Use of Learning, Game Theory and Optimization as Biomimetic Approaches for Self-Organization in Macro-Femtocell Coexistence

Ali Imran[1], Mehdi Bennis[2] and Lorenza Giupponi[3]

[1]. Center for Communication Systems Research, University of Surrey, UK, e.mail:a.imran@surrey.ac.uk

[2]. Centre for Wireless Communications, University of Oulu, Finland, e.mail: bennis@ee.oulu.fi

[3]. Centre Tecnológica de Telecomunicacions de Catalunya (CTTC)

Parc Mediterrani de la Tecnologia, Av. Carl Friedrich Gauss 7, Barcelona, Spain 08860 e.mail: lorenza.giupponi@cttc.es

Abstract**—**In this paper, we present the use of several Biomimetic approaches for Self Organization (SO) in heterogeneous scenarios where macrocell and femtocell networks coexist. Mainly these approaches are categorized in *indirect biomimetics* and *direct biomimetics*. Under indirect biomimetics we discuss 1) emerging paradigms in learning theory and 2) game theory for their potential to enable SO solutions in heterogeneous networks. By means of numerical results we demonstrate the pros and cons of these indirect biomimetic approaches for designing SO in macro-femto coexistence scenarios. Furthermore, we demonstrate the use of direct biomimetic approaches for designing SO by exploiting one to one mapping between a natural SO system and our system model for heterogeneous networks based on Outdoor Fixed Relays (OFR). Numerical results show that the proposed analytical solution can enhance wireless backhaul capacity of the OFR based femtocells by adapting the macro base station (BS) antenna tilts in a distributed and self organizing manner.

Keywords— Biomimetics, game theory, machine learning, self-organization.

I. INTRODUCTION

Biomimetic is a branch of science that investigates natural systems aiming at exploiting their working principles for improvement in the design and operation of man-made systems. In nature, there are myriad of examples of self-organizing behavior, e.g., ants colonies finding shortest routes to food sources, termites collectively building complex constructions without using a blueprint, fish shoals organizing themselves without a leader, and swarms of fireflies synchronously emitting light flashes. The fact that SO is originally a bio inspired phenomenon, and the abundance of perfect SO in biological systems, makes biomimetics a rich paradigm to investigate the constituents and working principles of SO for engineering systems.

Since SO has been defined as a fundamental cornerstone of future cellular networks [1] - due to the impromptu deployment of femtocells in future heterogeneous networks [2] - a number of bio inspired techniques have been ventured on in the literature to develop SO solutions in this context. These techniques include evolutionary heuristics like genetic algorithms and neural networks, machine learning, cellular automata and game theory. A detailed survey of these techniques can be found in [3]. Broadly speaking, these techniques can be classified under two different subcategories within biomimetics: the *direct* and the *indirect approach*, as explained below [4]:

Direct Biomimetics: In the direct (or top-down) approach, an engineering problem is tackled by looking for natural systems solving an equivalent problem. The biological solution and its principles are then analyzed and re-built in a technical application. Examples of the direct approach are the design of airplane wings that directly copy the gliding flight of birds or design of camera that copies design of human eye etc.

Indirect Biomimetics: In contrast, the indirect (or bottom-up) approach of bio-inspired design involves, first, the derivation of principles by analyzing natural systems. The principle is then abstracted from its biological context and used in the technical applications where it could be suitable. Examples of such indirect approach are the concept of artificial intelligence, which attempts to exploit the human learning behavior, or the concept of game theory, which aims to exploit the findings from dynamics of a free economy in various engineering applications.

The rest of this paper is organized as follows. In section II, we discuss and demonstrate the use of two key indirect biomimetic approaches, i.e., *learning* and *game theory,* and we compare their pros and cons through numerical results in a macro-femto coexistence scenario. In learning theory, we particularly focus on an emerging paradigm of *docitive learning* and demonstrate how it is more promising in heterogeneous network scenario compared to conventional learning techniques. While the indirect approaches are currently more popular in the literature for developing solutions for wireless systems, in section III, we present direct biomimetics for developing SO solutions for heterogeneous networks. To this end, we establish one to one analogy of our system model with a SO system in nature that addresses a problem analogous to ours. Based on this analogy we demonstrate the novel approach of developing SO solution for heterogeneous network scenario, by use of direct biomimetics. The proposed solution is developed for capacity enhancement on the backhaul access links for the Outdoor Fixed Relay (OFR) scenario, by SO of macro Base Station (BS) antenna tilts. Numerical results show that substantial increase in spectral efficiency can be achieved through SO of BS antenna achieved by proposed solution, without relying on heavy signaling. Finally, Section IV summarizes the main conclusions of the work.

II. Indirect Biomimmetic Approaches

A. Implementation of SO in distributed wireless systems in general, and in femtocells networks in particular, can be realized by taking advantage of the literature proposed by machine learning and game theory communities. In this section we discuss two solutions based on these techniques and we assess them through performance results in macro-femto coexistence scenario.Learning based approached for SO.

Learning can be defined as the capability of drawing intelligent decisions by self-adapting to the dynamics of the environment, taking into account the experience gained in past and present system states, and using long term benefit estimations. Learning is adamantly driven by the amount of information available at every femtocell. Complete information about neighbors can significantly improve performances with respect to the case of partial observability, but the signaling burden over the backhaul may lead to the lack of scalability of the proposed scheme. As a result, a tradeoff should be achieved, keeping in mind that the capability of making autonomous decisions is a desirable property of a self-organized network.

A particularly interesting framework in realistic decentralized wireless networks is the literature of *reinforcement learning* (RL) [1]. The reason is that RL provides model free and online learning features, which makes it suitable for taking decisions in realistic wireless settings characterized by a high degree of dynamism due to e.g., lognormal shadowing, fast fading, mobility of users, multiuser scheduling, random femtocell nodes activity patterns, etc. RL schemes, such as classical Q-Learning, possess a firm foundation in the theory of Markov Decision Processed (MDPs) and can be shown to optimally perform in situations where only one decision maker is present in the scenario (i.e. single-agent learning). In a distributed scenario, such as the femtocell case, however, (1) the intelligent decisions are made by multiple intelligent and uncoordinated nodes; (2) the nodes partially observe the overall scenario; and (3) their inputs to the intelligent decision process are different from node to node since they come from spatially distributed sources of information. This distributed system can be mapped onto a *multi-agent system*, which consists of multiple nodes who are similarly and simultaneously adapting. This may generate oscillating behaviors that not always reach an equilibrium. The dynamics of learning may thus be long and complex, with complexity increasing with an increasing observation space.

A possible solution to speed up the learning process and to create rules for unseen situations, is to facilitate expert knowledge exchange among learners [6][7]. We introduced then in [8] an emerging framework for femtocells, referred to as *docition*, from "docere" = "to teach" in Latin, which relates to nodes teaching other nodes. This concept perfectly fits a femtocell network scenario, where a femtocell is active only when the users are at home, so that it can take advantage of the decision policies learnt by the neighbor femtocells, which have been active during a longer time. Depending on the degree of docition among nodes, the following cases can be distinguished:

1) Start-up Docition:

Docitive radios teach their policies to any newcomers joining the network. In this case, again, each node learns independently; however, when a new node joins the network, instead of learning from scratch how to act in the surrounding environment, it learns the policies already acquired by more expert neighbours. Gains are due to a high correlation in the environments of adjacent expert and newcomer nodes. Policies are shared by exchanging Q-tables.

2) IQ-Driven Docition

Docitive radios periodically share part of their policies with less expert nodes, based on the degree and reliability of their expert knowledge. Policies are shared by exchanging (a weighted version) of the entire Q-table or rows thereof, corresponding to states that have been previously visited.

3) Performance-Driven Docition:

Docitive radios share part or the entirety of their policies with less expert nodes, based on their ability to meet prior set performance targets. Example targets are maximum created interference, achieved capacity. erfect Docition: the multi-user system can be regarded as an intelligent system in which each joint action is represented as a single action. The optimal Q-values for the joint actions can be learned using standard centralized Qlearning. In order to apply this approach, a central controller should model the Markov decision process (MDP) and communicate to each node its individual actions.Alternatively, all nodes should model the complete MDP separately and select their individual actions; whilst no communication is needed here, they all have to observe the joint actions and individual rewards. Due to an exponential growth of the states, this approach is typically not feasible.

B. Game theoretic approaches for SO

When information exchange among femtocells is allowed, the strategic coexistence among femtocells can be modeled using tools from evolutionary game theory (EGT) [11][12] which was explored as a means of mitigating interference towards the macrocell tier. EGT was shown to provide relatively high gains as compared to classical learning algorithms (Q-learning among others), by relying on a HeNB-GW, which acts as a semi-centralized entity through which femtocells exchange information in a two-way communication fashion (see [10] for more details). Yet another game theoretic approach is fictitious play (FP) where femtocells have complete and perfect information, i.e., they know the structure of the game and observe at each time t the power allocation vector taken by all other femtocells Formally speaking, FP can be written in a strategic form $g^{(FP)} = \left(K, \{A_k\}_{k\in K}, \{u_k\}_{k\in K}\right)$. Here, K denotes the set of players (i.e., FBSs). For all $k \in K$ the set of actions of FBS k is the set of power allocation vectors $A_k = \left\{q_k^{(l,n)} : l \in \{0,\ldots,L_k\}, \quad n \in N\right\}$, where $\mathrm{L}_k = \{1,\ldots,L_k\}$ and $L_k \in N$ is the number of discrete power levels of FBS $k \in K$. The power allocation vector when FBS k transmits over sub-carrier n with power level l is given by:

$$q_k^{(l,n)} = \frac{l}{L} p_{k,\max} \tag{1}$$

Finally, $u_k(.)$ denotes the utility function of FBS k. Each FBS k assumes that all its counterparts play independent and stationary (time-invariant) mixed strategies π_j, $\forall j$, where $\pi_j = \left(\pi_{j,q_j^{(1,1)}},\ldots,\pi_{j,q_j^{(L,N)}}\right)$ and $\pi_{j,q_j^{(l,n)}} = \Pr\left(p_j(t) = q_j^{(l,n)}\right)$. Under these conditions, femtocell k is able to build an empirical probability distribution over each action set A_j, for all j. Let $f_{k,p_k(t)} = \frac{1}{t}\sum_{s=1}^{t} 1_{\{p_k(s)=q_k^{(l,n)}\}}$ be the empirical probability with which player j observe that player k plays action $q_k^{(l,n)} \in A_k$. Hence, $\forall k$, $\forall p_k \in A_k$, the following recursive expression holds:

$$f_{k,p_k}(t+1) = f_{k,p_k}(t) + \frac{1}{t+1}\left(1_{\{p_k(t)=q_k^{(l,n)}\}} - f_{k,p_k}(t)\right) \tag{2}$$

Let $\bar{f}_{k,p_{-k}}(t) = \prod_{j \neq k} f_{j,p_j(t)}$ be the probability with which player k observes the action profile $p_{-k} \in A_{-k}$ at time t, for all k. Let the $|A_{-k}|$ dimensional vector $f_k(t) = \left(\bar{f}_{k,p_{-k}}\right)_{\forall p_{-k} \in A_{-k}}$ be the empirical probability distribution over the set A_{-k} observed by player k. In what follows, the vector $f_k(t)$ represents the belief of player k over the strategies of all its corresponding counterparts. Hence, at each time t, and based on its own beliefs, $f_k(t)$, each FBS k chooses its action $p_k(t) = q_k^{(l,n)}$, i. which maximizes its expected utility function.

It can thus be implied that by playing FP, players become myopic, by building beliefs of strategies used by all other players, and at each time t, players choose the action that maximizes their instantaneous expected utility.

On the other hand, when information exchange among femtocells is no longer possible, different decentralized learning algorithms can be adopted by femtocells so as to mitigate their interference toward the macrocell tier. Among these learning algorithms is the classical Q-learning which was studied in greater details in [9]. In short, every FBS first carries out an exploration phase in which it learns by interacting with the environment in a trials-and-errors manner. After building its Q-table, and provided that the network does not dramatically change, each FBS pick the strategies that maximize the observed rewards over the interaction time of the players.

C. *Performance Evaluation Game theoretic and Learning Based approach*

In what follows, we compare performance results obtained through classical Q-learning and EGT, bearing in mind that contrarily to Q-Learning which is able to work in autonomous manner, or with limited feedback through the X2 interface, EGT requires a centralized controller to gather, process, and broadcast information about the agents. We evaluate performances in a macrocell scenario with radius $R_m = 500$, underlaid with K femtocells of radius $R_f = 20$, transmitting over N = 8 sub-carriers. We assume that femtocells have L = 3 transmit power levels. The minimum SINR of the macrocell UEs is set to 3 dB for each sub-carrier. The macro BS transmission power is 43 dBm, and the maximum femto one is 10 dBm. The considered path-loss model is 3GPP compliant. We also assume fast fading and log-normal shadowing with standard deviation of 8 and 4 dBm for outdoor and indoor communications, respectively. The discount factor and exploration probability are set to 0.95 and, 0.5 respectively.

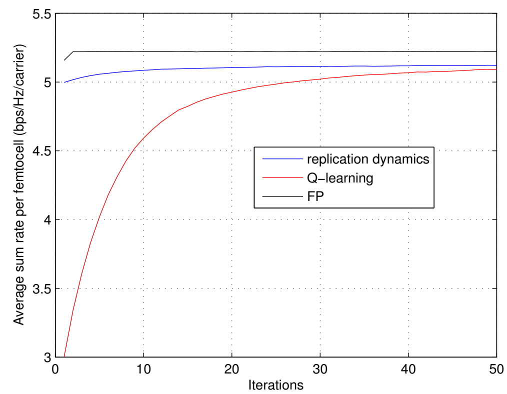

Figure 1: Convergence of the RL learning algorithms and their impact on the average femtocell sum-rate for K=50 femtocells and N=8 sub-carriers.

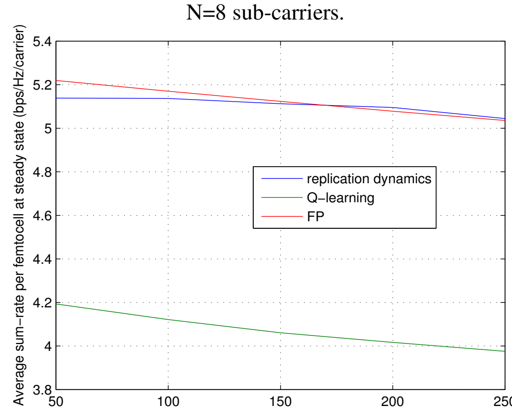

Figure 2: Effect of femtocell density on the average femtocell sum-rate, for different learning algorithms.

In Figure 1, we plot the average femtocell sum-rate for K=50 FBSs underlying one macrocell over N=8 sub-carriers, highlighting the convergence behavior for different learning algorithms. It can be noticed that the replicator dynamics, fictitious play, Q-learning schemes eventually converge to some steady state. Moreover, the fictitious play showcases the highest sum-rate, whereas the classical Q-learning needs an exploration phase until convergence is eventually reached. On the other hand, Figure 2 plots the impact of the femtocell density on the average femtocell sum-rate for the different learning algorithms, for K= 50, 100, 150, 200, 250 femtocells. A general decline in performance is perceived as the number of femtocells increases, which reflects the interference-limited nature of the network. Nonetheless, and quite interestingly, we see that the rate of decrease in performance is not the same for all algorithms. In particular, for K=250, the femtocell average sum-rate of around 5 bps/Hz is obtained using replicator dynamics and fictitious play, whereas approximately 4 bps/Hz is obtained with Q-learning. Convergence performances of Q-Learning can be improved by docition, which is shown in [8] to be able to improve the speed of convergence and precision of classical Q-Learning by up to 34%.

In addition, Figure 3 shows some performance results with respect to docition. Comparing the performance of 1) independent learning; 2) start-up docition; and 3) IQ driven docition, we observe a sharp improvement in precision of the docitive approach. In particular, it represents the complementary cumulative distribution function (CCDF) of the variance of the average SINR at the control point with respect to the set target of 20 SINR dB. It can be observed that due to the distribution of intelligence among interactive learners the paradigm of docition stabilizes the oscillations. At a target outage of 1 %, for instance, we observe that the IQ driven docition outperforms the start-up docition by a factor of two, and the independent learning algorithm by about an order of magnitude. This corroborates that docition facilitates a completely distributed and autonomous networking operation.

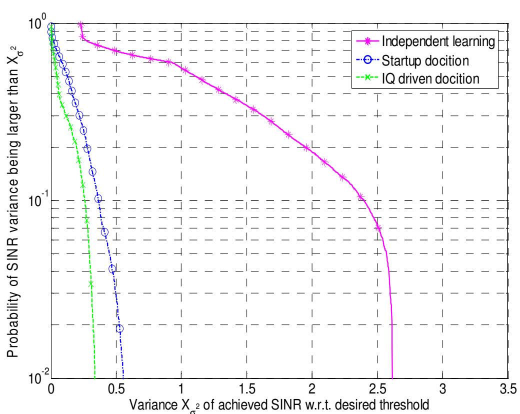

Figure 3: CCDF of the average SINR at macro user.

III. Direct Biomimmetic Approach for SO

In this section we present the use of direct biomimetic approaches to develop a novel SO framework for spectral efficiency enhancement on the access link between OFR and their donor Base Station (BS) through adaptation of system wide BS antenna tilts in a distributed manner.

The rest of this section is organized as follows. First we present the system model and problem formulation. Then a case study of SO system in nature is presented to describe how a solution to a similar problem is achieved in the natural system through SO. Same steps are applied to our problem by exploiting the one to one mapping between the two systems and SO solution is thus designed for our problem. Preliminary numerical results are also presented to demonstrate the gain proposed SO solution can achieve without relying on global coordination.

1) System Model and Problem formulation

We consider a sectorized multi cellular network with each base station having three sectors and each sector containing one OFR, as shown in **Figure 5**. Let N denote the set of points corresponding to the transmission antenna location of all sectors and K denote the set of location points (e.g. representing location of relay station (RS)) in the system. The geometric Signal to Interference Ratio (SIR) perceived at a location k being served by n^{th} sector can be given as:

$$\gamma_k^n = \frac{P^n G_k^n \alpha\left(d_k^n\right)^{-\beta}}{\sum_{\forall m \in \mathrm{N}\backslash n}\left(P^m G_k^m \alpha\left(d_k^m\right)^{-\beta}\right)} \qquad m,n \in \mathrm{N}, k \in \mathrm{K} \tag{4}$$

where P is transmission power, d is distance α and β are pathloss model coefficient and exponent respectively. G is the antenna gain and for 3GPP and LTE and LTE-A it can be modeled as in [13] i.e.

$$G_k^n = 10^{-1.2\left(\lambda_v\left(\frac{\theta_k^n-\theta_{tilt}^n}{B_v}\right)^2+\lambda_h\left(\frac{\phi_k^n-\phi_a^n}{B_h}\right)^2\right)} \tag{5}$$

where θ and ϕ are vertical and horizontal angles, from n^{th} sector to k^{th} location. Subscripts h, a and v denote horizontal, azimuth and vertical respectively. Subscript denotes the tilt angle of particular sector antenna. B represents beamwidth and λ is weighting factor to weight horizontal and vertical beam pattern of the antenna in 3D antenna model [13]. We assume that all the BSs transmit with the same power. This assumption is in line with LTE where no power control is applied to downlink. For such a scenario, by using (4) in (5), the SIR at location of k^{th} user can be determined as:

$$\gamma_k^n = \frac{h_k^n 10^{\mu\left(\left(\theta_k^n-\theta_{tilt}^n\right)^2+c_k^n\right)}}{\sum_{\forall m \in \mathrm{N}\backslash n}\left(h_k^m 10^{\mu\left(\left(\theta_k^m-\theta_{tilt}^m\right)^2+c_k^m\right)}\right)} \tag{6}$$

For the sake of simplicity, we use the following substitutions:

$$c_k^l = \frac{B_v^2 \lambda_h}{\lambda_v}\left(\frac{\phi_k^l - \phi_a^l}{B_h}\right)^2$$

$$h_k^m = \alpha\left(d_k^m\right)^{-\beta}; \qquad \mu = \frac{-1.2\lambda_v}{B_v^2}$$

Note that it can be seen that γ is a function of vector of tilt angles of all sectors i.e. θ_{tilt}^N where $N = |\mathsf{N}|$, but for the sake of simplicity of expression we will show this dependency only where necessary. Given the small sector size, we safely assume that a sector can have at most one randomly located OFR within it. Given the system model and assumption, our problem can be stated as: *optimizing system-wide antenna tilts to maximize the aggregate throughput* η *at access links of all the OFRs by minimizing the interference.* Mathematically:

$$\max_{\theta_{tilt}^N} \eta\left(\theta_{tilt}^N\right) = \max_{\theta_{tilt}^N} \sum_{\forall s \in \mathsf{S}} \log_2\left(1 + \gamma_s^n\left(\theta_{tilt}^N\right)\right) \tag{7}$$

where S is set of all points identifying locations of all OFR's.

2) Design of SO solution through direct biomimmetics

For exploiting a direct biomimetic approach we need to look for a system in nature that solves a similar problem in SO manner. A flock of common cranes is an excellent case study here. The problem it solves can be stated as "*optimise flock-wide flight attributes to maximise the group flight efficiency in the flock by minimising the air drag*." Replace *flock* with *system*, *flight attributes* with *antenna tilts, flight efficiency* with *spectral efficiency,* and *air drag* with *interference*; the two problems can be seen to be analogous providing opportunity for direct biomimetics to be exploited. Next step is to find out the underlying via flock of common cranes achieves a SO solution to its problem. For brevity, these steps are summarized in Figure 4.

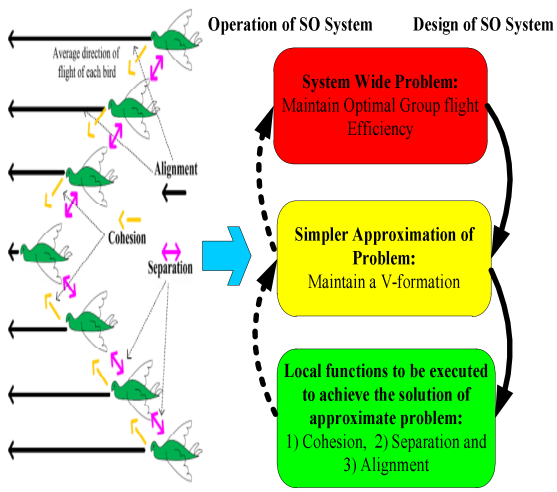

Figure 4: Design and operation of SO in flock of common cranes.

The three steps of designing self organisation (illustrated in Figure 4) when directly applied to our problem can be stated as 1) find a simpler *approximation of the actual problem* that can be then *2) decomposed* down into locally executable solution finally 3) determine the solution of the local sub-problem. In the following we follow these three steps to achieve a novel self organising solution for our problem in (7). In order to obtain simpler approximation we explore the analogy further and indentify that while controlling their flight attributes each bird only observes its two immediate neighbor birds and do not directly consider the behavior of rest of the birds in the flock. With this observation, the problem can be simplified as:

$$\max_{\theta_{tilt}^N} \hat{\eta}\left(\theta_{tilt}^N\right) = \max_{\theta_{tilt}^N} \sum_{\forall s \in \mathsf{S}} \log_2\left(1 + \hat{\gamma}_s^n\left(\theta_{tilt}^N\right)\right) \tag{8}$$

Where ^ shows the SIR is now dependent on the only neighboring two sectors. To accomplish the next step that requires decomposition of the system-wide problem into simpler local sub-problems, we propose the concept of *triplet* by exploiting the symmetry of our system model. The triplet is a fixed cluster of three adjacent and hence most interfering and immediate neighbor sectors as shown in Figure 5. Since the SIR is already a function of tilts of sectors within triplet only, the system wide problem in (8) can be now decomposed into local sub-problems for each triplet and can be written as:

$$\max_{\theta_{tilt}^{T_i}} \sum_{\forall s \in \mathsf{S}_i} \log_2\left(1 + \hat{\gamma}_s^n\left(\theta_{tilt}^{T_i}\right)\right) \tag{9}$$

This decomposition is similar to the decomposition of system-wide goal of maintaining V-formation into local tasks of cohesion, separation and alignment as shown in Figure 4. Now the problem in (9), is a very small scale optimization problem compared to that in (7), and therefore can be solved with standard techniques e.g. sequential quadratic programming.

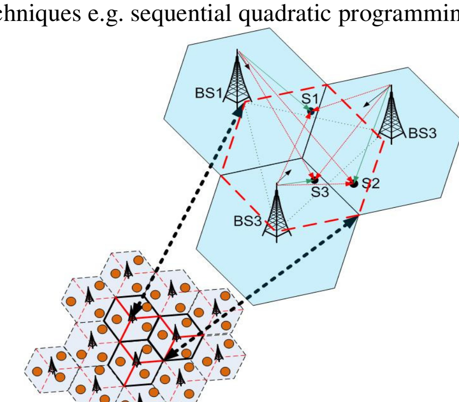

Figure 5: Circles indicate OFR's. Red dashed lines highlight a triplet.

The execution of this solution in each triplet independently, results in achievement of the system-wide simpler goal in (8) that in turn manifests the original objective in (9) approximately. This whole process is summarized in Figure 5. Here aiming for the *approximate* global objective is similar to the fact common cranes often do not achieve perfect V-shape but still by maintaining near V-shape they achieve substantial gain in spectral efficiency [13]. Similarly in next subsection we will show that the proposed solution can yield substantial gain while being self organizing at the same time.

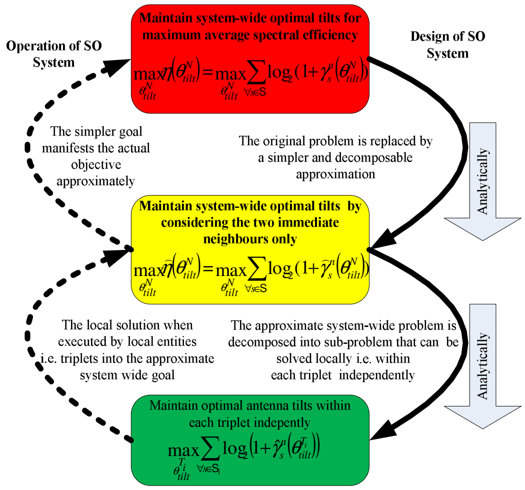

Figure 6: A use case of direct biomimetic approach

3) Numerical Results

In order to assess the potential gain of the proposed solution, numerical results for two different set of location of OFRs in a triplet are obtained, as shown in Figure 7.

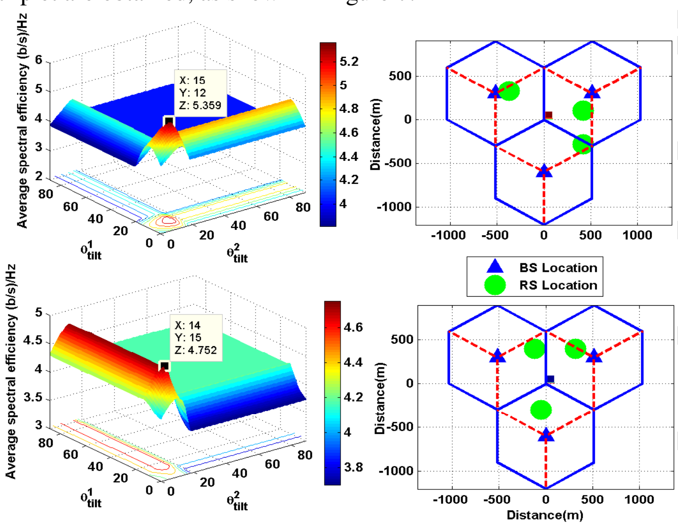

Figure 7: Average spectral efficiency per link within a triplet plotted as function of tilt angle of two sectors while third is fixed at 13^0.

It can be seen that depending on the location of OFRs, a gain in spectral efficiency from 1bps/Hz to 2bps/Hz can be achieved on average within each triplet if optimal tilt angle is obtained by the proposed solution. This can yield a system wide gain in spectral efficiency by SO of antenna tilts within each triplet in distributed manner without relying on global coordination or central control. The negligible amount of local signaling required among the immediate neighboring cells within a triplet can be done through X2 interface and it needs to be done only when location of OFR is changed.

IV. Conclusions

In this paper, use of directly and indirectly bio inspired techniques has been presented for enabling self-organization for macrocell and femtocell coexistence. Two key indirect bio mimetic techniques, namely learning and game theory, are explored and their performance is compared via simulation results. Another contribution of the paper is a novel demonstration of use of direct biomimetic approaches for developing SO for a system-wide antenna tilt non-linear and large scale optimization problem. The target scenario considers outdoor fixed relay based femtocells coexisting with a wireless backhaul to the macro system. Numerical results show that substantial gain in spectral efficiency on BS to relay station access links can be achieved applying the proposed solution that dynamically adapts antenna BS tilts in distributed and SO manner according to the relay station locations.

Acknowledgment

This work has been performed in the framework the ICT project ICT-4-248523 Broadband Evolved Femto networks (BeFEMTO), which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues from the BeFEMTO consortium.

References

[1] 3GPP,"Self-organising networks; concepts and requirements," 3GPP TS 32.500 v10.1.0, Tech. Rep., 2010.

[2] http://www.ict-befemto.eu/

[3] Osianoh Glenn Aliu, Ali Imran, Muhammad Ali Imran and Barry Evans. " A Survey of Self Organisation in Future Cellular Networks", Accepted in IEEE Journal of Survey and Tutorials on Communications.

[4] W. Elmenreich and H. de Meer, "Self-organizing networked systems for technical applications: A discussion on open issues," IWOS 2008, vol. LNCS 5343, pp. 1–9, 2008.

[5] M. E. Harmon and S. S. Harmon, "Reinforcement learning: A tutorial," 2000. [Online]. Available: http://www.nbu.bg/cogs/events/2000/Readings/Petrov/rltutorial.pdf.

[6] M. N. Ahmadabadi and M. Asadpour, "Expertness based cooperative Q-learning," IEEE Transactions on Systems, Man, and Cybernetics, Part B, vol. 32, no. 1, pp. 66–76, Feb. 2002.

[7] M. Tan, Multi-Agent Reinforcement Learning: Independent vs. Cooperative Agents. In M. N. Huhns and M. P. Singh, editors. Morgan Kaufmann, San Francisco, CA, USA., 1993, ch. 26, pp.

[8] L. Giupponi, A. Galindo, P. Blasco, M. Dohler, Docitive Networks - An Emerging Paradigm for Dynamic Spectrum Management, IEEE Wireless Communications Magazine, Vol. 17, No. 4, pp. 47-54, August 2010.

[9] EU FP7-ICT BeFEMTO project, ``Preliminary SON enabling & multi-cell RRM techniques for networked femtocells'', D4.1, Dec. 2010.

[10] M. Bennis, S. Guruacharya and D. Niyato, "Distributed Learning Strategies for Interference Mitigation in Femtocell Networks," IEEE GLOBECOM, Houston, USA 2011.

[11] Sandholm W. H, Population Games and Evolutionary Dynamics May 25, 2007, To be published by IT Press.

[12] D. Fudenberg and J. Tirole, "Game theory," MIT Press, 1991.

[13] Ali Imran, Muhammad Ali Imran and Rahim Tafazolli, "Relay Station Access Link Spectral Efficiency Optimization through SO of Macro BS Tilts", IEEE Communication Letters, Vol. 15, page 1-3, Nov-2011.