IRACST – International Journal of Computer Networks and Wireless Communications (IJCNWC), ISSN: 2250-3501
Vol.5, No 5, October 2015

Interference Mitigation in Femtocell Network using Q-Learning

Abdullah Hizam Alqhtani
Department of Business Administration
Prince Sattam bin Abdulaziz University, Saudi Arabia
alqhtaniabdullah99@gmail.com

***Abstract*—Efficient resource allocation in femtocell networks has become necessary owing to the enormous advantages of having femtocells deployed in a heterogeneous network. However, the interference arising from this deployment necessitate a mechanism for mitigation and optimal control of resource allocation. Q-learning, as an example of a reinforcement learning technique has been widely used for this purpose with more emphasis on downlink interference problems. Using a simplified heterogeneous model comprising of one macrocell and two femtocells, we extend the use of Q-learning to specifically model and address the uplink interference problem. We show by means of controlled experiments, that the proximity of the users in the network to their respective base stations plays a significant role in the total capacity of the network. This has the potential to enable networks to act in an improved manner. Results from the simulation can be used to configure any realistic network model.**

Keywords. ***femtocell; Q-learning; Interference mititgation; reinforcement learning.***

I. INTRODUCTION

Femtocells are the new devices that have emerged over the last decade to tackle the interference problem faced by wireless cellular network operators [1]. The need to enhance the capacity and high coverage of mobile users has caused network service providers to deploy these femtocells while trying to mitigate the uplink and downlink interferences caused by the deployment of these devices. The interference is more common when the existing macrocells and the deployed femtocells share the same frequency or they operate in a dedicated frequency band. The former is a more common deployment because of the advantages it offers to network service providers. However, this leads to a complicated type of interference. When these femtocells are set up by the end user at random positions away from the macrocells, the interference problem becomes complex which then requires an efficient interference mitigation strategy. In addition, as the number of femtocells in a network increases, managing the interference becomes more crucial in order to satisfy the quality of service of not just the macrousers, but also guaranteeing a certain measure of service for the femtousers. Coupled with mitigating the interference is to find a tool for optimal resource allocation in femtocell networks which is the main bottleneck in the effective deployment of these networks of femtocells and macrocells.

II. RELATED WORK

In managing the interference, one common approach is to allow the agents—femtocells and macrocells—learn from the dynamic environment created by the coexistence of both cells. By learning from their environment, these devices can adjust their parameters such as power transmission level to satisfy the quality of service of their respective users. In literature, one common tool that has been used widely to achieve this learning is a reinforcement learning technique called Q-learning [2]. An advantage of Q-learning is that it does not need any prior information on the state of the environment. This means that we do not need to be concerned about the number of femtocells in the system or their spatial locations. Furthermore, the learning can be independently performed by each of the agents or it could be cooperative learning [3] where, for instance the femtocells share their information with each other. The actions taken by any agent in this setting affects the state or environment and also affects the learning process. One merit of Q-learning is that the agents can take actions while still learning from the actions of other agents. This learning leads to network acting in an improved manner [4]. This improved performance of the network and the easiness of the algorithm motivates us to adopt Q-learning in this work. Details of this learning algorithm will be discussed in the following section.

III. Q-LEARNING

In this section, the concept of Q-learning (QL) is presented. QL is one of the reinforcement learning (RL) algorithms. RL is a machine learning (ML) algorithm where an agent learns from a dynamic environment through trial-and-error interactions. In QL, an agent can find the optimal decision policies through interaction with the environment. These interactions can be modelled as a Markov Decision Process (MDP) [5],[6].

The Q learning model can be defined as a tuple (S, A, $P_{s,s'}$, $R(s, a)$) where $S = (s_1, s_2, ..., s_n)$ is a finite set of environment states the agent can enter, $A = (a_1, a_2, ..., a_m)$ is a finite set of actions the agent may choose, $P_{s,s'}(a)$ is the state transition probability function from state s to the new state s' after taking action a, and $R(s, a)$ is the reward function that determines the reward for the agent when performing action a in state s. In a system of more than one agent, the agents may decide to share their actions with other agents or may not. A very thorough description of how many agents may interact or learn has been reported by [7]. Emphasis is on the learning goal of the interacting agents.

The objective of each agent is to find an optimal policy $\pi^*(s)$, which tells the agent the optimal action to choose in each state $s_t \in S$ in order to maximize the the total expected discounted reward. The expected discounted reward over infinite time given the policy π and initial state s can be written as:

$$Q(s,a) = E\left\{\sum_{t=0}^{\infty} \gamma^t R(s_t, \pi(s_t), s_{t+1}) | s_0 = s\right\} \qquad (1)$$

where $\gamma \in [0,1)$ is called the discount factor which determines the significance of future rewards. If $\gamma = 0$, the agent will ignore all future rewards. If $\gamma = 0.9$, the agent will put emphasis on future rewards. In QL, Q(s, a) is called the Q value. When the agent selects action a according to the optimal policy $\pi^*(s)$, then the Q value (expected discounted reward) is maximized. The optimal Q value $Q^*(s,a)$ to find the optimal policy $\pi^*(s)$ is defined as:

$$Q^*(s,a) = E\{R(s,a)\} + \gamma \sum_{s' \in S} P_{s,s'}(a) max_{b \in A} Q^*(s',b) \qquad (2)$$

where $Q^*(s',b)$ is the optimal Q value from the the next state s' after choosing the next action b. QL process finds $Q^*(s, a)$ in a recursive manner by using the following update rule:

$$Q(s, a) = Q(s, a) + \alpha[r + \gamma \max_{b \in A} Q(s', b) - Q(s, a)] \qquad (3)$$

where α is the learning rate which determine how much the new information will override the old information. It was proved that the update rule in Equation (3) converges to the optimal Q-value under certain conditions [2][8]. One of the conditions is that the agent has to try all the possible actions and visit all the possible states infinitely. To achieve this condition, the ε greedy exploration is introduced in the Q-learning algorithm.That is, in each iteration, the agent chooses a random action with probability $\varepsilon \in (0,1)$ and chooses the greedy action that will maximize the Q-value with probability $1-\varepsilon$.

Related works that have used Q-learning to address interference problems and resource allocation in HetNET include [9],[10],[11]. The main idea in these works is to model the femto network as a multi-agent system where the femto user equipments are the agents. Depending on the context, these agents are in charge of managing the resources to be allocated to their femto base stations. The interaction between the agents and the surrounding environment leads to a learning of an optimal policy to solve the interference problem. The vast majority of the literature on Q-learning proposes algorithms that attempt to solve the power control problem, while showing improved performance in terms of convergence of the capacity, e.g. as in [10],[11]. The role of the choice of the reward function and how it can greatly affect the performance of the Q-learning algorithm was presented by [12]. The authors showed how the distance between the interacting agents and the total number of agents in the network can affect the capacity and the SINR. In addition, they affirmed the learning algorithm as an efficient tool for resource allocation and management in femtocell networks. Another significant work by [3] was aimed at solving the power control problem with Q-learning to manage the interference caused by the femtocells on the macrocells in the downlink. The authors showed that, using Q-learning, the femtocells can either learn independently or cooperatively, i.e. they share their "Q-table", in order to enhance their performance.

The number of related work described above and the opportunity to adopt the powerful Q-learning tool for resource allocation in a HetNET motivates us to explore the role of this algorithm to manage and control the uplink interference in a two-tier network. In the next section, we set up the system model which depicts the various agents and the associated interference patterns.

IV. SYSTEM MODEL

Figure 1 illustrates the system model we shall be investigating. The system consists of one macro base stations (MBS) underlaid with N_f Femto access points (FAPs). The marco user equipment (MUE) is randomly located in the coverage area of the MBS, while the femto user equipment (FUEs) are randomly located in their associated FAPs. Orthogonal scheduling is assumed in every FAP, so only one FUE is allowed to transmit in a single time slot to avoid interference within a femtocell.

According to our model, there is only one active FUE in each femtocell at every given time slot. Hence, we are only interested in the interference caused by the neighbouring active FUE and the neighbouring active MUE at the designated FAP. In this system, we consider the uplink co-tier and cross-tier interferences between femtocell and macrocell contrary to the downlink interference studied by [9]. In addition, we also analyze the power control problem in a single sub-channel OFDMA in this system. To simplify our study, we assume all the MUE and FUEs are on the same sub-channel and have the same amount of available resource blocks (RB), which allows to increase the spectral efficiency per area through spatial frequency re-use [3][9]. In our model, the length of the RB is the length of the power transmission levels. The performance of our system is analyzed using the common parameter of signal interference noise ratio (SINR) and capacity (C). Although there are other measures of importance in characterizing a mobile network of femtocells and macrocells such as the total throughput, our choice of estimating the SINR and C follows from recent studies such as the references described previously.

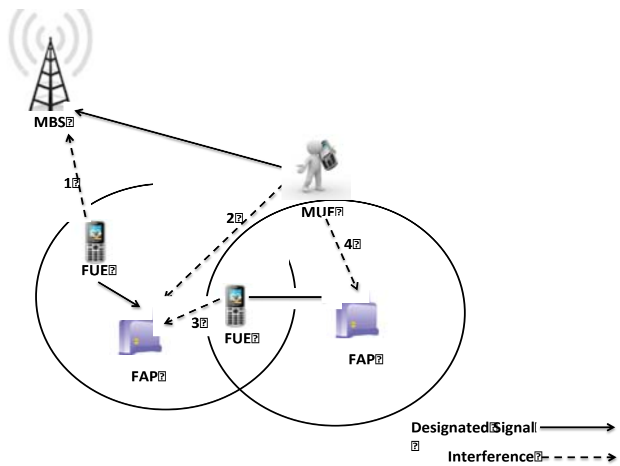

Figure 1. System Model

The throughput quantifies the volume of traffic in the network and this amount continues to increase steadily as technology advances. On the other hand, the capacity describes the maximum amount of signal strength the user equipments receives. Our study will be focused on the capacity and not the throughput which will be a subject of future research. The objective of the network service providers is to increase the capacity of the MUEs and FUEs. The SINR and the capacity are as defined in related literature such as [12], [9], [11], [13], and [14].

V. SIMULATION SCENARIO & PARAMETERS

The wireless network we consider consist of one macrocell underlaid with two femtocells. Each of the cells has an associated macro user and femtousers respectively. The channel gain between the transmitter i and receiver j depends on the distance between the transmitter and receiver. This is modeled as $g_{ij} = d^{-k}$, where d_{ij} is the physical distance between the ij transmitter i and receiver j and k is the path-loss exponent. In our simulations, we choose k = 4 just like [12]. In our network, we are interested in studying the uplink interference from the femtouser to the macro user and the interference between the two femtousers. Therefore, the signal to interference noise ratio of one FUE depends on the strength of the interference from the nearby MUE and the nearby FUE. Furthermore only one FUE is allowed to cause interference at the MUE while the MUE can cause interference at both FUEs because of the infinite range of possible distances between the transmitters and receivers, even in such a simplified network model, we tried several values for d in a bid to attain specific deterministic values for each pair of d_{ij}. Nevertheless, we initialized our simulations by setting $d_{ij} = 1$ to verify the usability of the simulation algorithm. The effect of this distance value will be explored in another related experiment.

The Q-learning aspect of our simulation considers two main factors - the reward function and the thresholds that were chosen adaptively. simulation. Computation of the SINR followed the conventional approach in literature, so we set the noise power $\sigma^2 = 0.1$ in line with [12]. In addition, the learning rate $\alpha = 0.5$ and the discounted factor used in the Q-learning algorithm is set to 0.9. Finally, the simulation is implemented in MATLAB on a desktop.

At the beginning of our simulation, we assume all the agents' distance is exactly the same, so we fix the distance to a unit of 1. This implies that the path loss is the same and this constant path loss is used in the Q-learning algorithm. This fixed condition on the distance means that the gains at each base station will be exactly the same and the SINR will simply be dominated by the choice of transmission level denoted by the random action the agents take at every iteration.

Setting the threshold parameters as ($\gamma_T = 2$), we performed 1000 Q-learning iterations where at each iteration, the agents choose a random action corresponding to one of the transmission power levels in the range 2dB, 4dB, 18dB. The number of iterations was chosen as a reliable statistic for the Q-learning simulation and it is similar in the related literature. In the first half of the iterations, the agents choose random actions, resulting in oscillating capacities as shown in Figure 2 while the agents continue to learn the actions of other agents. Finally information from the Q-table of each agent is used from the second half of the iteration leading to a convergence in the system for each of the FUEs and the MUE. As seen from the Figure, the capacity of the macrocell (labelled as MUE) is higher than the respective femtocells (labelled as FUEs), a condition we prefer for high QoS especially at the macrocell. In terms of scaling, we also see that the closer the agents, the stronger the interference leading to high magnitudes for the capacities.

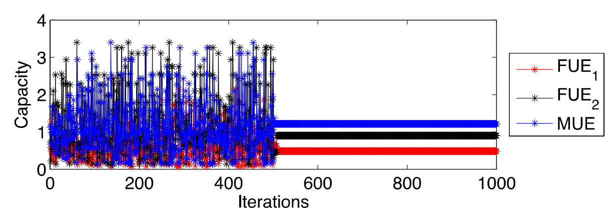

Figure 2. Macrocell and femtocells' capacity as a function of iterations. Independent Q-learning with agents at a common distance

On the other hand, a visualization of the cumulative capacity as the iteration progressed shows very interesting information. Figure 3 shows that the FUE that is not experiencing a cross-tier interference from the other FUE attains a capacity that approaches the capacity of the MUE. This phenomenon is probably attributed to the fact that the two FUEs did not share their information, calling for the prospect of a distributed learning or cooperative learning idea.

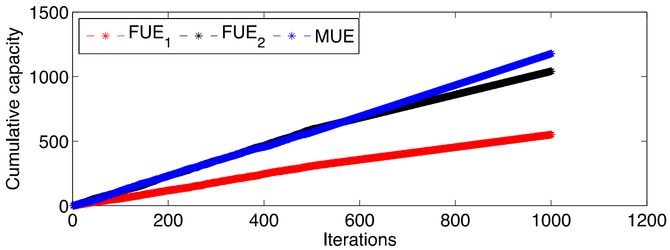

Figure 3. Cumulative capacity at the base stations as a function of iterations. Agents at common distance

Finally, in this first phase of experiments, we computed the aggregate capacity of the femtocells and the macrocells. This is necessary to determine the load or throughput in the network. By taking the sum of the individual femtocell capacities at each iteration and plotting it against the aggregate or sum capacity of the macrocell, we observe that the total FUE capacity is more than that of the macrocell as shown in Figure 4.

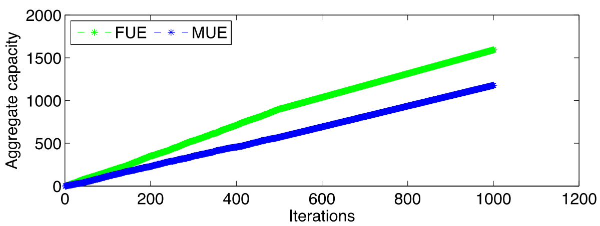

Figure 4. Total capacity of the FUE and the MUE in the network as a function of iterations using independent Q-learning of agents at common distance

Again this result is expected given that the agents (FUEs) do not cooperate in the learning process. Another possible explanation is the fact that we have assumed a constant gain which relates to the spatial location of the FUEs and MUEs from their associated access points or/and the stations they interfere with.

VII. CONCLUSION

By means of numerical simulations and results, we have evaluated the performance of Q- learning algorithm. Our system comprising of one macrocell and two femtocells can be extended to multiple macrocells and femtocells. We showed how convergence can be attained using the Q-learning algorithm. We also illustrated how we can achieve a high quality-of- service for the macro user equipment while dealing with the uplink co-tier interference from the femtouser(s). The position of the macrouser and femtousers can also be varied in the network. The closer the interfering agent is to the base station, the stronger the interference.

REFERENCES

[1] Saquib, Nazmus, Ekram Hossain, Long Bao Le, and Dong In Kim. ”Interference manage- ment in OFDMA femtocell networks: Issues and approaches.” Wireless Communications, IEEE 19, no. 3 (2012): 86-95.

[2] C. J. C. H. Watkins and P. Dayan “Technical note Q-learning,” Journal of Machine Learning, vol. 8, pp. 279-292, 1992.

[3] Saad, Hussein, Amr Mohamed, and Tamer ElBatt. ”Distributed cooperative Q-learning for power allocation in cognitive femtocell networks.” Vehicular Technology Conference (VTC Fall), 2012 IEEE. IEEE, 2012.

[4] Giupponi, Lorenza, Ana Galindo-Serrano, Pol Blasco, and Mischa Dohler. ”Docitive networks: an emerging paradigm for dynamic spectrum management [dynamic spectrum management].” Wireless Communications, IEEE 17, no. 4 (2010): 47-54.

[5] Hu, Junling, and Michael P. Wellman. ”Multiagent reinforcement learning: theoretical framework and an algorithm.” ICML. Vol. 98. 1998.

[6] Michael, T., and I. Jordan. ”Reinforcement learning algorithm for partially observable Markov decision problems.” Proceedings of the Advances in Neural Information Process- ing Systems (1995): 345-352.

[7] Busoniu, Lucian, Robert Babuska, and Bart De Schutter. ”A comprehensive survey of multiagent reinforcement learning.” Systems, Man, and Cybernetics, Part C: Applica- tions and Reviews, IEEE Transactions on 38.2 (2008): 156-172.

[8] E. H. Norman “Convergence of Q-learning: A simple proof,” Institute Of Systems and Robotics, Tech. Rep. Relations.

[9] Galindo-Serrano, Ana, and Lorenza Giupponi. ”Distributed Q-learning for interference control in OFDMA-based femtocell networks.” Vehicular Technology Conference (VTC 2010-Spring), 2010 IEEE 71st. IEEE, 2010

[10] Simsek, Meryem, Andreas Czylwik, Ana Galindo-Serrano, and Lorenza Giupponi. ”Im- proved decentralized Q-learning algorithm for interference reduction in LTE-femtocells.” In Wireless Advanced (WiAd), 2011, pp. 138-143. IEEE, 2011.

[11] Bennis, Mehdi, and Dusit Niyato. ”A Q-learning based approach to interference avoid- ance in self-organized femtocell networks.” GLOBECOM Workshops (GC Wkshps), 2010 IEEE. IEEE, 2010.

[12] The Japan Reader “A proximity-based Q-learning reward,” function for femtocell net- works” in IEEE Vehicular Technology Conference, Sept. 2013

[13] Dhahri, Chaima, and Tomoaki Ohtsuki. ”Learning-based cell selection method for fem- tocell networks.” Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th. IEEE, 2012.

[14] Saad, Hussein, Amr Mohamed, and Tamer ElBatt. ”A cooperative Q-learning approach for distributed resource allocation in multi-user femtocell networks.” Wireless Commu- nications and Networking Conference (WCNC), 2014 IEEE. IEEE, 2014.