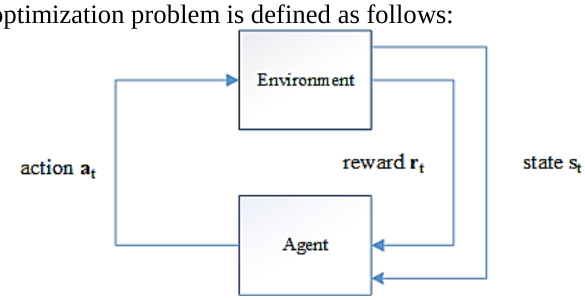

Fig-2: Reinforcement Learning (RL)

$$\max_{\tilde{p}} \quad \sum_{k=1}^{M} C_{FUE_k} \qquad \#(5)$$

$$subject\ to \quad P_i \le P_{max}, i = 1, \dots, M \#(5a)$$

$$C_{FUE_i} \ge \xi_{FUE_i}, i = 1, \dots, M \qquad \#(5b)$$

The fundamental objective in (5) is to maximize the sum capacity of the FUEs while maintaining the minimum QoS for the MUE and FUEs as described in (5b) and (5c) where ξ_{FUE_i} and ξ_{MUE} are QoS threshold for FUEs and MUE respectively. Another constraint, (5a) is to ensure the allocated power is below the maximum power. It is evident from the (2), (4) and (5) that optimization is a non-convex problem around densely populated femtocells. In fact, the interference terms in the denominator of (2) enforce that optimization problem stated in (5) as non-convex.

C. RL based Proposed Solution

The RL is a suitable ML technique for the environments or scenarios where single or multiple agents act to interact with environmental variables based on defined policy. Interaction of the agent results in the form of the feedback from the environment which is called reward and agents are updated accordingly. In the HetNets, FBS can be the agents that may interact with the environment and update their states accordingly. The overall process of the RL system is shown in Fig-2.

The RL can be implemented through a Q-Learning model which employee dynamic programming (Amiri *et al.*, 2018). The Q-Learning function can be taken as an approximator which depends upon the state, s_t, and action, a_t, at any time instant t. The Q function can be approximated through the following equation

$$Q(s_t, a_t) = \max_a \{E[R_t + \gamma Q(s_{t+1}, a_t)]\} \#(6)$$

$where$

$R_t = Reward\ Function\ at\ time\ t$
$\gamma = Discount\ Factor\ \ (0 \leq \gamma \leq 1)$

According to the (Amiri *et al.*, 2018), the (6) converges to a unique concave solution as $\rightarrow \infty$. The Algorithm 1, presents the simple Q-Learning scheme which is employed in this research paper. The parameter α used in the Q-Learning algorithm represents the rate of learning.

The four fundamental parts of the Q-Learning algorithms are briefly described below:

Algorithm 1: Q-Learning Algorithm

```
Define states S_t and actions A_t
Initialize Q-Table arbitrarily, i.e. Q(s_t, a_t)
for Iterations ≤ N_iterations do
    Initialize a_t
    for Step ≤ N_step do
        select a_t from A_t
        apply a_t
        observe R_t
        new state s_{t+1}
        update Q-Table Q(s_t, a_t) = Q(s_t, a_t) + α{max_a (R_t +
        γQ(s_{t+1}, a_t) ) − Q(s_t, a_t)}
        s_t ← s_{t+1}
    end
end
```

1) Actions: Actions are a step taken by the agents to maximize the objective function. In our case, the transmitting power for each FBS is considered as the actions. The transmit power of the FBS can be selected in an equally likely manner from the equally spaced set of powers

$$\boldsymbol{A} = [a_{\boldsymbol{1}}, a_2, \dots \dots, a_{N_{Power}}]\#(7)$$

between P_{min} and P_{max}. Therefore, the $Step_A$ can be defined as

$$Step_A = \frac{P_{max} - P_{min}}{N_{Power}}$$

2) States: In RL, state, s_t, describes the current situation of the agent. In our case, the state of the agent, i.e. FBS, is defined on the basis of the location of the FBS with respect to the nearby MUE and MBS. Therefore, to define the state of the FBS, the distance from the MBS, D_{MBS}, and distance from the MUE, D_{MUE}, is defined as follows on the basis of the distance rings N_{MBS} and N_{MUE} respectively:

$$D_{MBS} = [0, 1,2, \dots \dots, N_{\mathrm{MBS}}]\#(8)$$
$$D_{MUE} = [0,1,2, \dots \dots, N_{\mathrm{MUE}}]\#(9)$$

Based on the above-defined D_{MBS} and D_{MUE}, state, s_t, of i^{th} FBS at any time, t, is defined as follows:

$$s_t^i = [D_{MBS}, D_{MUE}]\#(10)$$

3) Q-Table: Based on the actions and states, a table is constructed that include all possible options of the actions and states where the actions are in columns and states in rows. This table is called Q-Table. The Q-Table may remain fix or vary according to change in the state of the FBS.

4) Reward: The reward function is a vital part of the Q-Learning technique. The accuracy of the algorithm depends upon the optimized reward function. There is not a qualitative method to drive the reward function. A reward function that maximizes the objective function can be proposed. In our case, the objective of the research is to maximize the sum capacity of the FUEs while maintaining the minimum defined QoS. The reward function is derived from the (5) and is given below:

$$R_i^t = \mathcal{T}\ \underbrace{C_{MUE_t}^n C_{FUE_t}}_{\mathcal{X}} - \mathcal{T}^{-1}\ \underbrace{(\beta_{MUE} + \beta_{FUE})}_{\mathcal{Y}}\ \#(11)$$

where $\beta_{MUE} = \left(C_{MUE_t} - \ \xi_{MUE}\right)^2$ and $\beta_{FUE} = \left(C_{FUE_t} - \boldsymbol{\xi_{FUE}}\right)^2$ are just the constants used to represent (11) in a concise form.

The reward function $R_i^t$is composed of two major parts $\mathcal{X}$ and $\mathcal{Y}$ as shown in (11). In the first part, $\mathcal{X}$ indicates that the reward is maximum when FUE and/or MUE capacities are higher. However, to give more weight to the MUE as the primary user, the capacity of MUE is powered by n; where $n \geq\ 2$. The second part, $\mathcal{Y}$, of the reward function ensures the QoS requirements where $\boldsymbol{\xi}_{MUE}$ and $\boldsymbol{\xi_{FUE}}$ defines minimum QoS requirements for the MUE and FUE respectively. The term $\mathcal{Y}$ is the deviation of MUE and FUE from the threshold QoS requirements, therefore, it is subtracted from the capacity maximizing part, $\mathcal{X}$, of the reward. A multiplier $\mathcal{T}$ is used to provide fairness to the MUE and FUEs as in the (Tefft and Kirsch, 2013). However, the value of the $\mathcal{T}$ is user dependent which may have value between 5-20. A constant value of the $\mathcal{T}$ is proven effective in the simulation. In our case, FBS acts as the agents and therefore each FBS runs the Algorithm 1 individually. The proposed solution is n cooperative which is based on the multi-agent RL methodology as discussed by (Whitehead, 1991). According to the authors, agents in the RL scheme can share their knowledge and experience about the environment with the other agents (Whitehead, 1991). It is also evident from the literature that cooperation among the agents also reduce learning and search time (Busoniu *et al.*, 2008; Whitehead, 1991). Therefore, in the proposed methodology, FBS shares Q-Table with other FBS in its vicinity. The approach for computing the new Q-Table from the shared Q-.Tables, is based on the methodology used by (Amiri *et al*., 2018).

D. Simulation Setup: To simulate the proposed methodology, a system model composed of one MBS and M number of FBS is assumed. To reduce the complexity of the system model only one MUE and FUE is considered. However, in the future, more number ofMUEs and FUEs can be simulated for a more realistic environment. The MBS is located in the center of the cell FBS, FUE, and MUE may take any position in the cell.

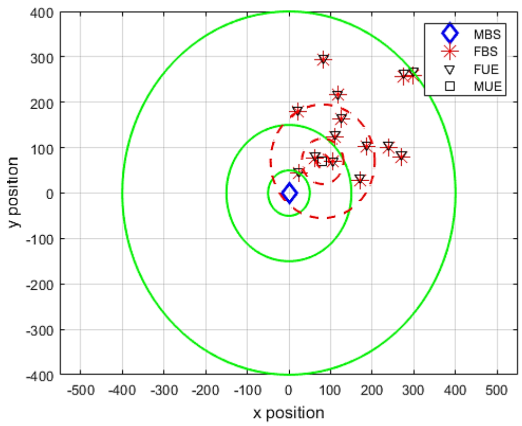

Fig-3: Simulation Model

However, to simulate the high density, M FBS, each with single FUE, are considered around the single MUE in the first quadrant at random locations. The random locations are simulated to make the system model close to a realistic situation. The system model is shown in the Fig-3.

To indicate the location of the FBS with respect to the distance from the MUE and MBS, three layers are used i.e. $N_{MBS} = N_{MUE} = N = 3$. However, in the future, a higher number of layers may be used to indicate a more precise location. The layer for MBS and MUE are indicated with green and red rings in the Fig-3. The state, s_t, of i^{th} FBS depends upon the distance from MUE and MBS as discussed in the last section.

In the simulations, the path loss is computed using the Log-Distance model and two assumptions are made to support the high-density realistic models which are *i)* residential area (Hossain *et al.*, 2012) and *ii)* indoor-outdoor propagation model for femtocells (Valcarce and Zhang, 2010).

$$PL = PL_0 + 10nlog_{10}\left(\frac{d}{do}\right)\#(12)$$

$where\ PL_0$ is Constant Path Loss and n is Path loss Exponent.

The parameters of the model are set as per the path loss model of the residential area in which $d_0 = 5, n = 4\ and\ PL_0 = 62.3dB$ (Hossain *et al.*, 2012). However, in the case of the indoor-outdoor propagation, which is suitable for simulations of the femtocells, path loss can be written as follows (Valcarce and Zhang, 2010):

$$PL = PL_i + PL_0\ \#(13)$$

$where$

$PL_i = Attenuation\ from\ the\ transmitter\ till\ the\ outer\ wall\ of\ the\ house$

$PL_0 = Outdoor\ attenuation$

Using the (6), (7) and Table-II proposed by the Valcarce & Zhang, 2010, PL_i and PL_0 can be written as follows:

$$PL_i = 6.1 + 10.6f - 1.8\,f^2\#(13a)$$

$$PL_0 = 62.3 + 32\log_{10}\left(\frac{d}{5}\right)\ \#(13b)$$

Where f is operating frequency in GHz. Rest of the simulation parameters are summarized in Table-2.

RESULTS AND DISCUSSION

In this section results of the simulation of the proposed scheme is presented and compared with the other state of the art solutions (Amiri *et al.*, 2018; Tefft and Kirsch, 2013). The results are compared in terms of the MUE capacity as function of FBS number, capacity of FUEs for every number of FBS and sum capacity of the FUEs as function of FBS number.

1) MUE Capacity: The comparison of the MUE capacity for the proposed solution and results presented by (Amiri *et al.*, 2018; Tefft and Kirsch, 2013) are shown in Fig-4. The simulation results show that the proposed solution significantly increased the capacity of the MUE in highly dense HetNets. The results of the proposed solution followed the trend of the results reported by Amiri et al., 2018 i.e. decrease in the MUE capacity

Table- 2: Simulation Parameters.

Parameter	Value
No. of MBS	1
No. of FBS	15
No. of MUEs	1
No. of FUEs per FBS	1
Radius of MBS	1000m
Radius of FBS	10m
$\boldsymbol{P_{MBS}}$	50 Watts
$\boldsymbol{P_{FBS}}$	-20dBm to 25dBm
$\boldsymbol{N_{Power}}$	31
$\boldsymbol{\xi_{MUE}}$ and $\boldsymbol{\xi_{FUE}}$	1 b/s/Hz
Learning Rate, $\boldsymbol{\alpha}$	0.5
Discount Factor, $\boldsymbol{\gamma}$	0.9
Number of Iterations	50000
Operating Frequency, $\boldsymbol{f}$	2.5GHz
Path Loss Model	Residential Area and Indoor-Outdoor Propagation

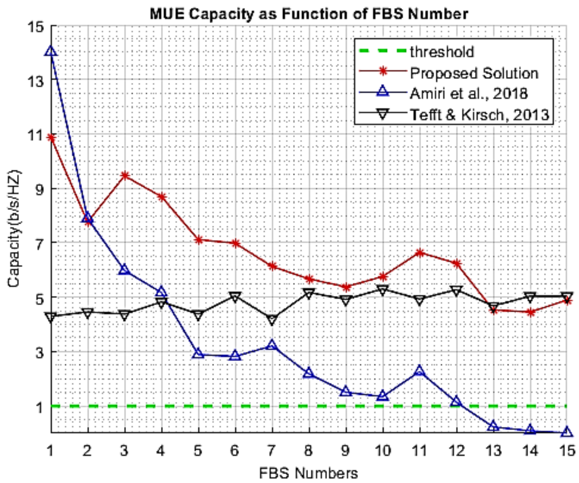

Fig-4: MUE Capacity, C_{MUE}

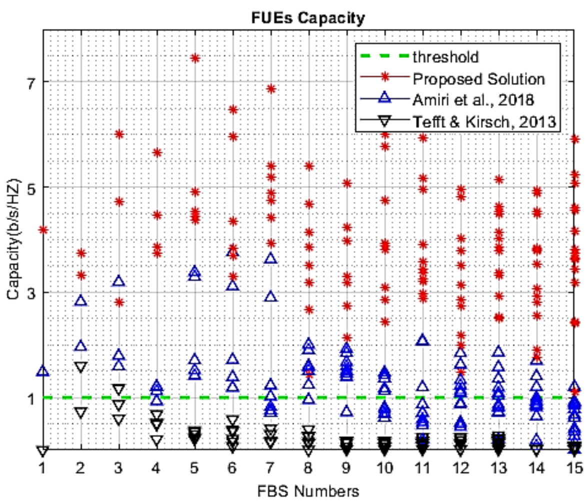

Fig-5: FUEs Capacity, C_{FUE_i}

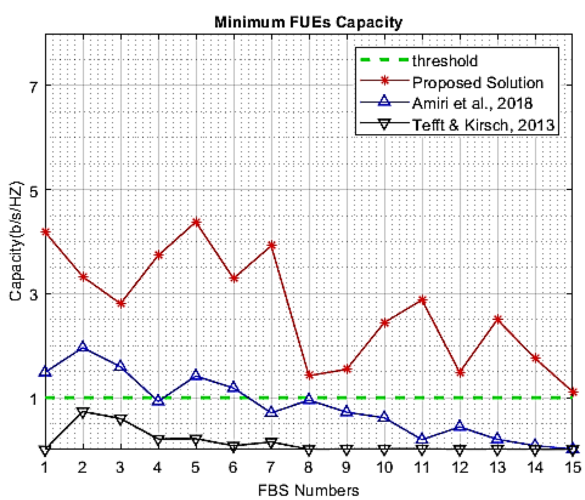

Fig-6: Minimum FUE Capacity, $C_{FUE_{min}}$

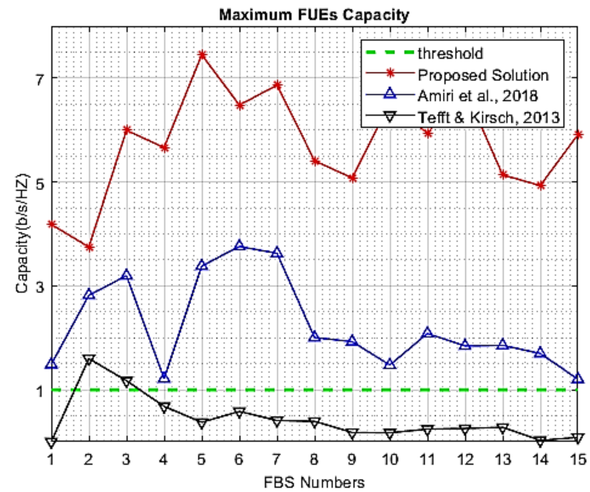

Fig-7: Maximum FUE Capacity, $C_{FUE_{max}}$

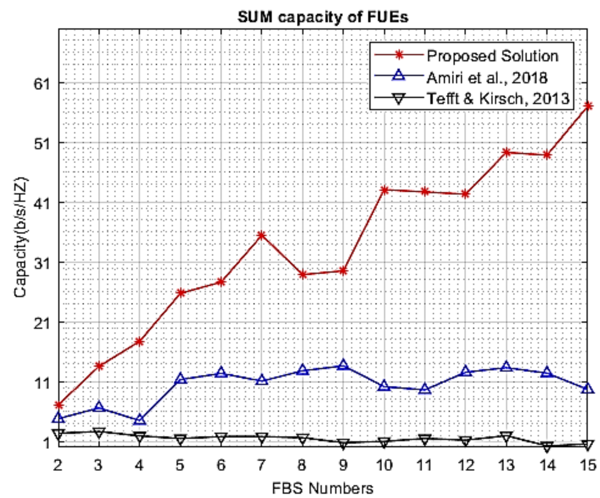

Fig-8: Sum Capacity of FUEs

with an increase in FBS number. However, for the higher number of FBS, it still remains higher the MUE capacity reported by (Tefft and Kirsch, 2013) and never falls below the QoS threshold.

2) FUEs Capacity: Comparison of the FUEs capacity using the proposed solution and the results presented by (Amiri *et al.*, 2018; Tefft and Kirsch, 2013) are shown in Fig-5. The capacity of the FUEs, using the proposed solution, is significantly higher than both of the state of the art solutions. Using the proposed solution, FUEs capacity never falls below the QoS threshold whereas for both (Amiri *et al.*, 2018; Tefft and Kirsch, 2013), the capacity of the FUEs decreases with an increase in FBS number and is almost zero after 14 FBS. A similar trend can also be observed for the minimum and maximum FUEs capacities shown in Fig-6 and Fig-7 respectively.

3) FUEs Sum Capacity: Like the FUEs capacity, the FUEs sum capacity also shows a remarkable increase with the increase in FBS as compared to the results of

(Amiri *et al*., 2018; Tefft and Kirsch, 2013) as shown in Fig-8. The proposed solution followed the rising trend of the result presented by (Amiri *et al*., 2018). However, sum capacity value for each FBS is significantly higher than that of (Amiri *et al*., 2018). The almost linear increase in the sum capacity of FUEs indicates that SINR of FUEs got significantly improved using the proposed solution.

Conclusion: In this paper a ML-based technique, Q-Learning is used for efficient power allocation to the FBS in highly dense HetNets. In high cross tier and co-tier interference scenarios, the power allocation optimization problem has a non-convex solution. However, the proposed ML technique solved the optimization problem successfully while maintaining the minimum QoS requirements for the MUE and FUEs. The simulation results show that despite the capacity of the MUE decreases with an increase in FBS number but a higher number of FBS can be added in the system while maintaining the minimum QoS for MUE. Similarly, a remarkable increase in the sum capacity of the FUEs is observed with the increase in FBS number which shows that the proposed solution effectively optimizes the FBS power to reduce the interference and increases the SINR of FUEs. In the future, more complex scenarios will be simulated keeping in view the other performance measuring parameters.

REFERENCES

Agiwal, M., A. Roy and N. Saxena (2016). Next Generation 5G Wireless Networks: A Comprehensive Survey. *IEEE Communications Surveys & Tutorials, 18*(3), 1617-1655.

Akpakwu, G.A., B.J. Silva, G.P. Hancke and A.M. Abu-Mahfouz (2018). A Survey on 5G Networks for the Internet of Things: Communication Technologies and Challenges. *IEEE Access, 6*, 3619-3647.

Amiri, R., H. Mehrpouyan, L. Fridman, R.K. Mallik, A. Nallanathan and D. Matolak (2018). *A Machine Learning Approach for Power Allocation in HetNets Considering QoS.* Paper presented at the 2018 IEEE International Conference on Communications (ICC).

Bin, W., G. Zhibin, H. Lianfen, T. Yuliang and C. Hongxiang (2014). *A Q-learning-based downlink resource scheduling method for capacity optimization in LTE femtocells.* Paper presented at the 2014 9th International Conference on Computer Science & Education.

Busoniu, L., R. Babuska and B.D. Schutter (2008). A Comprehensive Survey of Multiagent Reinforcement Learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 38*(2), 156-172.

Feng, Z., L. Tan, W. Li and T.A. Gulliver (2009). *Reinforcement learning based Dynamic Network Self-optimization for heterogeneous networks.* Paper presented at the 2009 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing.

Galindo-Serrano, A. and L. Giupponi (2010). *Distributed Q-Learning for Interference Control in OFDMA-Based Femtocell Networks.* Paper presented at the 2010 IEEE 71st Vehicular Technology Conference.

Gupta, A. and R.K. Jha (2015). A Survey of 5G Network: Architecture and Emerging Technologies. *IEEE Access, 3*, 1206-1232. doi:10.1109/ACCESS.2015.2461602

Hossain, E., V.K. Bhargava and G.P. Fettweis (2012). *Green radio communication networks*: Cambridge University Press.

Kurda, R., L. Boukhatem, T.A. Yahiya and M. Kaneko (2014). *Power adjustment mechanism using context information for interference mitigation in two-tier heterogeneous networks.* Paper presented at the 2014 IEEE Symposium on Computers and Communications (ISCC).

Parvez, I., A. Rahmati, I. Guvenc, A.I. Sarwat and H. Dai (2018). A Survey on Low Latency Towards 5G: RAN, Core Network and Caching Solutions. *IEEE Communications Surveys & Tutorials, 20*(4), 3098-3130.

Saad, H., A. Mohamed and T. ElBatt (2012). *Distributed Cooperative Q-Learning for Power Allocation in Cognitive Femtocell Networks.* Paper presented at the 2012 IEEE Vehicular Technology Conference (VTC Fall).

Sanchez, A., J. Arauz, J.W. McClure and Z. Miller (2016). *Cooperative self-organized optimal power control for interference mitigation in femtocell networks.* Paper presented at the 2016 IEEE Colombian Conference on Communications and Computing (COLCOM).

Tefft, J.R. and N.J. Kirsch (2013). *A Proximity-Based Q-Learning Reward Function for Femtocell Networks.* Paper presented at the 2013 IEEE 78th Vehicular Technology Conference (VTC Fall).

Tseng, F., L. Chou, H. Chao and J. Wang (2015). Ultra-dense small cell planning using cognitive radio network toward 5G. *IEEE Wireless Communications, 22*(6), 76-83. doi:10.1109/MWC.2015.7368827

Valcarce, A. and J. Zhang (2010). Empirical Indoor-to-Outdoor Propagation Model for Residential Areas at 0.9–3.5 GHz. *IEEE Antennas and*

Wireless Propagation Letters, 9, 682-685. doi:10.1109/LAWP.2010.2058085

Whitehead, S. D. (1991). *A complexity analysis of cooperative mechanisms in reinforcement learning*. Paper presented at the Proceedings of the ninth National conference on Artificial intelligence - Volume 2, Anaheim, California.

Yang, C., H. Dai, J. Li, Y. Zhang and Z. Han (2018). Distributed Interference-Aware Power Control in Ultra-Dense Small Cell Networks: A Robust Mean Field Game. *IEEE Access, 6*, 12608-12619. doi:10.1109/ACCESS.2018.2799138

Zhang, H., S. Chen, X. Li, H. Ji and X. Du (2015). Interference management for heterogeneous networks with spectral efficiency improvement. *IEEE Wireless Communications*, *22*(2), 101-107. doi:10.1109/MWC.2015.7096292.