Distributed Interference Mitigation in Two-Tier Wireless Networks Using Correlated Equilibrium and Regret-Matching Learning

Pawel Sroka and Adrian Kliks
Chair of Wireless Communications
Poznan University of Technology
Poznan, Poland
{pawel.sroka, adrian.kliks}@put.poznan.pl

Abstract—**In this paper we study the interference management in two-tier cellular system from a game theoretic perspective. We extend the work given in [1], [2] to apply the game theoretic approach based on correlated equilibrium and regret-matching learning to multi-tier decentralized interference mitigation. The proposed approach requires periodic information exchange between coordinating base stations (BSs), thus several simplifications to the original algorithm are proposed. Numerical results of Monte Carlo simulations of Long Term Evolution - Advanced (LTE-A) like system are presented, with the proposed solution providing significant increase in terms of average cell throughput and user rate comparing to the state-of-the-art schemes.**

I. Introduction

Recently, a new multi-tier cellular network architecture has been considered for future communication systems, with small cells introduced to increase the system spectral efficiency in areas with dense user equipments (UEs) deployment [3]. These can bring enormous benefits to both the users and network operators, with better received signal quality due to the reduced distance between the transmitter and receiver. From the network operator point of view, small cells can enhance system capacity and increase the number of simultaneously connected UEs. However, one of the main challenges reducing the benefits from introducing small cells is the cross-tier interference, resulting from macro BSs interfering with pico/femto BSs transmission.

Several approaches to distributed interference mitigation have been proposed in the literature, with several of them employing game theoretic approach to obtain the optimal solution [1], [2], [4]. However, these are more focused on single-tier systems or cognitive radio approach, thus assuming ideal knowledge of channel gains, interferences and taken decisions in all nodes.

Interference management for heterogeneous networks has been studied also as a part of 3GPP Long Term Evolution (LTE) Release 10 and beyond study items. The current 3GPP proposal adopts time-domain muting mechanisms known as almost blank sub-frames (ABS) [3], [5]. The main idea of ABS in the current standard is that macro-cell to small cell interference can be effectively reduced through muted time frames.

The interference management schemes can be classified into centralized and decentralized. The centralized approaches make use of a coordinator for smart radio resource management. Decentralized schemes, on the other hand, are characterized by the absence of such coordinator and the interference mitigation decisions are made individually by the BS. They are further split into implicit and explicit coordination, where for implicit coordination no signaling is assumed between neighboring BSs. On the contrary, explicit coordination assumes that BSs exchange signaling information to coordinate their radio resource usage and interference management.

In this paper, we extend the work given in [1], [2] to apply the game theoretic approach to multi-tier interference mitigation using a decentralized algorithm. The basic idea is that each BS learns the regrets of possible actions and aims to minimize its average regret over time. In this work we assume that BSs exchange the information on channel gains and selected actions periodically, with the game solved offline and the resulting actions applied afterwards. Therefore we propose a long-term (with a time span of multiple transmission time intervals (TTIs)) interference mitigation scheme based on resource partitioning and power allocation. Moreover, taking into account the practical application of the proposed solution, we suggest several simplifications to reduce the number of required computations necessary to solve the game.

The rest of the paper is organized as follows. In Section II, the system and game models are presented with the details on optimization metrics given. Section III outlines the idea of regret-matching learning algorithm employed to iteratively converge to correlated equilibrium state, as well as the simplifications proposed to reduce the number of necessary computations. The following Section IV presents the assumptions and numerical results of comparison of the proposed approach with state-of-the-art solutions through simulations. Finally, Section V concludes the paper.

II. Models

A. System Model

Let us assume a system using orthogonal frequency division multiple access (OFDMA) with $M = 9$ macro-cell BSs and $K = 11$ pico-cell BSs, as shown in Fig.1, operating in the same frequency band. Consider partitioning of the bandwidth available to each of the BSs into blocks in time and frequency, with the BSs transmitting in each of the blocks using one of the selected power per sub-carrier levels, selected out of the set $P = \{p_{\text{low}}, p_{\text{high}}\}$. Let J denote the set of users distributed in the rectangular area shown in Fig. 1, with higher user density in the vicinity of pico BSs, with J_i denoting the set of users served by BS i. At each time interval, each BS serves at most

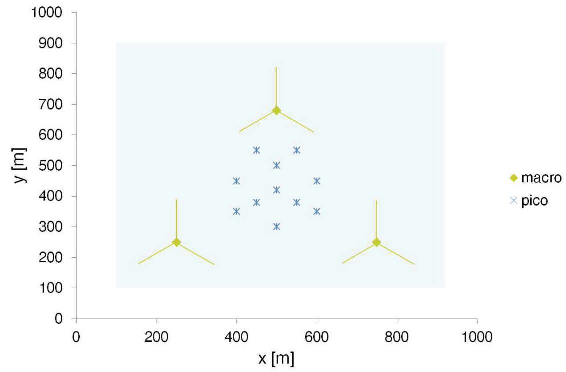

Fig. 1: BSs layout

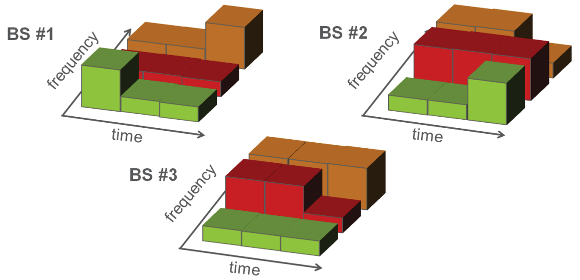

Fig. 2: Example of considered strategies

10 UEs, with the UEs scheduled according to the proportional fairness (PF) rule. Let $|h_{i,j}^{(s)}|^2$ denote the channel gain between the i-th BS and j-th UE on sub-carrier s ($h_{i,j}^{(s)} \in \mathbb{C}$), and σ_j^2 be the noise variance at receiver j. The signal to interference plus noise ratio (SINR) for UE j served by BS i on sub-carrier s is given as follows:

$$\gamma_{i,j}^{(s)} = \frac{|h_{i,j}^{(s)}|^2 p_i^{(s)}}{\sigma_j^2 + \sum_{l \in \{M \cup K\} \setminus i} |h_{l,j}^{(s)}|^2 p_l^{(s)}}, \tag{1}$$

where $p_i^{(s)}$ denotes the transmit power of BS i on sub-carrier s.

Let us assume that all BSs are interested in maximizing their rate R_i simultaneously minimizing the rate reduction (cost) of other BSs caused by interference, where the aggregate rate of BS i is given as follows:

$$R_i\left(\alpha_i, \boldsymbol{\alpha_{-i}}\right) = \sum_{j \in J_i} \sum_{s=1}^{S} \Psi\left(\gamma_{i,j}^{(s)}\right) \mathbf{1}_{j,s}, \tag{2}$$

where $\Psi\left(x\right)$ denotes the truncated Shannon formula [6] of x and $\mathbf{1}_{j,s}$ equals 1 when user j is scheduled over sub-carrier s and 0 otherwise. Each of the BSs selects its action α_i ($\alpha_i \in A$) that corresponds to selecting a set of power levels from P for each of the time-frequency blocks, as shown in Fig 2. Here A is a set of actions (strategies) that each of the BSs can play of the cardinality N. Let $\boldsymbol{\alpha_{-i}}$ be the set of actions selected by all other BSs than i. Thus the set of actions selected by all BSs can be formulated as $\boldsymbol{\alpha} = \{\alpha_i \cup \boldsymbol{\alpha_{-i}}\}$. The Vickrey-Clarke-Groves (VCG) [2] auction mechanism design can be introduced, where each of the BSs aims to maximize the utility U_i, $\forall i$, defined as:

$$U_i\left(\alpha_i, \boldsymbol{\alpha_{-i}}\right) \triangleq R_i\left(\alpha_i, \boldsymbol{\alpha_{-i}}\right) - \zeta_i\left(\alpha_i, \boldsymbol{\alpha_{-i}}\right), \tag{3}$$

where ζ_i denotes the cost (rate loss) introduced by BS i to all other BSs, which is evaluated as follows:

$$\zeta_i\left(\alpha_i, \boldsymbol{\alpha_{-i}}\right) = \sum_{l \neq i} R_l\left(\alpha_l, \boldsymbol{\alpha_{-i}}\right) - \sum_{l \neq i} R_l\left(\alpha_l, \boldsymbol{\alpha}\right). \tag{4}$$

B. Game Theoretic Model

The considered interference mitigation problem described above can be modeled using a normal-form game

$$\mathbb{G} = \left(\{M \cup K\}, A, \{U_i\}_{i \in \{M \cup K\}}\right),$$

where $\{M \cup K\}$ represents the set of players (BSs).
At each time t BS i selects its action from a finite set A following a probability distribution

$$\pi_i(t) = \left(\pi_i^{(\alpha_i^{(1)})}(t), \pi_i^{(\alpha_i^{(2)})}(t), \ldots, \pi_i^{(\alpha_i^{(N)})}(t)\right),$$

where $\pi_i^{(\alpha_i^{(n)})}(t)$ is the probability that BS i plays action $\alpha^{(n)}$. In general, each BS plays one of N strategies $\alpha^{(n)}, 1 \leq n \leq N$. Since the set A is discrete and finite, at least one equilibrium exists. The aim of each BS is to maximize its payoff U_i by cooperating with other BSs leading to the correlated equilibrium [7], formally defined as follows:

$$\sum_{\boldsymbol{\alpha}_{-i} \in A} \pi\left(\alpha_i^*, \boldsymbol{\alpha}_{-i}\right) \left(U_i(\alpha_i^*, \boldsymbol{\alpha}_{-i}) - U_i(\alpha_i', \boldsymbol{\alpha}_{-i})\right) \geq 0, \qquad \forall \alpha_i', \alpha_i^* \in A, \forall i \in \{M \cup K\}, \tag{5}$$

In (5) $\pi\left(\alpha_i^*, \boldsymbol{\alpha_{-i}}\right)$ is the probability of playing strategy α_i^* in a case when other BSs select their own strategies $\alpha_j, j \neq i$. The probability distribution π is a joint point mass function of the different combinations of BSs strategies. As in [1], the inequality in correlated equilibrium definition means that when the recommendation to BS i is to choose action α_i^*, then choosing any other action instead of α_i^* cannot result in higher expected payoff for this BS.

III. Regret-Matching Learning for Interference Mitigation

According to [8] if every BS plays with the rules of adaptive procedure of regret matching, the probability distribution π should converge to the set of correlated equilibrium as time goes to infinity. In [1], [2], the regret-matching learning algorithm is proposed to learn in a distributive fashion how to achieve the correlated equilibrium set in solving the VCG auction, which aims at minimizing the regret of selecting certain action. The regret $REG^{(T)}$ of BS i at time T for playing action $\alpha^{(n)}$ instead of other actions is given as follows:

$$REG_i^{(T)}\left(\alpha_i^{(n)}, \boldsymbol{\alpha_i^{(-n)}}\right) \triangleq \max\{D_i^{(T)}\left(\alpha_i^{(n)}, \boldsymbol{\alpha_i^{(-n)}}\right), 0\}, \tag{6}$$

where

$$D_i^{(T)}\left(\alpha_i^{(n)}, \boldsymbol{\alpha_i^{(-n)}}\right) =$$
$$= \max_{j \neq n} \frac{1}{T} \sum_{t \leq T} \left(U_i^t \left(\alpha_i^{(j)}, \boldsymbol{\alpha_{-i}}\right) - U_i^t \left(\alpha_i^{(n)}, \boldsymbol{\alpha_{-i}}\right)\right), \quad (7)$$

where $U_i^t\left(\alpha_i^{(\cdot)}, \boldsymbol{\alpha_{-i}}\right)$ is the utility at time t. $D_i^T\left(\alpha_i^{(n)}, \boldsymbol{\alpha_i^{(-n)}}\right)$ is the average payoff that BS i would have obtained if it had played other action than $\alpha_i^{(n)}$ every time in the past. Thus, if $D_i^T\left(\alpha_i^{(n)}, \boldsymbol{\alpha_i^{(-n)}}\right) > 0$ BS i would have obtained higher average payoff when playing different action than n and when $D_i^T\left(\alpha_i^{(n)}, \boldsymbol{\alpha_i^{(-n)}}\right) \leq 0$ BS i does not regret at all. Hence, given the regrets for all N actions, the probability of BS i selecting strategy n can be formulated as follows:

$$\pi_i^{(\alpha_i^{(n)})}(T) = 1 - \frac{1}{\mu^{(T-1)}} REG_i^{(T-1)}\left(\alpha_i^{(n)}, \boldsymbol{\alpha_i^{(-n)}}\right), \quad (8)$$

where

$$\mu^{(T-1)} = \frac{\sum_n REG_i^{(T-1)}\left(\alpha_i^{(n)}, \boldsymbol{\alpha_i^{(-n)}}\right)}{N-1}.$$

The regret-matching learning algorithm is an iterative approach that provides the means of convergence to correlated equilibrium in a distributed manner. However, an explicit knowledge of its utility function U_i (thus the costs introduced to other BS) and all other BSs actions is assumed, which are quite unrealistic in case of distributed optimization. Therefore, in this work it is assumed that the game is solved offline, with the results applied periodically on a long-term basis. Moreover, due to the long-term nature of optimization it is assumed that BSs exchange information on selected strategies and channel gains or interferences. However, when performing the calculation offline it can be computationally inefficient to calculate the sum of rates of all users, as given in (2) to obtain the costs of other BSs. Therefore two approaches are proposed:

- calculate the i-th BS rate R_i as an average wideband rate of users served by BS i (assuming all resources are allocated to considered user), which is given by:

$$R_i = \frac{1}{|J_i|} \sum_{j \in J_i} \sum_{s=1}^{S} \Psi\left(\gamma_{i,j}^{(s)}\right), \quad (9)$$

  with $|J_i|$ denoting the cardinality of the set J_i.
- calculate the cost of other BSs based on the rate of selected number of users experiencing highest interference from selected BS.

IV. Simulations

A. Assumptions

To compare the gains from proposed approach with state-of-the-art solutions Monte Carlo simulations of a LTE-A like system, shown in Fig. 1, with 9 macro BSs (forming 3 sites) and 11 pico BSs have been carried out. It was assumed that all BSs transmit using the same frequency and bandwidth, equal to 2.5 GHz and 10 MHz, respectively. Only downlink transmission has been considered with the interferences modeled for all possible links. The simulation parameter was the number of users distributed in the observation area, with the values selected from the set {100, 150, 200, 250, 300}. Users distribution was deterministic, with higher density of UEs near the pico BSs. The maximum transmission power for macro BSs was set to 46 dBm, whereas for pico BSs it was 24 dBm. The channel model is formulated using a combination of path-loss, log-normal shadowing and frequency-selective fading, with the path-loss and shadowing being inline with the ITU-T specification for IMT-Advanced [9]. It was assumed that all BSs and UEs equip only single antenna.

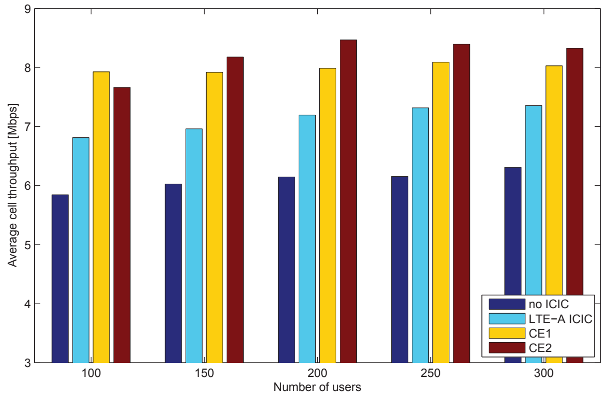

Fig. 3: Average BS rate vs. number of users

The results obtained for the proposed approach have been compared with results for system without interference mitigation and with system employing the LTE-A inter-cell interference coordination (ICIC) scheme based on ABS [5], with two ABSs used in each data frame (10 ms). When evaluating the performance of the proposed approach a long-term optimization has been assumed, with the game solved every 100 ms (100 TTIs). It was assumed that a total of 16 actions are available to each of the BSs, with the resources partitioned into six blocks (three in frequency and two in time) with different power level patterns. Two power levels have been considered with $p_{\text{low}} = p_{\text{high}} - 10\text{dB}$. Every BS decides on the employed action depending on the calculated probabilities distribution π_i. In order to reduce the computational burden of the game theoretic solution two algorithms have been considered;

- based on wideband user rate calculation - denoted as CE1,
- with cost calculated only for 10 users experiencing highest interference - denoted as CE2.

B. Numerical results

The considered solutions have been compared in terms of achieved cell rate, with also macro and pico BSs rates distinguished, and user throughput cumulative distribution for small cells UEs.

In Fig. 3 average BS rate is shown vs. the number of users in the observation area. One can notice the significant increase in achieved rate (about 30%) when comparing the results obtained for proposed algorithms with system with no

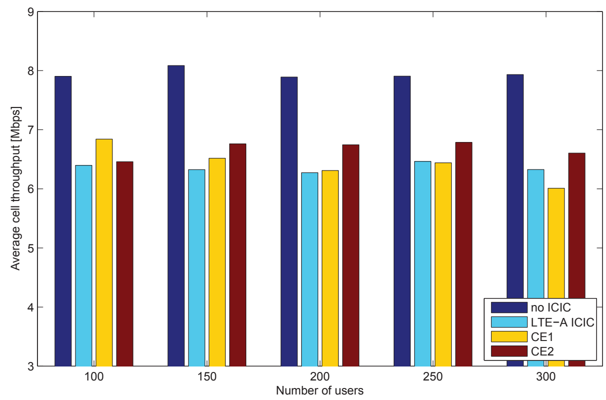

Fig. 4: Average macro BS rate vs. number of users

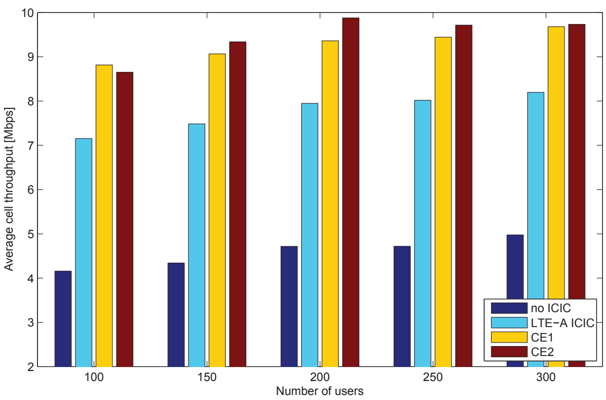

Fig. 5: Average pico BS rate vs. number of users

interference mitigation. Also when comparing the proposed approach with LTE-A ICIC scheme a throughput gain up to 1 Mbps can be observed. Additionally one can observe that when comparing the solution based on average wideband user rate calculation (CE1) and solution based on cost calculation for 10 UEs with highest interference (CE2) the former one performs better for lower number of users, while the latter outperforms other solutions for higher number of users. This is due to fact that when taking into account the rates of all users in cost calculation, the CE1 solution is more oriented on rate maximization than interference minimization, which combined with loss of accuracy caused by using wideband rate estimates may result in worse performance than for CE2.
Fig. 4 presents the average macro BS rate obtained with different solutions. One can notice that all interference mitigation schemes reduce the average throughput in macro-cells, with the achieved rate being similar for the proposed solutions and LTE-A ICIC. Therefore, the highest macro BS rate is achieved with no interference mitigation. One can also notice that the rate achieved with CE2 is slightly higher than in case of CE1.

When taking into account the throughput gains of interference mitigation schemes introduced to small cells, shown in Fig. 5, one can notice the huge rate increase with all interference mitigation schemes comparing to system with no ICIC. Again the proposed CE1 and CE2 solutions outperform also LTE-A ICIC, with CE2 achieving up to 2 Mbps gain over the state-of-the-art proposal.

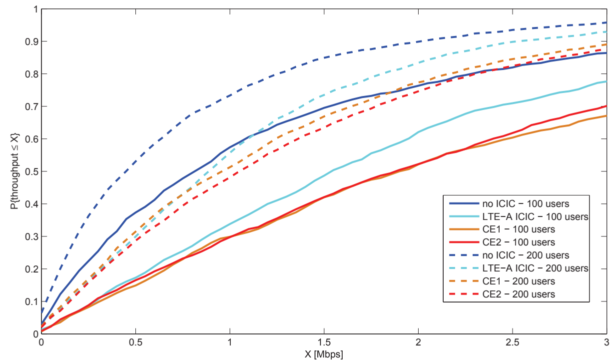

Fig. 6: Pico-cell user rate cdf

In Fig. 6 the pico-cell user rate cumulative distribution is presented for scenarios with 100 and 200 users. One can notice that for all considered solutions the lowest user rates are similar, meaning that there are some users in the system, whose transmission rate cannot be improved by means of interference mitigation. On the other hand, the probability of achieving higher data rate is higher with the proposed solutions, with CE2 being the best performer. One can also notice that the gain from proposed solutions comparing to LTE-A ICIC is mostly in the high rates region, meaning that more users achieve high data rate with the proposed solutions.
Concluding, the proposed approach based on correlated equilibrium and regret-matching learning outperforms the state-of-the-art solutions in terms of achieved average cell throughput and pico-cell user rate. Both versions of the proposed approach, employing different simplifications, provide similar results with the solution based on cost calculation for only limited number of UEs experiencing highest interference (CE2) performing slightly better.

V. Conclusion

In this paper the problem of cross-tier interference mitigation in heterogeneous cellular networks was studied from a game theoretic perspective. A decentralized approach based on regret-matching learning to achieve the correlated equilibrium and VCG auctioning has been proposed. The proposed solution requires exchange of information on channel gains and selected actions between the cooperating BSs, what is feasible under the assumption that the proposed interference mitigation scheme works on a long-term basis. Moreover, two simplifications have been proposed to reduce the number of necessary computations, which could be prohibitive otherwise when considering practical applications. Simulation results for the selected LTE-A like system showed that the proposed approach provides significant gains in terms of average cell throughput and small-cell users' rates comparing to the state-of-the-art solutions.

Acknowledgment

Part of this work has been performed in the framework of the FP7 project ICT-317669 METIS, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues in METIS, although the views expressed are those of the authors and do not necessarily represent the project.

References

[1] M. Charafeddine, Z. Han, A. Paulraj, and J. Cioffi, "Crystallized Rate Regions of he Interference Channel via Correlated Equilibrium with Interference as Noise "in *Proc. IEEE ICC 2009*, Dresden, Germany, June 14-18, 2009

[2] A. Kliks, P. Sroka, M. Debbah, "Crystallized Rate Regions for MIMO Transmission ", *Eurasip Journal on Wireless Communications and Networking*, 2010

[3] 3GPP TS 36.300, "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overal description", v. 11.3.0, September 2012

[4] H. Li,"Multi-agent Q-Learning of Channel Selection in Multi-user Cognitive Radio Systems: A Two by Two Case", in *Proc. IEEE Conference on SYStem, Man and Cybernetics (SMC)*, 2009

[5] LG Electronics, "3GPP TSG RAN WG1 Meeting #62 R1-104661 Comparison of Time-domain eICIC Solutions", August 2010

[6] ETSI TR 136 942, "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Frequency (RF) system scenarios (3GPP TR 36.942 version 10.2.0 Release 10)", v. 10.2.0, May 2011

[7] R. J. Aumann, "Subjectivity and correlation in randomized strategy", *Journal of Mathematical Economics*, vol. 1, no. 1, pp. 67-96, 1974

[8] S. Hart, A. Mas-Collel, "A simple adaptive procedure leading to correlated equilibrium", *Econometrica*, vol. 68, no. 5, pp. 1127-1150, September 2000

[9] ITU-R M.2135, "Guidelines for evaluation of radio interface technologies for IMT-Advanced", 2008