

A novel learning mechanism for traffic offloading with Small Cell as a Service

Panagiotis Trakas*, Ferran Adelantado*, Christos Verikoukis†

*Open University of Catalonia (UOC), Barcelona, Spain

†Centre Tecnologic de Telecomunicacions de Catalunya (CTTC), Barcelona, Spain

Email: ptrakas@uoc.edu, ferranadelantado@uoc.edu, cveri@cttc.es

Abstract—The densification of mobile networks with small cells is seen as the most promising solution to the explosive data traffic increase. Due to their financial implementation requirements, which could not be met by the service providers, the emergence of third parties that deploy and lease small cell networks opens up new business opportunities. In this paper, we study a proportionally fair auction scheme as an efficient way of small cell capacity distribution, both in network and financial terms. To improve the bidders' strategies, we propose a novel learning mechanism that alleviates the uncertainty incurred by variations in the traffic and the lack of information in the auctions. Extensive simulations prove the efficiency of our proposal, which also performs in equal terms with the ideal case of complete information.

Keywords—Traffic Offloading, SCaaS, Auction, LTE-A.

I. INTRODUCTION

The boost of the data traffic in cellular networks during the past few years forebodes future needs for additional capacity. Based on these predictions, both industry and academia have put the focus on solutions to increase the efficiency of the networks, thereby compensating such expected traffic growth. The densification of the Radio Access Network (RAN) with the deployment of small cells, and the resulting increase of the spectral efficiency, has emerged as one of the most promising solutions to address the problem.

Nevertheless, and despite its potential, network densification incurs important initial capital expenditure (CAPEX) as well as additional operational expenditure (OPEX) for mobile operators. These financial restrictions, that in principle could limit the actual deployment of dense networks, open up new business opportunities for third parties. In this described scenario the third parties, owners of small cells networks, will play the role of Small cell Service Providers (SSPs). Their infrastructure will be located at hotspots, either indoors or outdoors, where the Mobile Service Providers (MSPs) have increased capacity needs, and will create a heterogeneous network composite of several parties (i.e., MSPs and SSPs) to deliver pervasive broadband cellular access.

The aforementioned networks, despite coping with the current and future capacity needs, in turn pose new challenges. The arisen challenges are twofold: on the one hand the technological feasibility must be analyzed; on the other hand, an economic incentive must be generated to all the involved parties to assure the long-term sustainability.

Relevant studies have addressed this problem from a game theoretical approach, and most of them by modeling it with different auction schemes [1], [2].

An iterative double auction mechanism, which is managed by an independent broker, is proposed in [1]. The broker is

responsible for guaranteeing truthful bidding, thereby maximizing the market's efficiency while maintaining the profit. An auction-based incentive framework for leasing on-demand resources is presented in [2]. The MSP conducts reverse auctions with the third parties, generating the incentive for them to sell their capacity.

The traffic offloading through a third party, with WiFi or femtocell access points (APs), is considered in [3]. The authors use a two-stage multi-leader multi-follower game, called data offloading game (DOFF). In DOFF, the Base Stations (BSs) offer prices to the APs, which in turn decide the traffic volume to be offloaded by every proposing BS. An economic framework for traffic offloading to privately owned femtocells is described in [4]. The femtocells can be accessed by public users through a hybrid access mode and profit sharing is used to motivate their owners to offload traffic. A two-stage sequential game is modeled for revenue distribution, resource allocation and service selection. Similarly, in [5] the authors propose an Access Permission (ACP) transaction framework. This enables an MSP to buy ACP from multiple (geographically overlapping) SSPs. An adaptive strategy updating algorithm, which is based on an online learning process, is used for improving the SSPs strategies. The challenges of Mobility Management (MM) in Heterogeneous Networks along with various Handover (HO) decision algorithms are examined in [6]. Each MM scheme affects differently the traffic scheduled to the two tiers, and therefore the MSPs' requirements in traffic offloading.

In contrast to the SoA, this paper analyzes the traffic offloading problem under the Small Cell as a Service (SCaaS) approach [7]. In these scenarios, the third party, hereafter referred to as Small Cell Operator (SCO), is the owner of small cell infrastructure, composed of a cluster of co-channel operating small cells and the backhaul network to transport data from the small cells to the Internet or to the MSP core network. However, unlike the MSPs, the SCO does not have licensed spectrum and so it must be provided by the MSP. In particular, there is an economic transaction between the SCO and each MSP that determines the maximum MSP traffic that can be offloaded to the SCO, the spectrum sub-bands licensed to the MSP that can be used by the SCO, and finally the price that the MSP pays for the small cells infrastructure usage. The problem is modeled with an auction scheme in which the SCO is the auctioneer. In this context, and in order to adapt each actor's decisions to the traffic variations, we propose a novel learning mechanism that consists of a traffic forecasting method, a reinforcement learning algorithm and an adapting search range scheme for improving the MSPs strategies.

The rest of the paper is organized as follows; Section II describes the system model. The auction scheme, the involved parties' objectives and the learning mechanism are described

in Section III. Numerical results and analysis are presented in Section IV. Section V concludes the paper.

II. SYSTEM MODEL

The scenario under study is composed of N eNBs that belong to N different MSPs, namely MSP_i with $i = 1 \dots N$, and a single SCO cluster. In turn, the SCO cluster consists of a set of N_{sc} outdoor Home eNBs (HeNBs) deployed over a high traffic area (i.e., a hotspot) and a backhaul network with a maximum capacity C_{BH} .

Each MSP_i may be characterized by its spectral efficiency, SE_i , and its licensed bandwidth denoted by B_i . As for the offered load of the MSP_i (referred to as L_i), it may be divided into two components: the offered load generated within the hotspot (L_{hi}) and the offered load generated elsewhere (L_{ni}), i.e., $L_i = L_{hi} + L_{ni}$. Likewise, the SCO may also be characterized by the spectral efficiency, SE_{sc} , and the maximum bandwidth supported by the small cells according to the hardware limitations set by the respective technology, denoted by B_{sc} .

Load variations, both in time and space, are one of the key aspects of the stated problem. Thus, during time periods where the offered load in the hotspot (i.e., L_{hi}) is low, the need for the additional capacity provided by the SCO declines. Conversely, high hotspot loads result in a raising interest for the usage of the SCO infrastructure. In this sense, the day is divided into a set of T equal timeframes t_n , with $n = 1 \dots T$, during which the load can be regarded as a constant. It is nonetheless worth noting that, even though the load in a timeframe t_n is not necessarily the same every day, it is not independent from the load in the n th timeframe of the previous and subsequent days; in other words, the load tends to follow a day pattern.

At the beginning of each timeframe, an auction is started by the SCO. Each MSP_i bids an amount of money b_i to use the SCO infrastructure. After receiving the bids of all the MSPs, and based on them, the SCO distributes its capacity among the bidders. Thus, the MSP_i will be allowed to offload a maximum load through the SCO equal to $L_i^{sc} = N_{sc}x_iB_iSE_{sc}$, where $x_i \in [0, 1]$ is the part of the bandwidth provided by the MSP_i and used by the SCO. As defined before, the bandwidth allocated to the SCO cannot exceed B_{sc}

$$B_{sc} \geq \sum_{i=1}^N x_i B_i. \quad (1)$$

In turn, the offloaded traffic must be lower than the SCO backhaul capacity

$$C_{BH} \geq \sum_{i=1}^N L_i^{sc} \quad (2)$$

III. THE AUCTION

The distribution of the existing resources, i.e., the available spectrum and the cost of using the SCO infrastructure, are the response to two main objectives: firstly, the maximization of the throughput (the capacity objective), and secondly the optimization of the profits (the economic objectives of both the MSPs and the SCO). In this scenario the auction is the mechanism by which the interaction of these competing objectives results in incentives for all the involved parties.

A. The Auction Mechanism

In any auction there are three aspects that must be defined: the auctioneer, the bidders and the good to be auctioned. In the scenario under study, the capacity of the SCO is the auctioned good, the SCO plays the role of the auctioneer and the MSPs

are the bidders. Let us define the total load served by the SC cluster as $L_{sc} = \sum_i L_i^{sc}$ (in Mbps). If the distribution of the available SC capacity is based on proportional fairness [9], the maximum load that a given MSP_i can offload to the SCO is given by

$$L_i^{sc} = \frac{b_i}{b_i + \sum_{j \neq i} b_j} L_{sc}. \quad (3)$$

Accordingly, the maximum SC capacity gained by MSP_i after bidding b_i is subject to the bids of the rest of the MSPs and to the minimum profit demanded by the SCO. Given that neither the own load is known by MSP_i (since the auction is played at the beginning of the timeframe) nor the bids and the load of the rest of the MSPs are public (i.e., b_j and L_j , $\forall j \neq i$ are unknown), MSP_i must estimate L_{hi} , L_{ni} and $\sum_{j \neq i} b_j$ ¹ to properly select b_i . Hence, the more accurate the estimations are, the better the result of the auction.

In the following, a detailed analysis of the objectives of each party is presented to fully describe the auction. After this, a learning mechanism is proposed to enhance the bidding of the MSPs.

B. The SCO's objective

The SCO aims to maximize its profit by leasing high volumes of SC capacity at increased prices. The profit is defined as the difference of the MSPs' bids minus its expenses that are divided into the load and fixed costs, denoted as CL_{sc} and CF_{sc} respectively. Hence, the profit can be written as

$$P_{sc} = \sum_j b_j - CL_{sc} - CF_{sc}. \quad (4)$$

To ensure its objective, the SCO introduces a minimum profit, which is described as a percentage of its total costs

$$P_{sc}^{min} = (z - 1)(CL_{sc} + CF_{sc}), \quad z > 1. \quad (5)$$

If $P_{sc} < P_{sc}^{min}$, the SCO will not lease SC capacity to the MSPs. Regarding the load costs, they are given by convex functions that have been widely used in the literature [1]–[3], [10] for describing network congestion costs as well as the effect of subscriber churn. For low traffic, the decreased costs describe the already invested CAPEX for operating the network. Conversely, during peak traffic periods the cost increases rapidly, depicting the economic consequences of congestion. The load costs of the SCO are expressed as

$$CL_{sc} = \frac{a_{sc} L_{sc}^2}{C_{BH} + d_{sc} - L_{sc}}, \quad L_{sc} \in [0, C_{BH}]. \quad (6)$$

where factor a_{sc} defines the rate with which the cost increases (in €/Mbps), whereas d_i (in Mbps) moves the asymptotic discontinuity of the cost function to $L_{sc} = C_{BH} + d_{sc}$. Note that when the SCO network operates at its maximum capacity (C_{BH}), load costs are high but not infinite. Therefore, $d_{sc} > 0$.

Finally, using (3), (5) and (6) in (4), the reserve price for offloading L_i^{sc} traffic, denoted by b_i^{min} , can be written as

$$b_i^{min} = z \frac{a_{sc} L_i^{sc} L_{sc}}{C_{BH} + d_{sc} - L_{sc}} + z CF_{sc} \frac{L_i^{sc}}{L_{sc}}. \quad (7)$$

Note that the reserve price is a theoretical concept only valid when all MSPs bid in a complete information environment. In such a case, all MSPs are interested in bidding the reserve price to minimize the cost for offloading.

¹Let us henceforth denote as $b_{j \neq i}$ the sum of the opponents' bids.

C. The MSPs' objectives

The MSPs' objectives can be characterized by their financial and network aspects. Firstly, the MSPs aim to maximize their profit, which is defined as the revenue from serving traffic minus the expenses that are categorized into load (CL_i) and fixed (CF_i) costs, in addition to the bid b_i for leasing SC capacity. Therefore, the profit can be written as

$$P_i(x_i) = R_i(L_i^{mc} + L_i^{sc}) - CL_i - b_i - CF_i, \quad (8)$$

where L_i^{mc} denotes the load served by the macrocell and R_i the revenue per Mbps.

Secondly, they need to guarantee their subscribers' QoS, which can be accomplished by maximizing the throughput. When offloading occurs, the total load served by MSP_i , L_i^T is given by

$$L_i^T = L_i^{mc} + L_i^{sc} = (1 - x_i)B_iSE_i + N_{sc}x_iB_iSE_{sc}, \quad (9)$$

Since the throughput is a function of the transferred bandwidth, it is important to define the maximum and minimum limits of x_i (x_i^{max} and x_i^{min} respectively) for satisfying $L_i^T = L_i$. If $L_i > C_i$ ² is assumed, they can be expressed as follows

$$\begin{cases} x_i^{min} = \frac{L_i - B_iSE_i}{B_i(N_{sc}SE_{sc} - SE_i)}, \\ x_i^{max} = \frac{L_{h_i}}{N_{sc}B_iSE_{sc}}. \end{cases} \quad (10)$$

Similar to the SCO, MSP_i 's load cost is also described by a convex function of L_i^{mc}

$$CL_i = \frac{a_i(L_i^{mc})^2}{C_i + d_i - L_i^{mc}}, \quad L_i^{mc} \in [0, C_i]. \quad (11)$$

The values of a_i and d_i characterize the load cost function in (11). The costs can also be expressed in terms of Marginal Cost (MC) [3], [8], which is defined as the change in total cost for producing a unit of the offered good (L_i^{mc})

$$MC = \frac{dCL_i}{dL_i^{mc}} = \frac{a_iL_i^{mc}(2C_i + 2d_i - L_i^{mc})}{(C_i + d_i - L_i^{mc})^2} \left[\frac{\text{€}}{\text{Mbps}} \right]. \quad (12)$$

In order to highlight the dependencies of the characteristic cost parameters (i.e. a_i and d_i), the MC function is evaluated for $L_i^{mc} = C_i$,

$$MC(C_i) = \frac{a_iC_i(C_i + 2d_i)}{d_i^2} \Rightarrow a_i = \frac{MC(C_i)d_i^2}{C_i^2 + 2C_id_i}. \quad (13)$$

Applying the same analysis to the SCO, and assuming that both the MSPs and the SCO share the same congestion costs ($CL_{sc}(C_{BH}) = CL_i(C_i)$), a_{sc} can be written as

$$a_{sc} = \frac{d_{sc}CL_i(C_i)}{C_{BH}^2}. \quad (14)$$

Solving the system of (13) and (14) leads to the values of a_{sc} and d_{sc} .

D. Learning mechanism

The necessity of using a learning mechanism was pointed out in subsection III-A. However, in order to present the mechanism, it is essential to describe first the MSPs' decision making policy. This refers to the way the MSPs decide on the traffic to be offloaded to the SCO L_i^{sc} , as well as the corresponding bid at each auction. It was assumed that the MSPs have information about the SCO's distributing and charging

mechanism, that is, equations (3) and (7). By rearranging (7) with respect to L_{sc} , the following expression is obtained

$$za_{sc}L_{sc}^2 + (b_{j \neq i} - zCF_{sc})L_{sc} - (C_{BH} + d_{sc})(b_{j \neq i} - zCF_{sc}) = 0 \quad (15)$$

Solving the system of (3) and (15) results in MSP_i 's offloaded traffic L_i^{sc} . The solution of this new equation provides a function $b_i(L_i^{sc} | b_{j \neq i})$ that gives the winning bid for a demand L_i^{sc} , given the sum of the opponents' strategies $b_{j \neq i}$. This information would enable MSP_i to decide how much money it needs to pay for maximizing L_i^T and P_i . However, such information is not available. As a result, the MSPs need to choose their strategies based on estimates of $b_{j \neq i}$. Hence, MSP_i 's bidding strategy selection results from maximizing P_i given an estimate $\bar{b}_{j \neq i}$, while setting a constraint on the minimum throughput. This is formulated as

Maximize:

$$P_i(x_i | \bar{b}_{j \neq i}) = R_i(L_i^{mc} + L_i^{sc}) - CL_i - b_i(x_i | \bar{b}_{j \neq i}) - CF_i, \quad (16)$$

subject to:

$$x_i^{min} \leq x_i \leq x_i^{max} \quad (17)$$

Note that the limits imposed to x_i in (10) ensure the availability of resources, both in the hotspot and elsewhere, to serve the offered load of MSP_i . Therefore, the MSPs prioritize their throughput over their profits. The decision making could be described as MSP_i 's best response to $\bar{b}_{j \neq i}$.

We are now in a position to introduce the learning mechanism, which was designed in order to assist the MSPs' bid selection strategy. As stated above, it depends on MSP_i 's demand and its opponents' bids, indicating the importance of accurately estimating L_i and $b_{j \neq i}$ at each timeframe. Consequently, our proposed learning mechanism consists of three components. A forecasting method for predicting L_i , and a reinforcement learning algorithm along with an adaptive search range scheme for estimating $b_{j \neq i}$.

1) *Traffic Forecasting*: In order to forecast the MSPs' traffic loads L_i and L_{h_i} , the well-known and computationally efficient Holt-Winters (HW) method [11] was selected. Also known as Triple Exponential Smoothing, it takes into account the level, trend and seasonal changes in the observed dataset. There are two HW models according to the type of the seasonality³, known as multiplicative and additive seasonal models. The former refers to a proportional change in the values of the time series from season to season, whereas the latter refers to a particular absolute change. In our case, due to the behavior of the traffic pattern that is described by random, small, seasonal variations, the following equations for the multiplicative model were used

$$S_t = \alpha \frac{L_t}{I_{t-T}} + (1 - \alpha)(S_{t-1} + v_{t-1}), \quad (18)$$

$$v_t = \beta(S_t - S_{t-1}) + (1 - \beta)v_{t-1}, \quad (19)$$

$$I_t = \gamma \frac{L_t}{S_t} + (1 - \gamma)I_{t-T}, \quad (20)$$

$$\bar{L}_{t+m} = (S_t + mv_t)I_{t-T+m}, \quad (21)$$

where L_t denotes the observation of the offered load during timeframe t , and S_t a (smoothed) estimate of the level, that is, a local average of the dataset. Parameter v_t is an estimate of the linear trend (slope) of the time series, whereas I_t denotes

²Let us denote as $C_i = B_iSE_i$ the MSP_i 's macrocell sector capacity, when the former does not offload traffic at the SC cluster

³Seasonality can be described as the periodic repetition in time-series data

the seasonal component, in other words, the expectation for a specific timeframe based on its past season values. Parameter L_{t+m} denotes the forecast at m timeframes ahead, t the current timeframe, and T the number of daily timeframes (season length). In our scenario $m = 1$, since the estimate of the next timeframe was needed. The three smoothing factors $\alpha, \beta, \gamma \in [0, 1]$ show the dependence on the past values of the time series. Lastly, the method's parameters are initialized as

$$S_t = \sum_{t=1}^T \frac{L_t}{T}, v_t = 0, I_t = \frac{L_t}{S_t}, t = 1 \dots T. \quad (22)$$

For the initialization of the seasonal factors, a general rule suggests the use of at least two seasons of historical data.

2) *Learning Algorithm*: For a given continuous set of aggregate bid values $A_{b_{j \neq i}} = [b_{j \neq i}^{min}, b_{j \neq i}^{max}]$ that contains $b_{j \neq i}$, the learning algorithm should be able to estimate $b_{j \neq i}$. Nevertheless, using a continuous set could prove inefficient in our scenario, since the convergence requires thousands of iterations. Considering that, $A_{b_{j \neq i}}$ was discretized and the reinforcement learning algorithm Exp3 [12] was used. By doing so, the action selection was set up as a non-stochastic multi-armed bandit problem, where each $b_{j \neq i} \in A_{b_{j \neq i}}$ was regarded as one arm.

We propose Algorithm 1 to estimate the opponent's auction strategy, which is a modification of [12]. For each timeframe, an action set $A_{b_{j \neq i}}$ is defined and discretized to K values. When the algorithm is initialized in day $d = d_0$, each value is given an equal probability $p_t = 1/K$. Then, an action is drawn randomly based on the probability distribution, and the corresponding b_i is placed according to (16) and (17). A reward system is used for updating p_t . The rewards r_t are based on comparisons between the expected and the real results of the auction. As stated before, the MSPs know the SCO's distributing and charging mechanism. Hence, they can use their estimates L_i and $b_{j \neq i}$ to emulate the auction and its results. Since the actual bids are sealed, the gained SC capacity L_i^{sc} was chosen for the reward system. Particularly, the relative difference of the expected and actual value, ΔL_i^{sc} was mapped to a reward r_t in accordance with (24), denoting that for $\Delta L_i^{sc} = 0$ the reward is $r_t = 1$, whereas for $|\Delta L_i^{sc}| \geq 0.5$ it is $r_t = 0$.

Furthermore, another rule was included for improving already increased rewards. If the average ΔL_i^{sc} of the current and the last 4 days $\mu_{\Delta L}(d_{n-4} \dots d_n)$, as well as the standard deviation $\sigma_{\Delta L}(d_{n-4} \dots d_n)$ are below a certain value, r_t is given by (25). A small $\mu_{\Delta L}$ denotes good estimations, whereas a small $\sigma_{\Delta L}$ ascertains that these estimates do not reside within a large range of values, but they are close enough to $\mu_{\Delta L}$. The values in the condition were chosen after extensive simulations ($\mu_{\Delta L}(d_{n-4} \dots d_n) < 0.15$ AND $\sigma_{\Delta L}(d_{n-4} \dots d_n) < 0.2$). If it holds, the played action will receive a low reward, forcing the algorithm to explore neighboring values. Subsequently, the expected reward \hat{r}_t is calculated in (26), for updating all the probabilities with (27). Finally, $\eta \in (0, 1]$ denotes the learning speed; the higher its value, the faster the probabilities increase.

3) *Search Range Scheme*: In order for Algorithm 1 to work, $A_{b_{j \neq i}}$ must contain $b_{j \neq i}$. Hence, if the actual aggregate bid does not lie in the initial estimate of the action set, the algorithm will not be able to predict it. Then again, even if the initial estimate is correct, the algorithm will not be able to follow any changes of $b_{j \neq i}$ outside the predefined range. Therefore, a variable parameter space scheme was used [13]. The aim of this scheme is to center the probability distribution

Algorithm 1 Estimation of opponent's auction strategy

- 1: Initialization: $d = d_0$
 For one timeframe t
 Define $A_{b_{j \neq i}} = [b_{j \neq i}^{min}, b_{j \neq i}^{max}]$ and discretize it to K values $b_{j \neq i}^l, l = 1 \dots K$.
 Set probabilities: $p_t(d_0, b_{j \neq i}^l) = 1/K, \forall b_{j \neq i}^l \in A_{b_{j \neq i}}$
- 2: Draw $b_{j \neq i}^k$ randomly based on the probability distribution and place corresponding b_i according to (16) and (17).
- 3: Calculate relative difference

$$\Delta L_i^{sc} = (E[L_i^{sc}] - L_i^{sc}) / L_i^{sc} \quad (23)$$

- 4: Map ΔL_i^{sc} to a reward $r_t(d_n, b_{j \neq i}^k) \in [0, 1]$

$$r_t = -2|\Delta L_i^{sc}| + 1 \quad (24)$$

If $\mu_{\Delta L}(d_{n-4} \dots d_n) < 0.15$ AND $\sigma_{\Delta L}(d_{n-4} \dots d_n) < 0.2$

$$r_t = -|\frac{\Delta L_i^{sc}}{\mu_{\Delta L}}| + 1 \quad (25)$$

- 5: For $l = 1, 2, \dots, K$ calculate the expected reward

$$\hat{r}_t(d_n, b_{j \neq i}^l) = \begin{cases} r_t(d_n, b_{j \neq i}^k) / p_t(d_n, b_{j \neq i}^k) & \text{if } l = k \\ 0 & \text{otherwise,} \end{cases} \quad (26)$$

update and normalize probability distribution

$$p_t(d_{n+1}, b_{j \neq i}^l) = p_t(d_n, b_{j \neq i}^l) \exp(\eta \hat{r}_t(d_n, b_{j \neq i}^l) / K) \quad (27)$$

$$p_t(d_{n+1}, b_{j \neq i}^l) \leftarrow p_t(d_{n+1}, b_{j \neq i}^l) / \sum (p_t(d_{n+1}, b_{j \neq i}^l)) \quad (28)$$

obtained by the learning algorithm around a certain value by expanding and reducing its action set.

In this paper, it was utilized as an extension of the learning algorithm, so that $b_{j \neq i} \in A_{b_{j \neq i}}$. This scheme introduces parameters that determine the speed of the changes in $A_{b_{j \neq i}}$. These are the expansion and reduction rate, expressed as E_r and R_r respectively. Updating the action set rapidly or slowly might result in getting farther from the actual bid or not converging fast enough, respectively. The choice of expanding or reducing the range depends on conditions applied to the probability distribution. The parameters used in the conditions for updating $A_{b_{j \neq i}}$ are the minimum ($b_{j \neq i}^{min}$) and maximum ($b_{j \neq i}^{max}$) action values, as well as the 25th ($b_{j \neq i}^{25}$) and the 75th ($b_{j \neq i}^{75}$) percentile. The latter are compared with other percentiles that are defined by the expansion (E_c) and reduction (R_c) coefficients. In order for the range $A_{b_{j \neq i}}$ to expand or reduce, one of the following conditions [C1]-[C6] must be satisfied. The conditions for expansion [C1]-[C5] and reduction [C6] are defined as:

- [C1] $b_{j \neq i}^{25} < b_{j \neq i}^{25-E_c}$ AND $b_{j \neq i}^{75} > b_{j \neq i}^{75+E_c}$
- [C2] $b_{j \neq i}^{25} > (b_{j \neq i}^{max} + b_{j \neq i}^{min}) / 2$
- [C3] $b_{j \neq i}^{25} = b_{j \neq i}^{75}$
- [C4] $p_t(d, b_{j \neq i}) = 1/K \forall b_{j \neq i} \in A_{b_{j \neq i}}$
- [C5] $\mu_{\Delta L}(d_{n-4} \dots d_n) > 0.2$ AND $\Delta L_i^{sc}(d_n) < 0.15$
- [C6] $b_{j \neq i}^{25} > b_{j \neq i}^{25+R_c}$ AND $b_{j \neq i}^{75} < b_{j \neq i}^{75-R_c}$

Conditions [C1] and [C6] are from the initial scheme in [13]. The remaining conditions were introduced for specific cases that could not be handled by [C1] and [C6]. Condition [C2] is used when the distribution's mode is at $b_{j \neq i}^{max}$. The same

applies to [C3], but for the lower endpoint $b_{j \neq i}^{min}$. Condition [C4] is used when the algorithm is in a stalemate due to the actual bid being far from $A_{b_{j \neq i}}$, so that $r_t = 0$. Finally, [C5] is examined only when none of the rest is satisfied. It is applied when the mechanism is in a stalemate and does not place $A_{b_{j \neq i}}$ closer to the real bid. The new $A_{b_{j \neq i}}$ endpoints are given by:

$$\begin{cases} b_{j \neq i}^{min} \leftarrow b_{j \neq i}^{min} - E_r(b_{j \neq i}^{25} - b_{j \neq i}^{min}), & \text{for [C1]} \\ b_{j \neq i}^{max} \leftarrow b_{j \neq i}^{max} + E_r(b_{j \neq i}^{max} - b_{j \neq i}^{75}), & \text{for [C1]} \\ b_{j \neq i}^{min} \leftarrow b_{j \neq i}^{min} + E_r(b_{j \neq i}^{25} - b_{j \neq i}^{min}), & \text{for [C2]} \\ b_{j \neq i}^{max} \leftarrow b_{j \neq i}^{max} + E_r(b_{j \neq i}^{max} - b_{j \neq i}^{75}), & \text{for [C2]} \\ b_{j \neq i}^{min} \leftarrow b_{j \neq i}^{min} - E_r(b_{j \neq i}^{max} - b_{j \neq i}^{min}), & \text{for [C3]} \\ b_{j \neq i}^{max} \leftarrow b_{j \neq i}^{max} - E_r(b_{j \neq i}^{max} - b_{j \neq i}^{75}), & \text{for [C3]} \\ b_{j \neq i}^{min} \leftarrow b_{j \neq i}^{min} + E_r(b_{j \neq i}^{max} - b_{j \neq i}^{min}), & \text{for [C4]} \\ b_{j \neq i}^{max} \leftarrow b_{j \neq i}^{max} + E_r(b_{j \neq i}^{max} - b_{j \neq i}^{75}), & \text{for [C4]} \\ b_{j \neq i}^{min} \leftarrow 0.9b_{j \neq i}, & \text{for [C5]} \\ b_{j \neq i}^{max} \leftarrow 1.1b_{j \neq i}, & \text{for [C5]} \\ b_{j \neq i}^{min} \leftarrow b_{j \neq i}^{min} + R_r(b_{j \neq i}^{25} - b_{j \neq i}^{min}), & \text{for [C6]} \\ b_{j \neq i}^{max} \leftarrow b_{j \neq i}^{max} - R_r(b_{j \neq i}^{max} - b_{j \neq i}^{75}), & \text{for [C6]} \end{cases} \quad (29)$$

Note that, when $A_{b_{j \neq i}}$ is expanded or reduced, the range for which the probability distribution is defined vary. Therefore, when the range is reduced the probability distribution at the new points is set at the interpolated values, whereas for expansion these values are obtained through extrapolation.

Additional rules were introduced for changing dynamically the learning, expansion and reduction rates in order to improve the learning mechanism. The rationale is that the MSPs should use high rates when $\mu_{\Delta L}$ is increased, in order to detect faster the opponent's bid. On the other hand, when $\mu_{\Delta L}$ is below a certain value, they should reduce the rates for not deviating from a good prediction. The conditions are:

$$[C7] \quad \mu_{\Delta L}(d_{n-4} \dots d_n) > 0.2$$

$$[C8] \quad \mu_{\Delta L}(d_{n-4} \dots d_n) < 0.15 \text{ AND } \sigma_{\Delta L}(d_{n-4} \dots d_n) < 0.15$$

The new rate values are:

$$\begin{cases} \eta \leftarrow \eta + 0.1 \\ E_r \leftarrow 1.2E_r, & \text{for [C7]} \\ R_r \leftarrow 1.2R_r \end{cases} \quad \begin{cases} \eta \leftarrow \eta - 0.1 \\ E_r \leftarrow 0.5E_r, & \text{for [C8]} \\ R_r \leftarrow 0.9R_r \end{cases} \quad (30)$$

IV. PERFORMANCE EVALUATION

In this section, we evaluate our mechanism's performance by presenting the accuracy of its estimations, and its comparison with the ideal case of Complete Information (CI), where actual values are used instead of estimations. A scenario with two MSPs that share the same parameter values (given in Tables I and II) was implemented. Our work is focused on 3GPP's LTE-A networks, therefore, our values were chosen based on relevant studies [14]. Regarding the offered load, a traffic shape based on a bimodal distribution was used. Hence, a 24-hour pattern was devised with two peaks at t_{12} and t_{16} . The actual offered load of each MSP is generated from the pattern. Thus, in every timeframe, the offered load is a random variable centered at the load pattern value.

The mechanism's steady state (where MSPs converge) was compared with an ideal case that assumes CI for both MSPs, in other words, they know their opponent's network and financial parameters. This case was regarded as a multi-objective optimization problem, where the goal was the maximization of the profits (objective functions) of all the parties. Since no preferences were assumed, the Objective Sum Method was

TABLE I. MSP-SCO NETWORK-FINANCIAL PARAMETERS

MSP		
Bandwidth	B_i	20 MHz
Sector Spectral Efficiency	SE_i	1.7 bps/Hz
Transferred resources	x_i	[0,0.28]
Revenue per timeframe	R_i	3.375 €/Mbps
Marginal Cost	$MC(C_i)$	9.45 €/Mbps [8]
Cost shaping factor	a_i	0.495 €/Mbps
Cost shaping constant	d_i	10 Mbps
Fixed Costs per sector	CF_i	1.72 €/h
SCO		
Number of Small cells	N_{sc}	4
Max Bandwidth	B_{sc}	20 MHz
Spectral Efficiency	SE_{sc}	2.1 bps/Hz
Backhaul Capacity	C_{BH}	80 Mbps
Marginal Cost	$MC(C_{BH})$	9.45 €/Mbps [8]
Cost shaping factor	a_{sc}	0.06975 €/Mbps
Cost shaping constant	d_{sc}	7.5 Mbps
Fixed Costs	CF_{sc}	2.1 €/h
Profit factor	z	1.2

TABLE II. LEARNING MECHANISM PARAMETERS

Holt-Winters		
Constants	α	0.2
	β	0.1
	γ	0.1
Learning Algorithm		
Granularity of A_{b_j}	K	3
Learning Speed	η	0.3
Adaptive Search Range		
Percentage expansion coefficient	E_c	1
Percentage reduction coefficient	R_c	17
Expansion Rate	E_r	0.3
Reduction Rate	R_r	0.5

used [15], which maximizes the sum of the three profits, providing always a Pareto optimal solution. The now single-objective optimization problem was formulated as

$$\text{Maximize: } P(x_1, x_2) = P_1 + P_2 + P_{sc}$$

subject to: (2) and (17),

where the first constraint denotes the SC cluster capacity limitation, whereas the second one guarantees that $L_i^T = L_i$. It should be noted that since the bids are subtracted, function $P(x_1, x_2)$ only depends on the MSPs' revenue and the participating parties' costs. As a result, maximizing P is equivalent to minimizing the parties' costs, which can be achieved by serving L_i with the minimum use of transferred spectrum x_i .

Before presenting the results, it should be noted that the hotspot traffic L_{h_i} ranges between 20 – 60% of L_i . When offloading occurs, the macrocell's capacity C_i is reduced to L_i^{mc} , while the available SC capacity is L_i^{sc} , since a part x_i of MSP_i 's bandwidth is used at the SC cluster. During the low traffic timeframes ($L_i < C_i$), the MSPs offload their traffic to increase their profits. This happens because offloading small volumes of traffic is cheaper than serving it with the macrocell. Hence, our focus is mainly on the peak traffic hours (i.e., t_{11-17}), where $L_i > C_i$ and L_{h_i} ranges between 50 – 60% of L_i .

As it can be observed in Fig. 1, the mechanism is able to properly estimate b_2 , since the error is below 5% in most of the timeframes. The bad estimate in t_{18} ($\Delta b_2 \approx 10.5\%$) does not hinder the mechanism from performing well, since the bid and the estimation error are quite small ($b_i \approx 1.4\text{€}$, see Fig. 3) to affect significantly the profit and the auction result, respectively. The same applies to t_{10} .

Fig. 2 compares our proposal and the CI case with regard

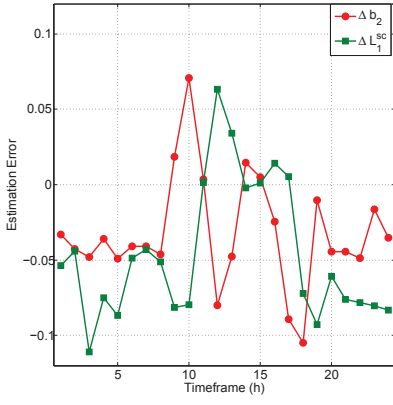


Fig. 1. MSP_1 Estimation Performance

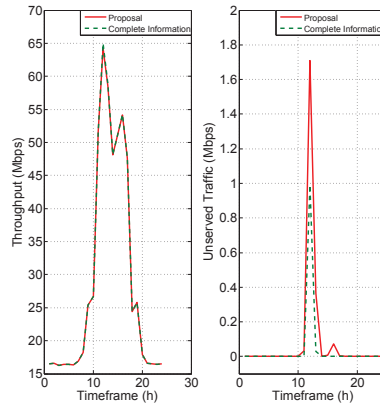


Fig. 2. MSP_1 Throughput and Unserved Traffic

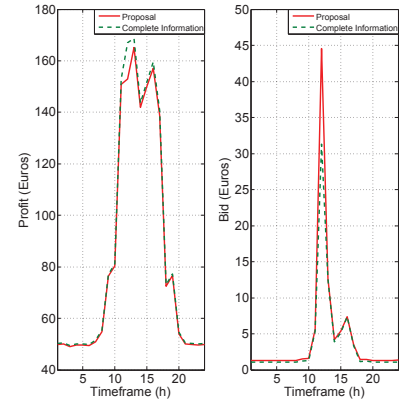


Fig. 3. MSP_1 Profit and Bid

to the throughput and unserved traffic, while in Fig. 3 they are compared in terms of profit and bid. The aforementioned good estimates provide a performance almost identical to the ideal case in comparison. The latter is marginally superior with 1.2% maximum difference for the throughput and 9.3% for the profit. Both of them are observed in t_{12} , where the offered load has its peak. At this timeframe, the MSPs' increased needs for offloading cannot be covered by the SC cluster. Therefore, a small volume of unserved traffic is observed. The small differences in the throughput are due to the estimation errors of b_2 that are seen in Fig 1. The corresponding b_1 is not sufficient for winning the auction, which results in offloading less traffic.

The difference in the profit can be explained by the difference observed in b_1 (Fig. 3). In the CI case the MSPs lease the necessary L_i^{sc} by bidding the reserve price b_i^{min} . On the other hand, our proposal is applied in a competitive environment, where due to the lack of information it has the tendency of increasing the MSPs' bids. This behavior is justified by the conditions used for expanding $A_{b_j \neq i}$ in the search range scheme, in particular [C1], [C2] and [C4]. This is rather intensified in t_{12} , where the MSPs' demands cannot be realized concurrently, leading to a daily increase of their bids. Therefore, the bids become higher than the ones in the CI case, reducing the profits.

V. CONCLUSIONS

In this paper we studied traffic offloading under the SCaaS approach, where a small cell operator owns the small cell infrastructure and the macro service providers transfer spectrum resources to serve their users. The problem of the efficient capacity distribution, both from a network and a financial perspective, has been modeled with a proportionally fair auction scheme. In this framework, it has been shown that the uncertainty about future traffic load poses the necessity to develop learning mechanisms to assist the auction. Modified versions of the HW method and the Exp3 algorithm have been proposed to deal with the load forecasting and the opponents' bid estimation, respectively. Results show that the proposed forecasting, learning and auction scheme copes with the uncertainty efficiently and provides results comparable with scenarios without uncertainty.

ACKNOWLEDGMENT

This work has been funded by the MITN Project CROSS-FIRE (PITN-GA-2012-317126)

REFERENCES

- [1] Iosifidis, G.; Lin Gao; Jianwei Huang; Tassiulas, L., "An iterative double auction for mobile data offloading," Modeling & Optimization in Mobile, Ad Hoc & Wireless Networks (WiOpt), 2013 11th International Symposium on , pp.154,161, 13-17 May 2013.
- [2] Dong, W.; Rallapalli, S.; Jana, R.; Qiu, L.; Ramakrishnan, K.K.; Razoumov, L.; Zhang, Y.; Cho, T.W., "iDEAL: Incentivized Dynamic Cellular Offloading via Auctions," Networking, IEEE/ACM Transactions on , vol.22, no.4, pp.1271,1284, Aug. 2014.
- [3] Lin Gao and Iosifidis, G. and Jianwei Huang and Tassiulas, L., "Economics of mobile data offloading", INFOCOM, 2013 Proceeding IEEE, pp.3303,3308, 14-19, April 2013.
- [4] Chien-Han Chai; Yuan-Yao Shih; Ai-Chun Pang, "A spectrum-sharing rewarding framework for co-channel hybrid access femtocell networks," INFOCOM, 2013 Proceedings IEEE , pp.565,569, 14-19 April 2013.
- [5] Yanjiao Chen; Jin Zhang; Qian Zhang, "Incentive mechanism for hybrid access in femtocell network with traffic uncertainty," Communications (ICC), 2013 IEEE International Conference on , pp.6333,6337, 9-13 June 2013.
- [6] Xenakis, D.; Passas, N.; Merakos, L.; Verikoukis, C., "Mobility Management for Femtocells in LTE-Advanced: Key Aspects and Survey of Handover Decision Algorithms," Communications Surveys & Tutorials, IEEE , vol.16, no.1, pp.64,91, First Quarter 2014
- [7] Real Wireless, "An assessment of the value of small cell services to operators", Report for Virgin Media, Oct 2012.
- [8] Joe-Wong, C.; Sen, S.; Ha, S., "Offering Supplementary Network Technologies: Adoption Behavior and Offloading Benefits," Networking, IEEE/ACM Transactions on , to be published.
- [9] R. T. Maheswaran and T. Basar, "Nash Equilibrium and Decentralized Negotiation in Auctioning Divisible Resources," in Group Decision and Negotiation, vol. 12, no. 5, pp. 361-395, September 2003.
- [10] Yangming Zhao; Sheng Wang; Sizhong Xu; Xiong Wang; Xiujiao Gao; Chunming Qiao, "Load balance vs energy efficiency in traffic engineering: A game Theoretical Perspective," INFOCOM, 2013 Proceedings IEEE , pp.530,534, 14-19 April 2013.
- [11] P.R. Winters, "Forecasting Sales by Exponentially Weighted Moving Averages," Management Science, vol. 6, pp. 324-342, 1960.
- [12] Peter Auer , Nicolò Cesa-Bianchi , Yoav Freund , Robert E. Schapire, "The Nonstochastic Multiarmed Bandit Problem," SIAM Journal on Computing, v.32 n.1, p.48-77, 2003.
- [13] M.N. Howell, T.J. Gordon, "Continuous action reinforcement learning automata and their application to adaptive digital filter design", Engineering Applications of Artificial Intelligence, Volume 14, Issue 5, October 2001, Pages 549-561, ISSN 0952-1976.
- [14] Real Wireless, "The Business Case for Urban Small Cells," Report for Small Cell Forum, Feb 2014.
- [15] Marler, R. T., and Arora, J. S. (2004), "Survey of Multi-Objective Optimization Methods for Engineering," Structural and Multidisciplinary Optimization, 26 (6), 369-395