Power Optimization in a Non-Coordinated Secondary Infrastructure in a Heterogeneous Cognitive Radio Network

Tauseef Ahmed[1], Yannick Le Moullec[1]
[1]*Thomas Johann Seebeck Department of Electronics, Tallinn University of Technology,*
Ehitajate tee 5, 19086 Tallinn, Estonia
tauseef@elin.ttu.ee

Abstract—**In this paper we describe a novel approach that combines dynamic spectrum allocation and transmission power optimization for the secondary network users in an heterogeneous cognitive radio network. The proposed approach builds upon reinforcement learning and convex optimization procedures. Furthermore, the several key components, i.e. inter-cell interference, path loss, and fading have been considered when designing the power optimization algorithm. Simulation results show that the proposed approach improves the QoS of the system by up to 10 dB in terms of SINR and by up to 4% in terms of spectral efficiency while maintaining the average dissatisfaction probability close to that of the non-optimized approach.**

Index Terms—**Cognitive radio, dynamic spectrum allocation, heterogeneous network, power optimization, radio spectrum management.**

I. INTRODUCTION

Cognitive Radio has been attracting a significant interest during the last decade. It was triggered by DARPA's approach on Dynamic Spectrum Access network, with the so-called NeXt Generation (xG) program to solve the current spectrum inefficiency, claimed to be a real bottleneck for the progress of wireless telecommunication. Since then, the problem has been recognized to be not so much spectrum scarcity per se, but more its efficient exploitation. At this point, the term *opportunistic network* has been coined, which devises a plan to effectively and efficiently use the available radio resources. The opportunistic use of the radio spectrum is one of the key benefits of cognitive radio. Thus, many contributions dealing with the sensing of primary users spectrum and its related link layer issues (e.g. power control, modulation schemes, etc.) have been published.

However, a major challenge to realizing the potential benefits of cognitive radio lies in the interference management between non coordinated secondary users and primary users, with the aim of sharing the available spectrum.

In this paper, we consider uncoordinated secondary networks that are asking to opportunistically share, in an optimum way, the spectrum owned by primary networks without damaging the QoS of the licensed users beyond certain agreed limits. In this work the secondary networks consist of a unique base station which is providing services to the secondary users. We also consider the static load traffic for which each secondary network has to allocate spectrum in an adaptive way. Novel procedures relying on reinforcement learning (RL) [1]–[4] based algorithms are presented (see II.B) to deal with the uncoordinated and opportunistic spectrum sharing problems. We present the study of a decentralized approach for the dynamic spectrum and power allocations in multi-cell orthogonal frequency division multiple access (OFDMA) networks. Each cell independently decides i) the frequency allocation using the RL algorithms and ii) the power allocation based on convex optimization algorithms. In OFDMA, the broad frequency spectrum is divided into smaller bandwidth frequency resources called chunks. While assigning the frequency units, i.e., chunks, the aim is to reduce the inter-cell interference i.e., the interference caused to each other by two or more neighboring cells that use the same frequency resources. The assignment of the power levels is based on convex optimization algorithms [4], where the key factor in deciding the power allocation is inter-cell interference and other degradations.

II. SYSTEM MODEL

A. Decentralized Network Architecture

We consider a decentralized network architecture composed of a hybrid environment of primary and secondary networks. Each secondary entity, i.e. cell, comprises an independent RL agent which performs the spectrum allocation task keeping in mind the objective function of maximizing the signal to noise and interference ratio (SINR) while keeping in consideration the cell users QoS requirement (i.e. spectral efficiency). Considering that a cell has U users at any moment, the secondary base station (SBS), before every assignment, checks the generated inter-cell interference by the U users, and the interference to the primary base station. Note that in this particular example, for simplicity's sake, we are assuming that primary users are not present; more advanced cases will be presented in a future publication. A generalized OFDMA radio interface is

considered for the downlink for users' data transmission. The total system bandwidth W is divided into N chunks, the smallest unit that can be allocated. A chunk is a group of contiguous OFDMA subcarriers with bandwidth $B = W/N$ Hz. Frames are divided into time slots. The minimum radio resource block which is available to users is one chunk per frame. There is an uplink control channel where users send frame-by-frame measurement reports.
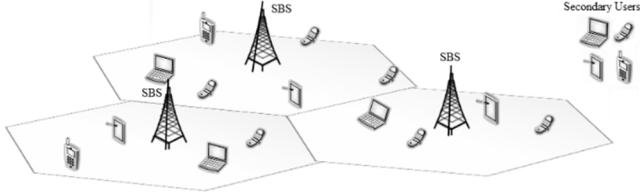

Fig. 1. A typical contiguous 3 Cells' deployment for Secondary Network used in these simulations.

A typical macro cell (MC) based cellular scenario on a geographical location, as shown in Fig. 1, consists of 3 cognitive radio SBSs which are serving secondary users in their vicinity. For simplicity, we consider only secondary users that are using various services and sharing the primary spectrum among themselves. These SBSs allocate both spectrum and power to their users in a non-coordinated or decentralized way. There could also be overlapping areas covered by several SBSs because they are not coordinated and could be run by different operators/vendors. However, in this work we assume no overlap between the cells.

B. Cell Operation

In the short term the cell handles users' traffic and performs the OFDMA fast link adaptation following the channel aware strategy proportional fair (PF) [5]. On the other hand, the spectrum assignment is done on a medium term basis. Specifically, each cell tries to learn the best resources assignment scheme, i.e., frequency and power, by executing the reinforcement learning dynamic spectrum assignement (RL-DSA) algorithm [1], [6] and convex optimization algorithm [3], in an execution period of L frames. On the first execution, a cell randomly selects the initial time to start the proposed combined algorithm; the algorithm first assigns the initial frequencies and then receives the reward signal (SINR) from the environment. The RL-DSA is internally based on random variables and Bernoulli logic.

The key steps of RL-DSA algorithm (described in appendix A) tries various assignments and the one which gives the highest reward (once the the algorithm has converged) is selected, i.e. its frequencies are assigned to the cell. The next execution occurs after L frames. Hence, large values of L are expected for a medium range execution of RL-DSA and water-filling algorithm. The probability that adjacent cells select the same initial time becomes negligible. The individual steps of the algorithm are further detailed in [1], [6].

The objective is to perform both an optimal frequency allocation and power allocation to each SBS so that a maximum throughput (or efficiency) per SBS can be attained, while at the same time the following constraints are satisfied:

– Each SBS should provide service to $U = 15$ users, ensuring a minimum bit rate to each of them in accordance with the considered service. There could be several service types;
– Generated interference should be minimum, i.e., interference to the primary users should remain below the primary threshold value; since we are considering that no primary user exists in the area, the condition of the interference is for the inter-cell interference.

In order to perform a reliable spectrum allocation, the requirement is the user satisfaction. In order to fulfill the users' QoS, we should estimate the spectrum usage in the adjacent cells to calculate the potential inter-cell interference. Previously, frequency allocation optimization with constant chunk powers [1] has been used; in this paper we propose a new spectrum assignment method in which both frequency and power are optimized. The assignment procedure is a two-step process, in which deciding i) the frequency allocation (chunks) is performed as summarized in Appendix A (for details see [1], [6]) and ii) the transmitted power for each frequency chunk is performed as described in the next section. The RL-DSA in our spectrum management has been revised in order to take inter-cell interference into consideration.

C. Power Allocation

Power allocation is based on a convex optimization problem with the objective function given in (1)

$$f = max_{P_{n,l}} \sum_{n \in C(l)} \log_2 \left(1 + \frac{\overline{x} P_{n,l}}{\Gamma \dagger_{n,l}^2} \right), \qquad (1)$$

where $C(l)$ is the set of chunks currently allocated to cell l, $P_{n,l}$ is the power assigned to chunk n in the lth cell.   is the average fading. $^2_{n,l}$ is the average noise plus interference defined in (2) and is reported or measured by a generic user at chunk n coming from each one of the interfering cells $c \in A(n)$ (where $A(n)$ is the set of cells with chunk n allocated) at the time when the resource allocation is updated. $\gamma_{n,l}$ is the channel gain (in accordance with the propagation model including slow fading) associated to chunk n in cell l

$$\dagger_{n,l}^2 = P_{noise} + \sum_{\substack{c \in A(n) \\ c \neq l}} I_n^c, \qquad (2)$$

where P_{noise} is the noise power and I_n^c is the received interference for that particular frequency chunk from the other cells which are also using that chunk. There are two main constraints for the power algorithm. The first constraint, which is described in (3), is the maximum power at cell l

$$\sum_{n \in C(l)} P_{n,l} \leq P_{max,l} = \min(P_{max,l}, \sum_{n \in C(l)} P_{TH,n,l}), \qquad (3)$$

where $P_{max,l}$ is the total maximum power available at cell l and $P_{TH,n,l}$ is the maximum power allowed in chunk n in order not to interfere.

The latter is the second constraint described in (4)

$$P_{TH,n,l} \geq P_{n,l} \geq 0. \tag{4}$$

If chunk n is not used then $P_{TH,n,l} = $, and thus the second constraint has no effect. The solution to the power optimization problem is given by the classical water-filling approach [3], [7]. The detailed formulation of the power optimization is beyond the scope of this paper and will be presented in subsequent work.

III. SIMULATIONS

We consider a downlink OFDMA-based 3 MC scenario; we focus our study on two case studies. First, we use RL-DSA with constant power assignments where all assigned chunks are assigned equal powers (*Case A*). In the second situation, we use the power optimization algorithm in which all the chunks use different powers based on the surrounding situation (*Case B*). Users are homogenously scattered in the cellular zone and they are not moving, i.e., for simulation purposes the users do not change their geographical positions and handovers are not considered. Also during the entire course of action, the cell load is static, i.e., the numbers of users do not change. Users always have data ready to send, which means every user will try to occupy as much bandwidth as they can, (full buffer traffic model [8-10]). The performance of the system is measured on the basis of spectral efficiency, SINR and the users' dissatisfaction probability, over one simulated hour. The spectral efficiency is the QoS parameter defined as a performance metric that measures the amount of successfully delivered bits per unit of time and spectrum. The dissatisfaction probability is defined as the percentage of seconds in which the user throughput is below a target throughput called the satisfaction throughput. In the simulations, the user satisfaction throughput is set to 256 Kbps. Other simulation parameters are presented in Table I.

TABLE I. SIMULATION PARAMETERS.

General	
Frame time	2 ms
Chunk Bandwidth [B]	375 kHz
Number of Chunks [N]	6
UE Thermal Noise	-174 dBm/Hz
UE Noise Factor	9 dB
Short Term Scheduling (STS) Method	Proportional Fair (PF)
PF Averaging Window	50 Frames
Spectral Efficieny (theoretical maximum)	5 (bits/s)/Hz [11]
Secondary BS Cell	
Cell Radius	500 m
Maximum BS Power	43 dBm
Minimum Distance to BS	35 m
Antenna Pattern	Omnidirectional
Path Loss at d Km in dB	128.1 + 37.6log10(d)
Shadowing Standard deviation	8 dB
Showing decorrelation distance	5 m
Small Scale Fading Model	ITU Ped. A
RL-Spectrum Algorithm	
Measurement Averaging Period [l]	2500 Frames
RL-DSA Execution Period [L]	60000 Frames
RL-DSA parameters [, , , l	[100, 0.00001, 0.05, 10^{-6}]
RL-DSA Exploratory Probability [$p_{explore}$]	0.1 %
RL-DSA Steps [MAX_STEPS]	1000000

We are simulating for the two above scenarios (Case A and Case B), i.e., with and without power optimization algorithms, and then the results are compared. All simulations have been performed with Matlab.

A. Case A: Frequency Allocation with Constant Chunk Power

There are 15 users in each cell and 6 chunks to be allocated. Each cell requires 3 chunks to satisfy the users' communications. The users are satisfied most of the time, and they do not suffer from resource scarcity. Usually when one cell's users obtain higher spectral efficiency, the other cells experience reduced spectral efficiency due to the inter-cell interference. Since there are only 6 chunks available to be assigned for each cell, some of the chunks are reused, giving birth to the inter-cell interference. When one cell uses the chunk, which is being used in other neighboring cell or cells, inter-cell interference is generated.

B. Case B: Frequency Allocation with Power Optimization

In this part of the simulations we have evaluated the proposed allocation scheme. The combined frequency and power allocation based on RL-DSA and Convex Optimization algorithm is a sub-optimal approach because we do not optimize the frequency and power while performing the resource allocation algorithms. The procedure is as follows:

1. The frequency allocation is carried out assuming a feasible constant power setting as done in the first part of the simulations so that the conditions on the power can be satisfied.

2. The set of allocated frequencies, $C(l)$, to cell l is retained and then the convex optimization is used to obtain the power setting $P_{n,l}$, from (1), per chunk in each cell l.

3. Steps 1 and 2 are repeated for the cell l with the new power settings to obtain the new frequency and power allocations.

The concept behind the whole procedures is that the first time the frequency allocation is performed by the RL algorithms using constant powers, exactly as described in the previous section, and then once the frequency allocation is known, the power allocation algorithm computes the powers for the individual chunks based on how much it received inter-cell interference and fading. When this power allocation is done for all chunks in the cell, then the RL algorithm is executed for these optimized powers to obtain the new frequency allocations. This process is continued until we reach the convergence in the power optimization algorithm. This procedure is done by all the SBS cells after the L frames. Now the chunks are assigned powers individually and the total power which the SBS can allocate is assigned to the chunks depending upon the parameters from the environment taken into account by the power allocation algorithms. Two of the most important parameters which the algorithm considers are the inter-cell interferences and fading.

C. Results

The simulation results from Cases *A* and *B* are presented in Fig. 2–Fig. 4. When comparing the results, it can be seen that better performance is achieved when using the power optimization. Firstly, as shown in Fig. 2, the spectral

efficiency of the power-optimized system (Case *B*) is higher than that of the non-optimized case throughout the simulation by up to 4 %. Secondly, as shown in Fig. 3, the average SINR of the system increases by up to 10 % thanks to the power optimization. Although the average SINR somehow decreases and fluctuates in the middle of the experiment for the power-optimized case, it still has better results than Case *A (constant power);* in the worst case, the gain is 0 dB. Finally, as shown in Fig. 4, the average user dissatisfaction probability is somewhat similar to that of Case *A*. Thus it can be concluded that in general the system offers better performance in terms of throughput (spectral efficiency) and SINR while providing the same level of user satisfaction.
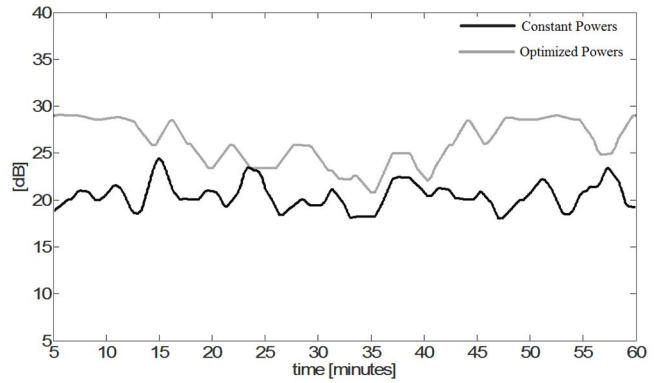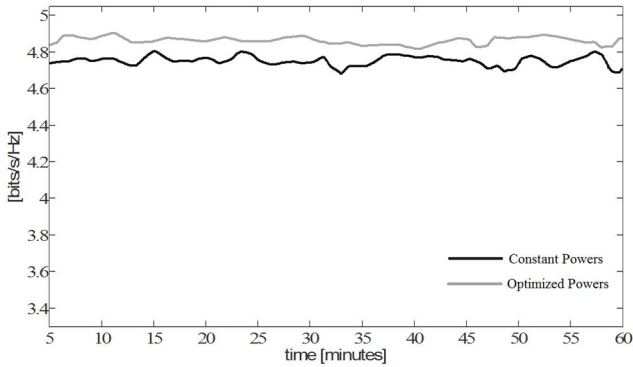

Fig. 3. Average SINR.

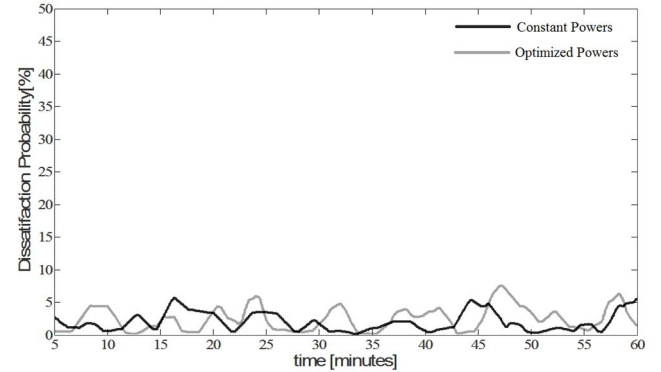

Fig. 2. Average Spectral Efficiency.

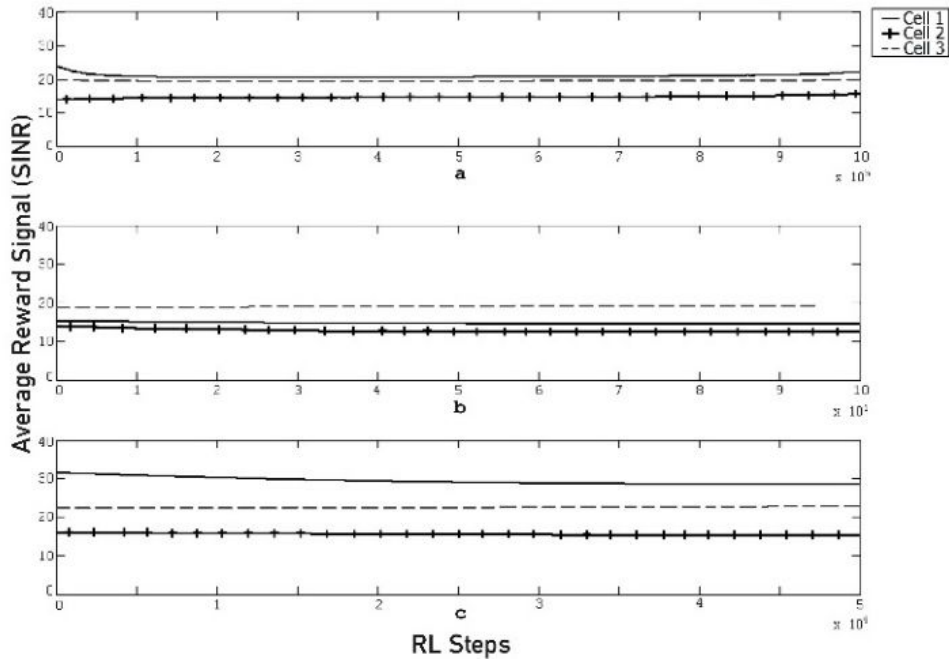Fig. 4. Average Dissatisfaction Probability.

Fig. 5. RL Convergence Behaviour.

D. Convergence Study

The convergence behavior of the RL-DSA coupled with power optimization algorithm is given in Fig. 5. The convergence behavior is studied over three different maximum steps of (RL_MAX_STEP), i.e., a = 1000000, b = 100000, and c = 50000, where RL convergence steps are set to 5000 (which is experimentally chosen over multiple iterations). The convergence condition is set to 0.01. The convergence behavior is studied for three cells; from Fig. 5 it is quite evident that with the inclusion of the power algorithm with RL-DSA, the convergence behavior is quite in accordance with [6] and convergence is achieved for a, b and c (typically, for RL-based method, this value should be ca. 3000).

IV. CONCLUSIONS

Despite the sub-optimality of the RL-DSA, its combination with power optimization offers better performance than the techniques proposed in [1] and [6], while converging reasonably well for all 3 cells. Future

work will address more complex scenarios with dynamic system and higher numbers of cells and users for the power management algorithm. Furthermore, we will evaluate the applicability of such approaches when adding cognitive capabilities to wireless sensor networks. Indeed, adding cognitive capabilities to wireless sensor networks is highly desirable since the resulting cognitive wireless sensor networks (CWSN) could then feature, among other things, dynamic spectrum allocation and energy optimization, thereby enabling them to better cope with spectrum scarcity and limited battery life-times. In particular, we will address the question of designing such dynamic algorithms so that their implementation on computationally and energy limited resources do not outweigh the expected benefits. Another key aspect that should be investigated is the design and implementation of power management and optimization techniques to deal with fluctuating energy sources in CWSN powered by energy harvesters.

APPENDIX A

RL-DSA is based on the Bernoulli logi unit. The internal architecture of the RL works on the weighted probabilities which are updated on every iterations including the interaction with the environment. The key steps involved in the frequency allocations are listed here and the details of every step is available in [1], [6].

```
1.   REPEAT
2.     Received reward signal from the
       environment.
3.     Update the average reward.
4.     FOR all cells AND chunks
5.       Update the internal probabilities of
         the RL – agent.
6.     END FOR
7.     FOR all cells AND chunks
8.       IF internal probabilities for the
         cell status is greater than the
         threshold value (criteria set by user)
9.         Assign that frequency chunk to the
           cell
10.      ELSE
11.        Do not assign the chunk.
12.      END IF
13.    END FOR
```

REFERENCES

[1] F. Bernardo, R. Agusti, J. Perez-Romero, O. Sallent, "Intercell interference management in OFDMA networks: a decentralized approach based on reinforcement learning", *IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Reviews,* vol. 41, no. 6, pp. 968–976, 2011. [Online]. Available: http://dx.doi.org/10.1109/TSMCC.2010.2099654

[2] R. J. Williams, "Simple statistical gradient-following algorithms for connectionist reinforcement learning", *Machine Learning*, vol. V8, no. 3, pp. 229–256, 1992. [Online]. Available: http://dx.doi.org/10.1007/BF00992696

[3] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2009.

[4] R.-S. Sutton, A.-G. Barto, *Reinforcement Learning: An Introduction*, The MIT Press, 1998.

[5] C. Wengerter, J. Ohlhorst, A. G. E. von Elbwart, "Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA", in *Proc. IEEE 61st Veh. Technol. Conf.,* 2005, vol. 3, pp. 1903–1907. [Online]. Available: http://dx.doi.org/10.1109/vetecs.2005.1543653

[6] F. Bernardo, R. Agusti, J. Perez-Romero, O. Sallent, "A novel framework for dynamic spectrum management in multicell OFDMA networks based on reinforcement learning", *IEEE Wireless Communications and Networking Conf., (WCNC 2009)*, 2009, pp. 1–6. [Online]. Available: http://dx.doi.org/10.1109/wcnc.2009.4917524

[7] T.-M. Cover, J.-A. Thomas, *Elements of Information Theory*, John Wiley & Sons, 2006.

[8] NGMN Alliance, "Next Generation Mobile Networks Radio Access Performance Evaluation Methodology", White paper, 2008. [Online] Available: http://www.ngmn.org

[9] IEEE 802.16 Broadband Wireless Access WG, "IEEE 802.16m Evaluation Methodology Document", 2008. [Online] Available: http://www.ieee802.org

[10] 3GPP, "Further Advancements for E-UTRA Physical Layer Aspects", Tech Rep 36.814 v9.0.0, 2010. [Online] Available: http://www.3gpp.org

[11] R. Schoenen, R. Halfmann, B. H. Walke, "MAC Performance of a 3GPP-LTE Multihop Cellular Network", *IEEE Int. Conf. Communications, (ICC 2008)*, 2008, pp. 4819–4824.