

Improving spectrum efficiency in self-organized femtocells using learning automata and fractional frequency reuse

Maryam Nasr-Esfahani¹ · Behrouz Shahgholi Ghahfarokhi¹

Received: 4 September 2015 / Accepted: 17 July 2017 / Published online: 13 August 2017
© Institut Mines-Télécom and Springer-Verlag France SAS 2017

Abstract Deploying heterogeneous networks (HetNets) and especially femtocell technology improves indoor cell coverage and network capacity. However, since users install femtocells which usually reuse the same frequency band as macrocells, interference management is considered a main challenge. Recently, fractional frequency reuse (FFR) has been considered as a way to mitigate the interference in traditional as well as heterogeneous cellular networks. In conventional FFR methods, radio resources are allocated to macrocell/femtocell users only according to their region of presence ignoring the density of users in defined areas inside a cell. However, regarding the unpredictability of cellular traffic, especially on the femtocell level, smart methods are needed to allocate radio resources to the femtocells not only based on FFR rules, but also traffic load. In order to solve this problem, new distributed resource allocation methods are proposed which are based on learning automata (LA) and consider two levels of resource granularity (subband and mini-subband). Using the proposed methods, femto access points learn to choose appropriate subband and mini-subbands autonomously, regarding their resource requirements and the feedback of their users. The goal of the proposed methods is reduction of interference and improvement of spectral efficiency. Simulation results demonstrate higher spectral efficiency and lower outage probability compared to traditional methods in both fixed and dynamic network environments.

Keywords Resource allocation · Self-organized femtocell · FFR · Learning automata · Heterogeneous networks

1 Introduction

The usage of smartphones by home users is widespread [1], and the traditional cellular design is not capable of supporting high numbers of users through the expensive deployment of macrocells. Accordingly, low-power short-range femto access points (FAPs) have recently been introduced. Femtocell technology is a highly promising technique which can be used to tackle high amounts of traffic load, increase cell coverage and capacity, and improve the quality of home and edge users in a cost-effective manner [2]. In spite of femtocell benefits, femtocell deployment faces many challenges such as security, mobility and handoff management, auto-configuration, and especially interference management [3]. Since FAPs are installed and managed by users, their location as well as the pattern of their on/off switching is not known for the underlying cellular systems [4]. Moreover, as femtocells usually exploit the same frequency band as macrocells, interference mitigation techniques should consider not only co-tier interference (i.e., the interference occurring between neighboring femtocells) but also cross-tier interference (i.e., macro-to-femto and femto-to-macro interference) [5–8]. Therefore, a self-organized mechanism is required for FAPs to avoid such interferences [7].

Fractional frequency reuse is considered to be an effective method capable of reducing co-channel interference and increasing the network capacity for orthogonal frequency-division multiplexing (OFDM)-based HetNets [9, 10]. In most FFR schemes, the macrocell area is split into a number of subregions, and the frequency band is divided into subbands to allocate the different subbands to those subregions [5]. In

✉ Behrouz Shahgholi Ghahfarokhi
shahgholi@eng.ui.ac.ir

Maryam Nasr-Esfahani
m.nasresfahani@eng.ui.ac.ir

¹ Department of Information Technology Engineering, Faculty of Computer Engineering, University of Isfahan, Isfahan, Iran

general, FFR deployment schemes are categorized as static or dynamic. In static FFR schemes (such as strict FFR, soft FFR, FFR-3, and optimal static fractional frequency reuse (OSFFR) [6]), the assignment of subchannels to the user equipment (UEs) is performed statically with regard to the subregions in which the UEs are located [6]. For instance, in OSFFR, the macrocell coverage is divided into two zones: the center and edge zones, each including six sectors. From the other side, the entire frequency band is partitioned into seven parts: one is considered for the center zone and the other six are assigned to sectors of the border area, one part per each sector, as shown in Fig. 1. Accordingly, each macrocell UE (MUE) in the edge zone may experience intercell interference from one of the neighboring macrocells assuming the first tier. For example, in Fig. 1, the MUEs located in sector $\times 1$ of macrocell, one may experience interference only from macrocell 2. Also, since different subbands are used in the center and edge zones, intracell interference is avoided. In this scheme, the FAPs located in the center zone use only the subbands used by the MUEs in other sectors of the edge zone and not used in the adjacent border area of the current sector. However, the FAPs in the edge zone can exploit all subbands excluding the subbands assigned to the MUEs in that sector of the edge zone [6].

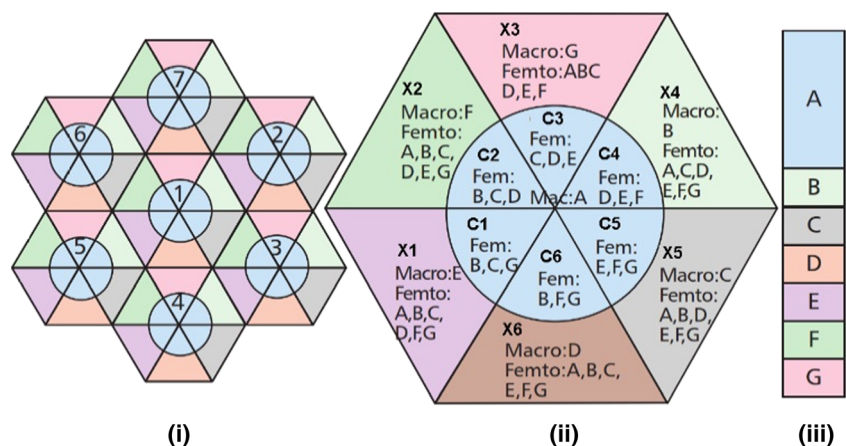
In contrast to static FFR methods which are based on hard reservation, dynamic FFR schemes can adaptively allocate radio resources to subregions and/or intelligently adjust the optimal radius of the center zone with regard to channel conditions and traffic load [11, 12]. However, these FFR methods usually select the best spectrum partitioning concerning the signal-to-interference-plus-noise ratio (SINR) level [4, 9, 10, 13]. Moreover, most of the contributions in the literature focus on centralized resource allocation techniques [7, 14]. Since centralized approaches run the risk of the single point of failure on the one hand and since femtocells are installed by the users themselves on the other, such techniques are not appropriate for femtocell networks.

The literature also includes distributed resource allocation methods (e.g., [15–17]) based on multi-agent learning techniques, where FAPs select appropriate subchannels with regard to local user feedbacks. With such schemes, due to a large number of non-organized subchannels, the search space is large and, consequently, performance is not appropriate with regard to the dynamic nature of the HetNet environment. Therefore, to cope with this problem and to better reduce interference and improve spectral efficiency as a result, in this paper, a combination of the FFR idea with the multi-agent learning concept is proposed to ensure self-organized resource allocation methods for femtocells. The proposed methods can choose adequate resources with respect to the requirements of users and the status of network.

In this paper, two self-organized fractional resource allocation methods are proposed where each FAP independently tries to learn the best resource allocation strategy using leaning automata (LA). The methods consist of two learning stages. In the first stage, each FAP tries to determine the best subband based on received feedbacks from its UEs and the macro base station (MBS) independently. However, as the size of the subbands may be large compared to femtocell needs, a more quantized level of resource partitioning, namely, mini-subband, is introduced and the second stage of the learning process is performed on the mini-subband level to further improve the utilization of the radio resources. The second stage only considers the feedback of FAP users (in terms of SINR level) and chooses one or more mini-subbands based on the aggregate demand of FAP users as regard the selected subband from stage 1. The second proposed method differs from the first in its second stage. In this method, distinct learners are exploited in the second stage according to the total resource demands of FAP users in order to improve convergence speed. It is noteworthy that the proposed methods give more priority to MUEs and no interaction is required between FAPs.

The rest of this paper is organized as follows: Related work is explained in the next section. Subsequently, the system model, problem formulation, and proposed methods are

Fig. 1 OSFFR deployment scheme for HetNets [6]



described in Sect. 3. Simulation results and discussions are presented in Sect. 4. Finally, the conclusion and future work are given in Sect. 5. Table 1 shows the abbreviations used in this paper and their definitions.

2 Related work

In OFDM systems, three different radio resource allocation techniques are available: power control, time-frequency resource allocation, and hybrid schemes. Besides, there are two different approaches for employing these techniques, i.e., centralized methods and distributed methods. However, considering the ad hoc nature of femtocell networks and regarding the fact that femtocells are installed and managed by users at unknown locations, it is essential to focus more on distributed mechanisms. Several studies aim at reducing interference and improving spectral efficiency using power control [1, 2, 4, 5]. However, the complexity of power control mechanisms, especially when the movement of the users is considered, is significant. On the other hand, numerous studies [3, 6–12] are dedicated to time-frequency resource allocation. The authors in [6] propose a distributed resource allocation technique based on game theory and Gibbs sampler to mitigate interference. Sudepta Mishra and Siva Ram Murthy [18] present a distributed power and physical resource block (PRB) assignment method for Long-Term Evolution (LTE) femtocell networks. The proposed method considers the location of the UEs on their interfering FAPs. The method includes an

iterative elimination algorithm for PRB allocation and an adaptive power control algorithm to find the required transmission power of each PRB. In [19], subchannel and transmit power allocation is formulated as an optimization problem in which the goal is to maximize the overall uplink throughput while guaranteeing minimum rate requirements of the femto user equipments (FUEs) and MUEs. The matching theory has been exploited to model the problem, and distributed algorithms are proposed on its basis. Moreover, several studies [7, 8, 14] suggest resource allocation algorithms based on reinforcement learning. In these approaches, each FAP is defined as an autonomous agent with local knowledge from the environment, assuming no communication between FAPs.

Distributed Q-learning mechanisms are proposed in [7, 8], in which femtocells choose the appropriate resource allocation policy based on local information, aiming to mitigate cross-tier interference in a self-organized manner. There, MBS periodically receives RSSI reports from MUEs and sends some control messages to the FAPs. Then, each FAP uses a Q-table to select the best action with respect to those environment feedbacks. This scheme is extended further in [14] using the combination of fuzzy inference system and multi-agent reinforcement learning.

Similarly, Bernardo et al. [10, 11] have proposed a distributed mechanism to maximize spectral efficiency in HetNets where FAPs are introduced as autonomous agents consisting of three different parts: RL dynamic spectrum assignment strategy (RL-DSA), cell characterization entity (CCE), and observer status. The RL-DSA algorithm plays the most important role, i.e., learning and selecting the best channel as its action. Moreover, CCE is a kind of system model, which estimates the feedback of the environment. Finally, observer status is responsible for triggering the RL-DSA algorithm and defining the cell state and updating allocation policy according to network conditions. Additionally, two distributed mechanisms have been proposed in [20] that exploit game theory and Q-learning. In the first scheme, each FAP adaptively chooses resources according to instant reward and the behavior of other agents while the second approach applies a trial-and-error mechanism to attain the best performance. Similarly, Semasinghe et al. [21] propose an evolutionary game theory-based distributed subcarrier and power allocation scheme for OFDMA-based small cells to learn from the environment and make individual decisions with minimum information exchange. The utility function of the small cells depends on average SINR level and data rate. Likewise, Xu et al. [22] formulates the resource allocation problem as a non-cooperative rate maximization game where the utility of each FAP is its capacity. Later, a utility-based learning model was proposed that requires no information exchange between FAPs. However, in the above learning methods introduced in the literature, the possible actions of the learners include

Table 1 Abbreviations used in this paper

Abbreviations	Definition
BS	Base station
CCE	Cell characterization entity
FAP	Femto access points
FFR	Fractional frequency reuse
FUE	Femto user equipment
HetNets	Heterogeneous networks
LA	Learning automata
MBS	Macro base station
MUE	Macrocell UE
OFDM	Orthogonal frequency-division multiplexing
OSFFR	Optimal static fractional frequency reuse
PDF	Probability density function
RL-DSA	RL dynamic spectrum assignment strategy
RSRP	Reference signal received power
RSSI	Received signal strength indication
SINR	Signal-to-interference-plus-noise ratio
UE	User equipment
PRB	Physical resource block

the set of subchannels/PRBs where the large size of the set affects the convergence speed. As a result, it seems only logical to exploit FFR resource allocation rules to improve convergence speed in self-organized learning-based methods.

FFR has recently been considered as an intercell interference mitigation technique designed to improve spectral efficiency. Researchers provide a wide comparison among four different static FFR schemes for downlink resource allocations in [15]. However, as the static resource allocation may not be preferable for dynamic conditions, some previous studies have focused on centralized dynamic resource allocation approaches based on FFR. For example, Oh et al. and Lee et al. [13, 23] employ a central controller to assign subbands to FAPs and MBS at the expense of higher signaling overhead and complexity. In [16], each FAP repeatedly transmits the preamble signal. So, each MUE provides a neighboring list of the FAPs to its MBS, which is finally used to create a table to determine available resources for each FAP. Lee et al. [17] propose a centralized mechanism to reduce interference between femtocells using FFR. They employ a centralized server which evaluates the interference between the femtocells and allocates orthogonal resources based on evaluation.

Part of the literature focuses on distributed resource assignment methods based on FFR. The authors of [3] propose a self-organizing spectrum allocation mechanism for femtocell networks based on the FFR technique. The method exploits random frequency hopping where each user selects a random subcarrier for transmission with respect to a given probability density function (PDF). Zhang et al. [9] propose a new interference management strategy to select the appropriate FFR scheme that improves the system throughput with respect to system constraints such as spectral efficiency and outage probability. For this purpose, at first, the spectrum is allocated to macrocells according to an FFR scheme. Then, each FAP sorts neighboring macrocells according to the reference signal received power (RSRP) and chooses the subband with the least RSRP. In [24], a distributed resource allocation mechanism is proposed for HetNets that not only improves user throughput but also reduces the interference on the uplink of the UEs. Whenever a FAP is plugged in, it calculates the distance from neighboring MBSs and the proper angle to all nearby MBSs to adjust its transmission power. Then, the FAP evaluates the set of possible transmission rates and selects the best subband that supports the target rate. Saadat et al. [25] present an algorithm for cognitive FAPs that guarantees a minimum QoS for FUEs regarding bandwidth while ensuring frequency reuse at the macrocell level. The solution considers a specific FFR scheme where part of the PRBs are allocated to the outer region of the macrocell (with a reuse factor of 1/3), and the remaining RBs are dynamically assigned to each sector in the inner region of the macrocell based on the demand of the MUEs.

As mentioned before, most FFR techniques restrict femtocells to the usage of specific frequency subbands while it is possible for intelligent FAPs to choose other subbands that are out of the frequency plan of a static FFR scheme, to improve the reuse and, consequently, spectrum efficiency. On the other hand, most of the previous works assume only one user with a constant request rate in each femtocell and mostly assume a network with static conditions. The primary purpose of this paper is to introduce more inventive self-organized resource allocation methods for femtocells to improve spectral efficiency in FFR-based HetNets. The proposed methods can also be rapidly adapted to the dynamic conditions of the network. In our proposed learning-based methods, the convergence speed has been studied by reducing the search/action space exploiting the FFR concept. Also, the proposed methods do not need any information exchange between the FAPs and only local femto and macro feedbacks are considered which diminishes the signaling overhead. A preliminary version of this research has already been presented in [26]. In that work, a learning automata is employed by each FAP to choose the appropriate subband based on previous feedbacks received from its local UEs. However, the study considers equal-size subbands which is lacking when FAPs with various user demands are considered. In the current research, we have extended our previous work regarding two levels of resource granularity in the learning process and also addressed macrocell users.

3 Proposed methods

In this section, the system model is presented followed by problem formulation; finally, the proposed methods are described in detail.

3.1 System model

A two-tier cellular network with n macrocells is considered where the MUEs, FAPs, and femto user equipments (FUEs) are randomly distributed in cells according to normal distributions. By evaluating the strength of the reference signals from the adjacent MBSs and FAPs, each UE connects to the base station (BS) with the highest signal strength. Here, it is supposed that the FAPs do not use power control. Therefore, FAP transmission power is fixed. Furthermore, it is assumed that the allocation of radio resources to the MUEs is based on OSFFR (shown in Fig. 1). In OSFFR, the number of subchannels allocated to the center zone (K_{center}) is obtained from Eq. (1) where K is the total number of available subchannels and r_{center} represents the radius of the center zone. r_{center} is assumed to be 0.65 times the radius of the macrocell, R as shown in Eq. (2) [15].

$$K_{center} = \left\lceil K \left(\frac{r_{center}}{R} \right)^2 \right\rceil \tag{1}$$

$$r_{center} = 0.65^* R \tag{2}$$

In Eq. (3), the number of subchannels allocated to the edge zone, K_{edge} , is calculated where N is the reuse factor of the edge zone in the mentioned FFR scheme, i.e., 6 for OSFFR.

$$K_{edge} = \frac{(K - K_{center})}{N} \tag{3}$$

In the proposed method, the subchannels are categorized into certain subbands. Also, as shown in Fig. 2, each subband is divided into a number of mini-subbands. The number of mini-subbands (N_{Msb}) in each subband is a function of subband size (S_{Sb}) and mini-subband size (S_{Msb}) as shown in Eq. (4). The sizes represent the number of subchannels in the subband and the mini-subband, respectively. It is assumed that all users of each FAP can be served by at most one subband. However, as the number of FUEs associated with each FAP and their requested bandwidth are changeable, the number of mini-subbands allocated to each FAP is variable. It is assumed that the round-robin scheduling algorithm is used to assign the allocated subchannels to the FAP FUEs.

$$N_{Msb} = S_{Sb} / S_{Msb} \tag{4}$$

3.2 Problem formulation

In the proposed methods, the goal is to improve spectral efficiency and consequently aggregate network capacity (C_{total}) given by Eq. (6). In this equation, $\tau_{x_m,m}^k$ and $\tau_{y_f,f}^k$ show the

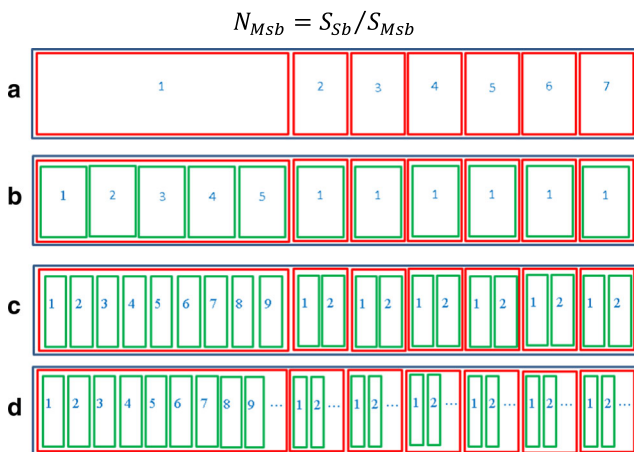


Fig. 2 **a** Subbands. **b** Mini-subbands in each subband when the size of mini-subbands is the same as subbands 2–7. **c** Mini-subbands in each subband when the size of mini-subbands is half of the size of subbands 2–7. **d** Mini-subbands with different sizes

allocation of subchannel k to MUE x_m (that is associated with macrocell m) and FUE y_f (that is related to femtocell f), as shown below:

$$\tau_{\alpha_i,i}^k = \begin{cases} 1 & \text{subchannel } k \text{ is allocated to UE } \alpha_i \text{ which is associated with BS } i \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Similarly, $C_{x_m,m}^k$ and $C_{y_f,f}^k$ represent the maximum achievable data rate for MUE x_m and FUE y_f on subchannel k . In Eq. (6), we use the aggregation of the log-scaled data rates to achieve proportional fairness [27] in addition to spectrum efficiency. In (6), M represents the set of all macrocells and X_m denotes the set of users associated with macrocell m . Likewise, F_A shows the set of all femtocells and Y_f means the set of users associated with femtocell f . It is considerable that the values of both $C_{y_f,f}^k$ and $C_{x_m,m}^k$ are functions of vectors SB and MSB where SB and MSB denote the subband and the mini-subband allocation matrices of the femtocells.

$$C_{total} = \sum_{m \in M} \sum_{x_m \in X_m} \sum_{k \in K} \left[\tau_{x_m,m}^k \times \left[\log_{10} C_{x_m,m}^k(SB, MSB) \right] \right] + \sum_{f \in F_A} \sum_{y_f \in Y_f} \sum_{k \in K} \left[\tau_{y_f,f}^k \times \left[\log_{10} C_{y_f,f}^k(SB, MSB) \right] \right] \tag{6}$$

Equation (7) shows the definition of SB and MSB.

$$SB_{i,j} = \begin{cases} 1 & \text{subband } j \text{ is allocated to FAP } i \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

$$MSB_{i,j} = \begin{cases} 1 & \text{mini-subband } j \text{ is allocated to FAP } i \\ 0 & \text{otherwise} \end{cases}$$

It should be noted that the MSB columns, the relevant subband of which is assigned to the FAP, are considered in Eq. (7) and the values of the other columns are zero.

Therefore, the values of $C_{x_m,m}^k$ and $C_{y_f,f}^k$ are calculated using Eqs. (8) and (9) where the SINR levels of the downlink transmissions to MUE x_m and FUE y_f on subchannel k are denoted as $SINR_{x_m,m}^k$ and $SINR_{y_f,f}^k$, respectively. In (8) and (9), the subchannel bandwidth is shown by ΔB . It is clear that $SINR_{x_m,m}^k$ and $SINR_{y_f,f}^k$ depend on subband and mini-subband allocations and are, therefore, functions of matrices SB and MSB .

$$C_{x_m,m}^k = \Delta B \cdot \log_2 \left(1 + SINR_{x_m,m}^k \right) \tag{8}$$

$$C_{y_f,f}^k = \Delta B \cdot \log_2 \left(1 + SINR_{y_f,f}^k \right) \tag{9}$$

Based on the above definitions, the resource allocation problem is formulated as Eq. (10) where RR_{y_f} represents the requested rate of FUE y_f .

$$s.t. \sum_{y_f \in F} \log \left(\sum_{k \in K} \tau_{y_f, f}^k C_{y_f, f}^k (SB, MSB) \right) \geq \sum_{r \in R} \log (RR_{y_f, f}) \quad \forall f \in F \tag{10}$$

3.3 Learning automata

As mentioned before, FAPs are installed and managed by users. Therefore, it is necessary to have self-organized mechanisms to allocate radio resources to FAPs with the aim of improving spectral efficiency while providing user demand. In this paper, to solve the optimization problem of Eq. (10) in a distributed manner, the stochastic learning automata method is used by each FAP. In this case, the femtocell network is considered as a learning automata team. The learning automata team has already been used to solve similar problems such as relaxation labeling and graph coloring [28].

In general, the stochastic learning automata (Fig. 3) is defined as $LA = \{\alpha \cdot \beta \cdot p\}$ where $\alpha = \{\alpha_1 \cdot \alpha_2 \cdot \dots \cdot \alpha_r\}$ is the set of actions (here, subbands as well as the possible combinations of mini-subbands) and r is the number of actions. Also, $\beta = \{\beta_1 \cdot \beta_2 \cdot \dots \cdot \beta_r\}$ is the set of environment feedbacks for each action which is a set of continuous values in the range [0,1] according to the S-model environment [29] (here, the quality feedbacks from FAP users and MBS as denoted by $\beta_{femtoSB}$ and $\beta_{macroSB}$). Finally, the action selection probability vector is represented by $p = \{p_1 \cdot p_2 \cdot \dots \cdot p_r\}$ [30] (in the first stage of

the proposed methods, the action probability vector is denoted by P_{SB} while in the second stage, it is denoted by P_{MSB}).

The action selection probability vector is updated according to environment feedbacks from chosen actions during the training of LA. Linear update methods are used in this research. Linear methods are categorized as linear reward-inaction (L_{R-I}), linear reward-penalty (L_{R-P}), and linear reward- ϵ -penalty ($L_{R-\epsilon P}$). As shown in Eqs. (11) and (12), the algorithms update the values of the probability vector based on received feedback, $\beta_i(n)$ after performing action i . The reward and penalty coefficients are denoted as constants a and b , respectively, and n represents the time step. The type of the learning algorithm is related to the values of these coefficients. When $b = 0$, i.e., the penalty coefficient is zero, the algorithm is named the L_{R-I} scheme, while if the values of the coefficients are the same ($a = b$), the update algorithm is called L_{R-P} . When the value of the penalty coefficient is too small compared to the reward coefficient ($b \ll a$), the algorithm is represented by $L_{R-\epsilon P}$ [29]. In L_{R-I} , since it is possible for one of the elements of the action probability vector to converge to one (consequently, all others converge to zero), it may not be suitable for dynamic conditions. However, the L_{R-P} and $L_{R-\epsilon P}$ update algorithms converge in distribution and, thus, are better for dynamic conditions. Here, the L_{R-P} scheme is used by the proposed methods [28].

$$p_i(n+1) = p_i(n) + a(1-\beta_i(n)) \cdot (1-p_i(n)) - b \cdot \beta_i(n) \cdot p_i(n) \tag{11}$$

$$p_j(n+1) = p_j(n) - a(1-\beta_i(n)) \cdot p_j(n) + b \cdot \beta_i(n) \cdot \left[\frac{1}{r-1} - p_j(n) \right]; \forall j \neq i \tag{12}$$

3.4 Two-stage LA-based channel allocation

The first proposed method, namely two-stage LA-based channel allocation (TSLACA), is a distributed resource allocation method that consists of two stages where one LA structure is used in each stage to train the subband and mini-subband selection probabilities respectively, as shown in Fig. 4. In the first stage, each FAP uses an LA to select the best subband according to the feedbacks received from MBS ($\beta_{macroSB}$) and its FUEs ($\beta_{femtoSB}$). For this purpose, in each time step, each MBS calculates its user feedback ($\beta_{macroSB}$) per each subband and sends the feedback vector (one entry per subband) to the underlying FAPs. Then, each FAP merges $\beta_{femtoSB}$ and $\beta_{macroSB}$ to attain an aggregate feedback (called β_{SB}) for each subband with regard to its region of presence (in the OSFFR scheme) according to Eq. (13).

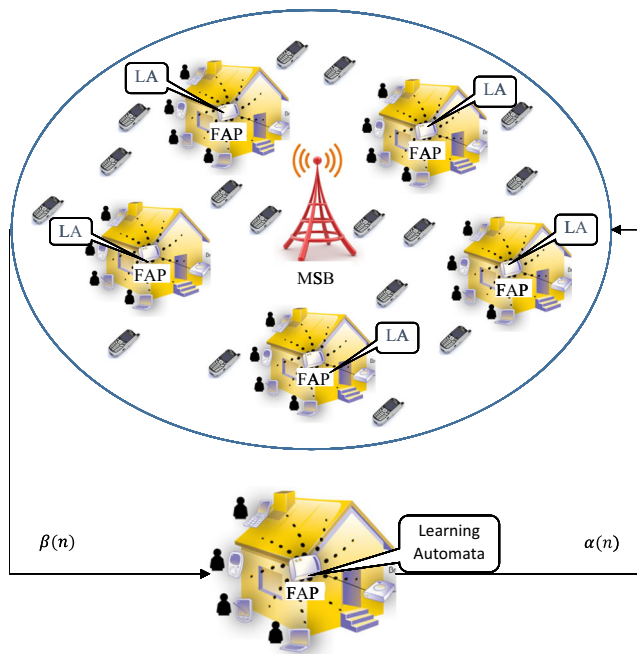


Fig. 3 Stochastic LA-based channel allocation team

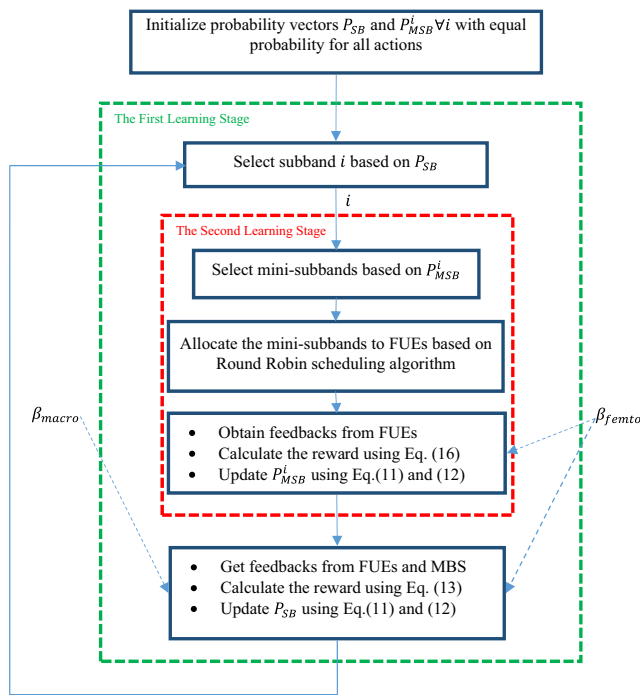


Fig. 4 The flowchart of two-stage LA-based channel allocation (TSLACA)

$$\beta_{SB} = \begin{cases} \beta_{femtoSB} & \text{if } subband_{FAP} \notin \{subband_{BS}^{reg}\} \\ Coef_f \times \beta_{femtoSB} + Coef_m \times \beta_{macroSB} & \text{otherwise} \end{cases} \quad (13)$$

In this equation, if the FAP’s selected subband is different from the subbands that the MUEs can use in that region (i.e., $\{subband_{BS}^{reg}\}$), $\beta_{macroSB}$ does not affect the overall feedback. Otherwise, the weighted sum of the feedbacks is used as the overall feedback. To give more priority to the MUEs, $Coef_m$ is given a higher value than $Coef_f$. β_{SB} is used to update the subband probability vector as in Eqs. (11) and (12). In fact, the main purpose of the first stage is not only to improve the spectral efficiency of the FUEs but also select the appropriate subband with respect to MUE satisfaction feedback.

Calculating FUE feedback, each FAP evaluates the total capacity attained by its users (C_{x_m}) and calculates $\beta_{femtoSB}$ on its basis. For this purpose, the log-scaled aggregate capacity that is achieved by the FUEs ($AchievedR$) is compared to their aggregate log-scaled requested rates ($TotalRR$) as demonstrated in Eqs. (14, 15, 16). In these equations, D_{y_f} and c_{y_f} are the requested and allocated rates for FUE y_f , respectively. It is noteworthy that in calculating feedback, $\beta_{femtoSB} = 0$ shows the most favorable response, while the closer $\beta_{femtoSB}$ to 1, the more unfavorable the response.

$$TotalRR = \sum_{y_f \in f} \log(D_{y_f}) \quad (14)$$

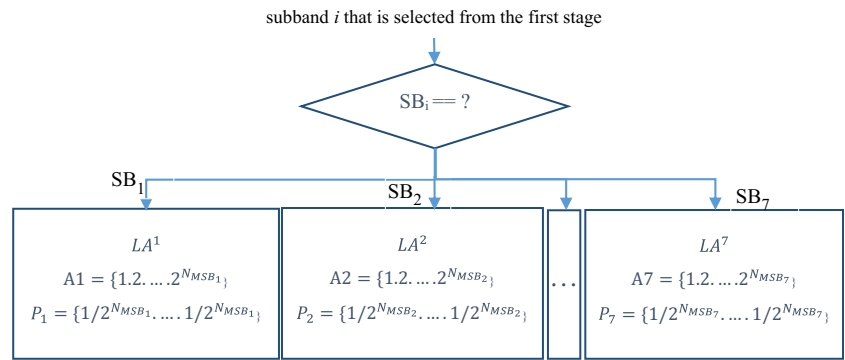
$$AchievedR = \sum_{y_f \in f} \log(c_{y_f}) \quad (15)$$

$$\beta_{femtoSB} = \begin{cases} 0 & \text{if } AchievedR > TotalRR \\ 1 - \frac{(AchievedR)}{(TotalRR)} & \text{otherwise} \end{cases} \quad (16)$$

Subsequently, LA updates its action selection vector, P_{SB} , using feedback and selects the best subband according to it. In each time step, the selected subband is used as the input for the second stage of the proposed method.

Furthermore, to improve the resource utilization of the femtocells, the proposed method employs the second learning stage to choose a portion of the subband (in terms of some mini-subbands). In this learning stage, only FUE feedbacks and demands and the selected subband from the first stage are used. This stage learns which mini-subbands should be selected with regard to the requested data rate of the FAP users. For this stage, the FAP maintains distinct LAs per each subband to learn the mini-subband selection strategy per each subband, independently. As shown in Fig. 5, LA^i is the LA structure used to choose the appropriate mini-subbands from subband i . Regarding the selected subband from the first stage, the set of possible actions of each LA is different, and the size of the action set depends on the number of mini-subbands (N_{MSB_i}) that constitute that subband (i.e., the size of the action set is $2^{N_{MSB_i}} - 1$). The set of actions is modeled by integers, i.e., $\alpha^i = \{1, 2, \dots, 2^{N_{MSB_i}} - 1\}$. where the location of the j^{th} bit in the binary representation of each integer shows that the j^{th} mini-subband has been selected. For instance, for a selected subband with two mini-subbands, the action set is defined as $\{1, 2, 3\}$ with binary equivalents $\{01, 10, 11\}$ that represent the selection of the first, second, or both mini-subbands. Note that the action 00 is removed from action set. Each LA (LA^i) has a probability vector, i.e., P_{MSB}^i for selecting mini-subbands. Initially, the probabilities of all the actions are the same and are equal to $(1/N_{MSB_i})$. In each time step τ , each FAP autonomously selects subband i with regard to P_{SB} . Then, each FAP chooses one or more mini-subbands from the selected subband according to P_{MSB}^i and allocates them to its users according to the round-robin scheduling algorithm. Subsequently, according to the quality of user feedbacks, P_{MSB}^i is updated (using Eqs. (11) and (12)).

Fig. 5 Independent learners in the second stage of TSLACA



3.5 Enhanced two-stage LA-based channel allocation

In this section, an extended version of TSLACA, namely enhanced TSLACA (ETSLACA), is proposed to further improve spectral efficiency and convergence speed. In this method, the second learning stage is different from that of TSLACA. As mentioned in the previous section, the second stage of TSLACA considers one LA per each subband, the action set of which includes all possible combinations of mini-subband selection. However, for wide subbands (such as the first subband in Fig. 2), this action space is large and lowers convergence speed. In ETSLACA, each FAP has a group of learners with a smaller action set per each subband. As a result, convergence speed rises and performance improves especially in fast varying network conditions.

Here, for each subband i , each FAP maintains independent learners for different states that are separated based on the aggregate demand of its users. In each time step, each FAP estimates the number of required mini-subbands according to its users’ requested bit rate. Then, the FAP selects one of the learners with regard to this estimation. The action set of the selected learner (LA^i_j) includes all possible combinations that have an estimated number of mini-subbands (j mini-subbands). If the number of the required mini-subbands is bigger than that of the available mini-subbands, then the LA that is relevant to the number of the latter is used. Therefore, actions are represented by integers, where the integers that constitute an LA action set have the same number of 1s in their binary representation. Therefore, the size of the action set for learner j (i.e., LA^i_j) is equal to $(N_{MSB} - j)$. In other words, each FAP contains a series of LAs with a different number of actions (i.e., $(N_{MSB} - 1) \cdot (N_{MSB} - 2) \dots (N_{MSB} - N_{MSB})$ actions). Figure 6 shows the flowchart of the second proposed method.

For example, if the selected subband from the first stage is SB_i with nine mini-subbands and the total requested bandwidth of the FAP is two mini-subbands, the FAP uses the second LA with the action set $\alpha = \{3.5 \dots 384\}$. The binary representation of all integers in set α contains two 1s that indicate the selected mini-subbands. For instance, if the

selected action were 5, the FAP could use only the first and third mini-subbands. Restricting the action set to a limited number of mini-subbands makes the convergence faster. Figure 7 shows the action set of all second stage learners, assuming the scheme in Fig. 2c.

4 Simulation results

This section evaluates the proposed methods through several simulation scenarios where the performance metrics are utilization, the average of the log-scaled data rate of all users (given by Eq. (17)), and outage probability (as shown in Eq. (18)). In Eq. (17), MUE_{total} and FUE_{total} represent the total

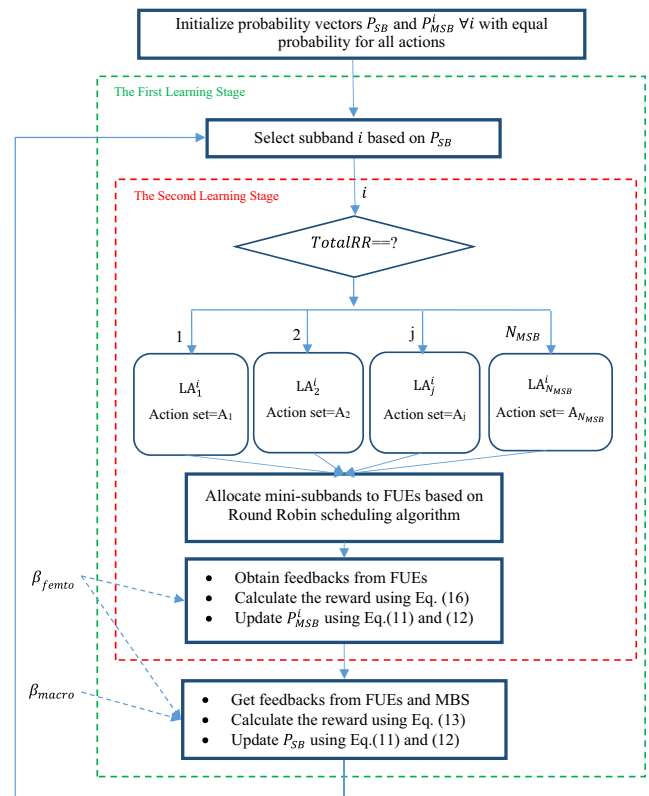


Fig. 6 The flowchart of enhanced TSLACA method

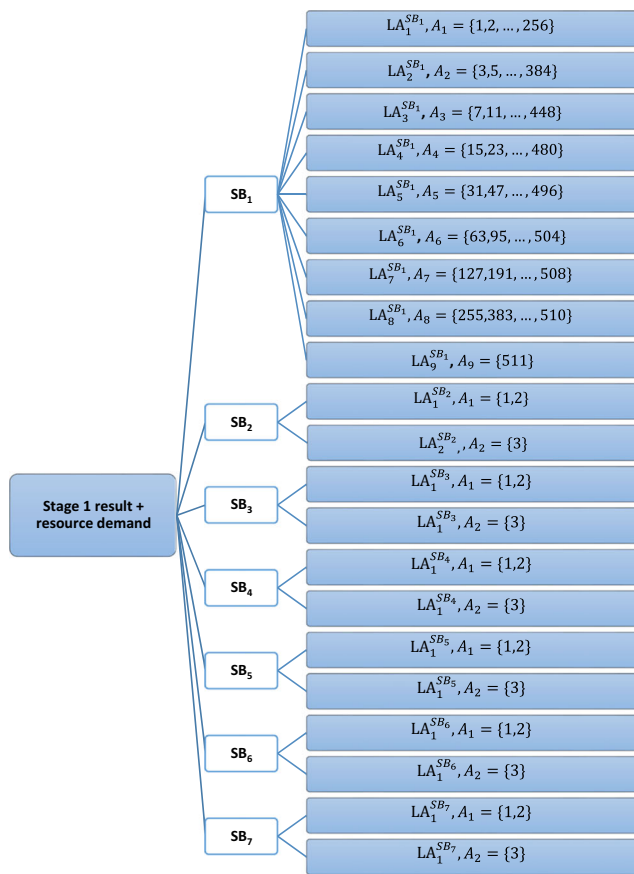


Fig. 7 The learners and their action sets for the second stage of ETSLACA (assuming the scheme of Fig. 2c)

number of macrocell and femtocell users, respectively. In Eq. (18), the outage probability is defined as the probability of the instantaneous SINR of the user falling below the threshold, γ . It is clear that P_{outage} depends on SB and MSB .

$$C_{avg} = C_{total} / (MUE_{total} + FUE_{total}) \tag{17}$$

$$P_{outage} = \Pr(SINR < \gamma) \tag{18}$$

The downlink communication is considered in an OFDM-based HetNet with nine cells (Fig. 8) which has been simulated at link level in MATLAB. Simulation parameters are given in Table 2. The proposed methods are evaluated under fixed and dynamic network conditions compared to strict FFR and OSFFR schemes. All simulation results are the average of ten simulation runs with different seeds.

In the simulation scenarios, the MBSs are located in the center of the cells. However, femtocells are randomly distributed in the cell area. The number of femtocells in each cell is a function of the normal distribution with parameters $(\mu_{FAP}, \sigma_{FAP})$. Moreover, macro/femto users are randomly distributed in the cell area. Similarly, the number of MUEs and the number of FUEs are a function of the normal distribution with parameters

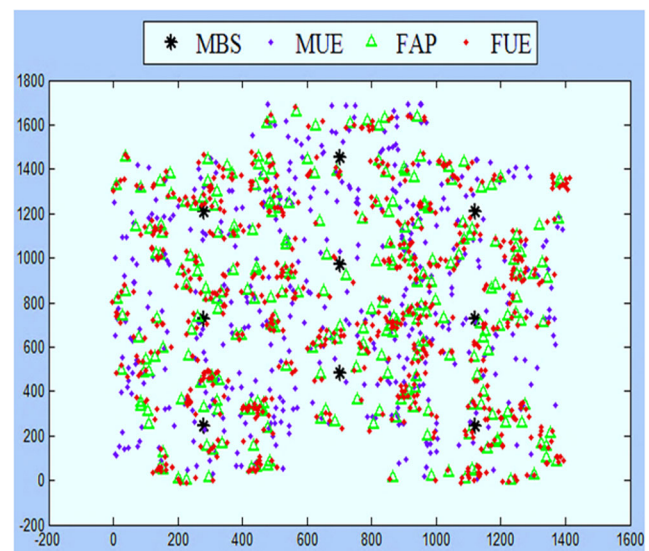


Fig. 8 Simulated network (including MBSs, MUEs, FAPs, and FUEs)

$(\mu_{MUE}, \sigma_{MUE})$ (and) $(\mu_{FUE}, \sigma_{FUE})$ (respectively). Each user requests a random number of subchannels based on normal distribution function with an average of three subchannels and standard deviation of 1. In the following subsections, we evaluate the performance of the proposed schemes for $N_{MSB} = 11$ and $N_{MSB} = 21$, under fixed network conditions, under network condition change at the 1000th time step, and under iterative changes of network conditions.

4.1 The simulation results under fixed network conditions

In this section, the simulation results of the proposed methods are compared to strict FFR and OSFFR under fixed network conditions (i.e., fixed number of users with constant demand).

Table 2 Simulation parameters

Parameter	Value
Radius of a femtocell	30 (m)
MBS transmit power	1.742 (w)
FAP transmit power	0.02 (w)
Average number of MUEs in a macrocell (μ_{MUE})	50
Standard deviation(SD) of the number of MUEs in a macrocell (σ_{MUE})	10
Average number of FAPs in a macrocell (μ_{FAP})	30
SD of the number of FAPs in a macrocell (σ_{FAP})	5
Average number of FUEs in a femtocell (μ_{FUE})	3
SD of the number of FUEs in a femtocell (σ_{FUE})	1
Channel bandwidth	1.5 (MHz)
Number of subchannels	100
Subcarrier spacing (ΔB)	15 (KHz)
White noise power spectral density (N_0)	-174 (dBm/Hz)
Learning parameter (a)	0.1
ϵ_p	0.5

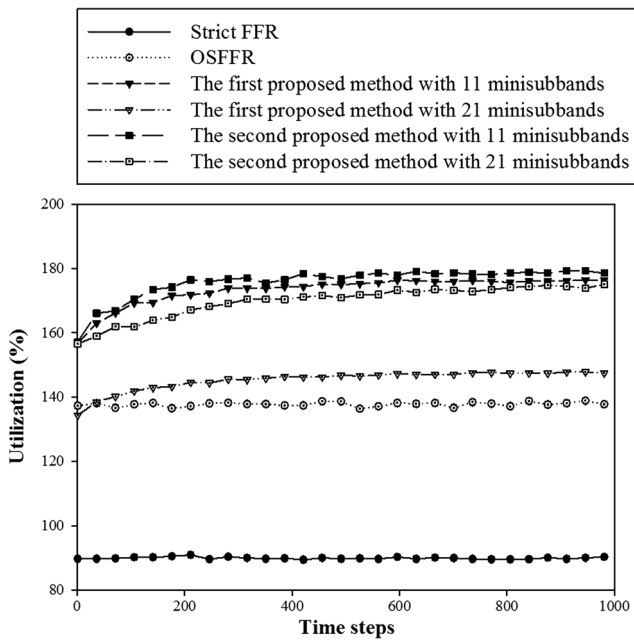


Fig. 9 Utilization comparison under fixed network conditions over 1000 time steps

As shown in Figs. 9 and 10, using the proposed methods, utilization and log-scaled average available capacity for each UE (C_{avg}) improve during the learning stages, and the average (over multiple runs) is higher than that of conventional strict FFR and OSFFR schemes. The results show that for the configuration with $N_{MSB} = 11$, both proposed methods demonstrate better results while strict FFR has the worst results.

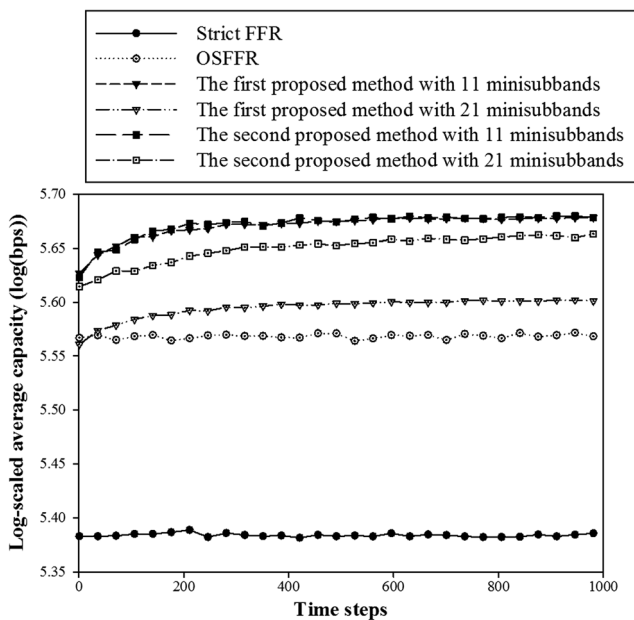


Fig. 10 Comparison of log-scaled average link capacity of UEs under fixed network conditions over 1000 time steps

Moreover, the results show that the second proposed method is better than the first. In fact, selecting the appropriate learner with regard to user requirements leads to better utilization of radio resources. Note that for the second proposed scheme with $N_{MSB} = 11$, C_{avg} and utilization are the best. Besides, the second proposed method with $N_{MSB} = 21$ shows considerable improvement compared to the first. In fact, the ability of the second proposed method in better utilization of radio resources leads to higher performance when the quantization of frequency fractions is finer.

Figure 11 compares the methods regarding expectation from outage probability, i.e., $E(P_{outage})$. As demonstrated, using the proposed methods, $E(P_{outage})$ decreases over the learning steps. The values of $E(P_{outage})$ for both methods under $N_{MSB} = 11$ and $N_{MSB} = 21$ are very close to each other and are near to zero while strict FFR has the worst results.

4.2 Simulation results under random changes at the 1000th time step

Here, we evaluate the proposed schemes over 2000 time steps with network condition changes at the 1000th time step. In this time step, UEs turn off with the probability $p = 0.1$ and the FAPs without any connected UEs are consequently turned off. In addition, some new UEs and FAPs are turned on randomly. As demonstrated in Figs. 12 and 13, the utilization and log-scaled average available capacity of the UEs (C_{avg}) increase over time. However, by random changes in the 1000th step, both utilization and C_{avg} fall and begin to rise again by the following learning steps. Note that the results of the proposed

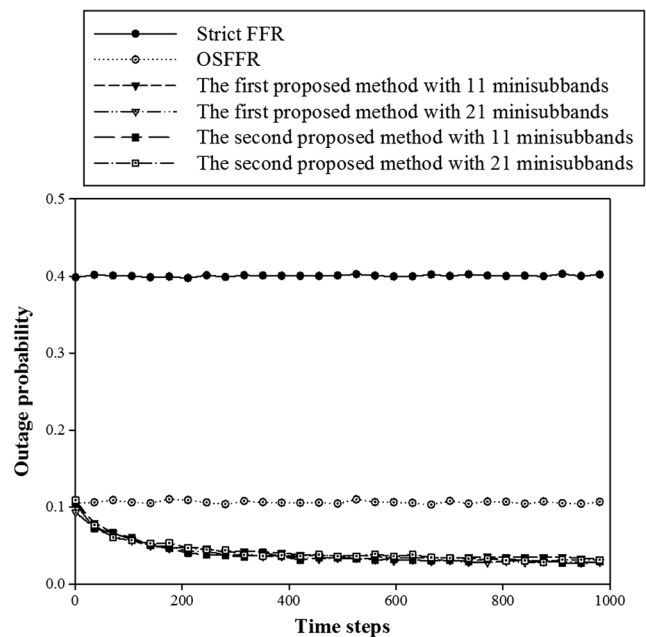


Fig. 11 $E(P_{outage})$ comparison under fixed network conditions over 1000 time steps

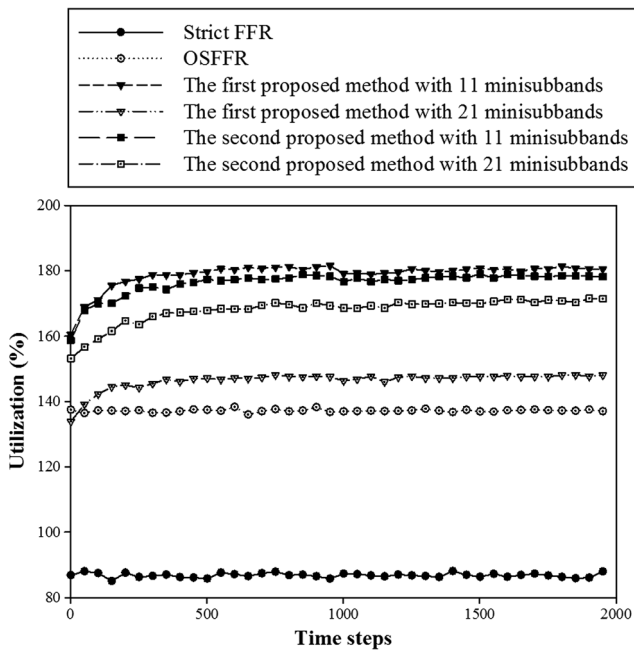


Fig. 12 Utilization comparison over 2000 time steps assuming random network changes at the 1000th time step

methods are better than those of both strict FFR and OSFFR. With $N_{MSB} = 21$, the results of the second proposed method are better than those of the first. Also, the results of both proposed methods with $N_{MSB} = 11$ are close.

Furthermore, as shown in Fig. 14, $E(P_{outage})$ rises very little in the 1000th step and, by following the learning process, decreases again. The results of the proposed methods for both

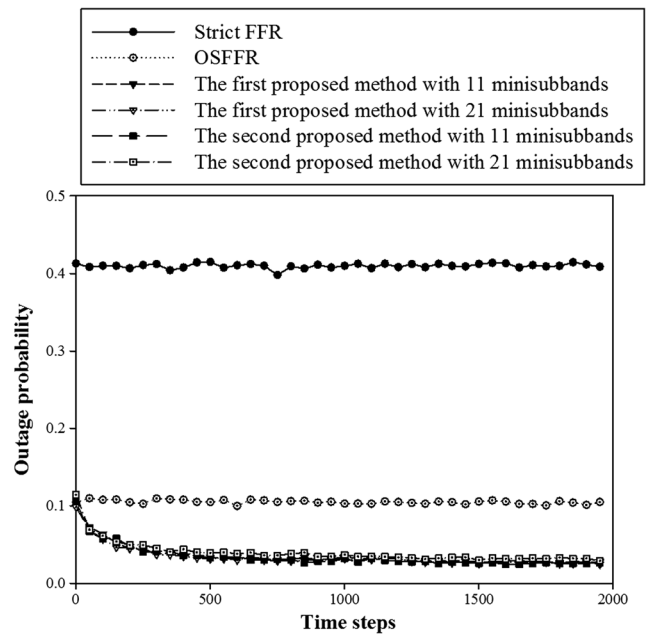


Fig. 14 $E(P_{outage})$ comparison over 2000 time steps assuming random network changes at the 1000th time step

subband quantization levels are very close and are much better than those of strict FFR and OSFFR.

4.3 Simulation results under continuous network changes

Finally, we evaluate the performance of a network with frequent changes. As shown in Figs. 15 and 16, utilization and

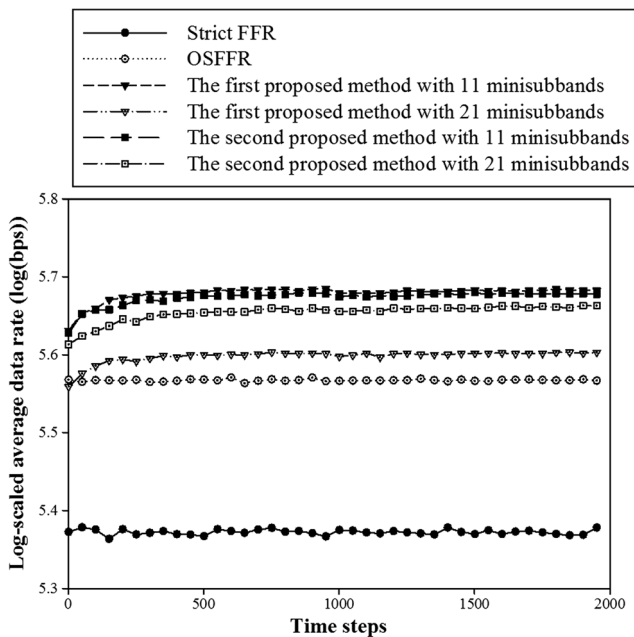


Fig. 13 Log-scaled average capacity comparison over 2000 time steps assuming random network changes at the 1000th time step

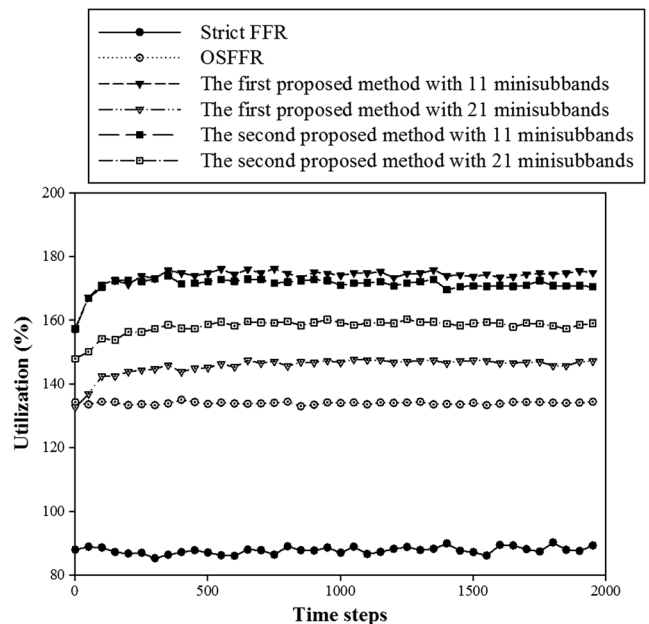


Fig. 15 Utilization comparison under continuous network changes

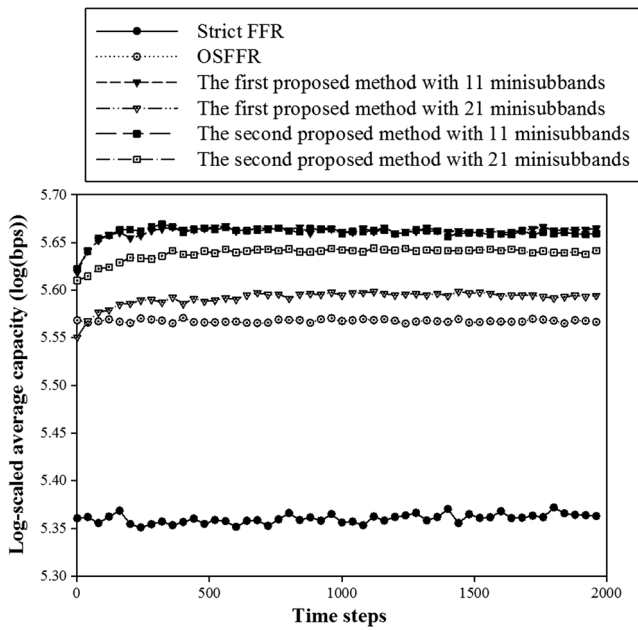


Fig. 16 Log-scaled average data rate comparison under continuous network changes

C_{avg} in the proposed methods are higher and improve over time compared to strict FFR and OSFFR. The first proposed scheme has the best utilization and C_{avg} under $N_{MSB} = 11$ and the results of the second proposed method are close those of the first. Every 200 time steps once, random changes occur, and utilization and C_{avg} fall. However, following the learning process, these values increase again.

Moreover, the proposed methods have the best $E(P_{outage})$ compared to the strict FFR and OSFFR methods. As shown in

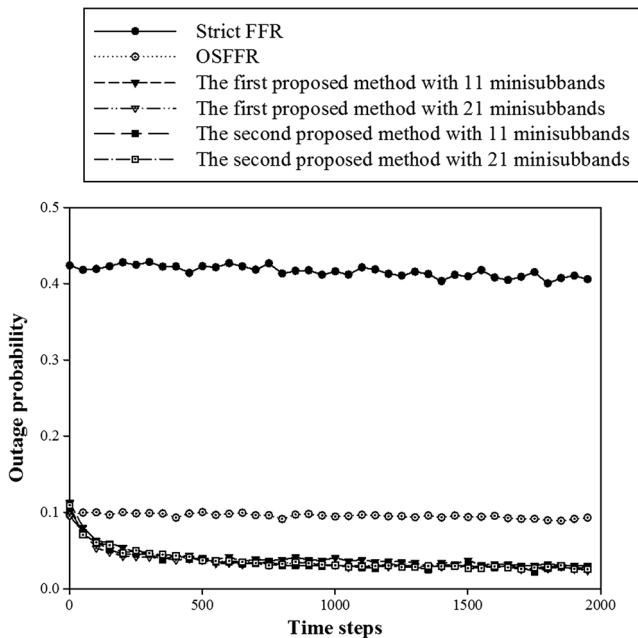


Fig. 17 Outage probability comparison under continuous network changes

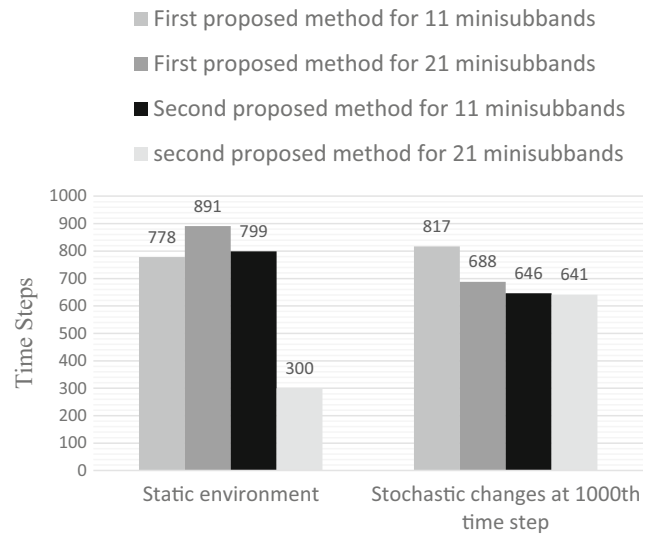


Fig. 18 The number of the time steps for convergence

Fig. 17, the worst $E(P_{outage})$ belongs to strict FFR. However, in the proposed methods, as random changes occur, $E(P_{outage})$ rises a little, and by following the learning process, its value decreases again.

4.4 Comparison of convergence speed

In this subsection, the number of time steps required for the proposed methods to converge to a stable condition is evaluated. To this aim, the standard deviation of the log-scaled average data rate is evaluated once every ten time steps and the number of time steps required to reach a specified standard deviation is calculated.

As shown in Fig. 18, the best convergence speed belongs to the second proposed method with 21 mini-subbands, i.e., 300 time steps under fixed network conditions. Also, after random changes at the 1000th time step, the number of time steps required for the second proposed method to reconverge is less than that of the first. Assuming LTE networks where each Transmission Time Interval (TTI) is 1 ms, we need two TTIs per each learning time step, namely the first TTI as the time of making the decision and the second as the time of allocating resources, which leads to 2 ms per each time step. As a result, the convergence time of the second proposed method is less than 1 s which is reasonable for common network changes.

5 Conclusion and future work

This paper presents two distributed resource allocation mechanisms for femtocell networks based on the stochastic learning automata structure. These methods employ a two-stage learning process for the selection of appropriate radio resources. In the proposed methods, each FAP selects the best fraction of frequency resources without any previous information about

network conditions and based only on the demand of its UEs, their received feedbacks, and MBS quality feedbacks. Simulation results show that the proposed methods display better spectrum utilization and outage probability under fixed and dynamic network conditions compared to strict FFR and OSFFR. Moreover, the results show that the second proposed method is better than the first in terms of convergence speed. Extending the proposed methods by employing learning techniques in MBSs as well as learning the size of different frequency fractions can be addressed in future work.

References

- Garcia V, Chen CS, Lebedev N, Gorce J-M (2011) Self-optimized precoding and power control in cellular networks. In 22nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC). IEEE, pp 81–85
- Barbarossa S, Sardellitti S, Carfagna A, Vecchiarelli P (2010) Decentralized interference management in femtocells: A game-theoretic approach. In Proceedings of the Fifth International Conference on Cognitive Radio Oriented Wireless Networks & Communications (CROWNCOM). IEEE, p 1–5
- Putzke M, Wietfeld C (2013) Self-organizing fractional frequency reuse for femtocells using adaptive frequency hopping. In Wireless Communications and Networking Conference (WCNC). IEEE, pp 434–439
- Dirani M, Altman Z (2010) A cooperative reinforcement learning approach for inter-cell interference coordination in OFDMA cellular networks. In Proceedings of the 8th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt). IEEE, pp 170–176
- Saad H, Mohamed A, and ElBatt T (2014) Cooperative Q-learning techniques for distributed online power allocation in femtocell networks, *Wireless Communications and Mobile Computing*
- Cheng S-M, Lien S-Y, Chu F-S, Chen K-C (2011) On exploiting cognitive radio to mitigate interference in macro/femto heterogeneous networks. *IEEE Wirel Commun* 18(3):40–47
- Stefan AL, Ramkumar M, Nielsen RH, Prasad NR, Prasad R (2011) A QoS aware reinforcement learning algorithm for macro-femto interference in dynamic environments. In 3rd International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). IEEE, pp 1–7
- Bennis M and Niyato D (2010) A Q-learning based approach to interference avoidance in self-organized femtocell networks, in *GLOBECOM workshops (GC Wkshps)*, IEEE, 2010, pp. 706–710
- Zhang W-D, Wang Y, Xu M-Y, Zhang P (2013) Interference management with adaptive fractional frequency reuse for LTE-A femtocells networks. *J China Univ Posts Telecommun* 20:86–91
- Bernardo F, Agustí R, Pérez-Romero J Sallent O (2009) Distributed spectrum management based on reinforcement learning, in *Cognitive radio oriented wireless networks and communications, 2009. CROWNCOM'09. 4th International Conference on*, pp 1–6
- Bernardo F, Agustí R, Perez-Romero J, Sallent O (2011) Intercell interference management in OFDMA networks: a decentralized approach based on reinforcement learning. *IEEE Trans Syst Man Cybernet C Appl Rev* 41:968–976
- Lien S-Y, Tseng C-C, Chen K-C, Su C-W (2010) Cognitive radio resource management for QoS guarantees in autonomous femtocell networks. In *International Conference on Communications (ICC)*. IEEE, p 1–6
- Lee P, Lee T, Jeong J, Shin J (2010) Interference management in LTE femtocell systems using fractional frequency reuse. In 12th International Conference on Advanced Communication Technology (ICACT), vol. 2. IEEE, pp 1047–1051
- Galindo-Serrano A, Giupponi L (2011) Downlink femto-to-macro interference management based on fuzzy Q-learning. In *International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*. IEEE, pp 412–417
- Saquist N, Hossain E, Kim DI (2013) Fractional frequency reuse for interference management in LTE-Advanced HetNets. *IEEE Wirel Commun* 20:113–122
- Kim J, Jeon WS (2012) Fractional frequency reuse-based resource sharing strategy in two-tier femtocell networks. In *Consumer Communications and Networking Conference (CCNC)*. IEEE, pp 696–698
- Novlan TD, Andrews JG (2013) Analytical evaluation of uplink fractional frequency reuse. *IEEE Trans Commun* 61:2098–2108
- Mishra S, Murthy CSR (2016) An efficient location aware distributed physical resource block assignment for dense closed access femtocell networks. *Comput Netw* 94:164–175
- LeAnh T, Tran NH, Saad W, Moon S, Hong CS (2016) Matching-based distributed resource allocation in cognitive femtocell networks, in *IEEE/IFIP Network Operations and Management Symposium*, pp. 61–68
- Nazir M, Bennis M, Ghaboosi K, MacKenzie AB, Latva-aho M (2010) Learning based mechanisms for interference mitigation in self-organized femtocell networks. In *Signals, Systems and Computers (ASILOMAR), 2010 Conference Record of the Forty Fourth Asilomar Conference on*, IEEE, pp 1886–1890
- Semasinghe P, Hossain E, Zhu K (2015) An evolutionary game for distributed resource allocation in self-organizing small cells. *IEEE Trans Mob Comput* 14:274–287
- Xu C, Sheng M, Wang X, Wang CX, Li J (2014) Distributed subchannel allocation for interference mitigation in OFDMA femtocells: a utility-based learning approach. *IEEE Trans Veh Technol* 64:2463–2475
- Oh C-Y, Chung MY, Choo H, Lee T-J (2013) Resource allocation with partitioning criterion for macro-femto overlay cellular networks with fractional frequency reuse. *Wirel Pers Commun* 68: 417–432
- Salati AH, Nasiri-Kenari M, Sadeghi P (2012) Distributed subband, rate and power allocation in OFDMA based two-tier femtocell networks using Fractional Frequency Reuse. In *Wireless Communications and Networking Conference (WCNC)*. IEEE, pp 2626–2630
- Saadat S, Chen D, Jiang T (2016). QoS Guaranteed Resource Allocation Scheme for Cognitive Femtocells in LTE Heterogeneous Networks with Universal Frequency Reuse. *Mob Netw Appl* 21(6):930–942
- Esfahani MN and Ghahfarokhi BS (2014) Improving spectrum efficiency in fractional allocation of radio resources to self-organized femtocells using Learning Automata, in *Telecommunications (IST), 2014 7th International Symposium on*, pp. 1071–1076
- Li QC, Hu RQ, Xu Y, Qian Y (2013) Optimal fractional frequency reuse and power control in the heterogeneous wireless networks. *IEEE Trans Wirel Commun* 12(6):2658–2668
- Thathachar MA, Sastry PS (1986) Relaxation labeling with learning automata. *IEEE Trans Pattern Anal Mach Intell* 8(2):256–268
- Narendra K, Thathachar M (1989) *Learning automata: an introduction*. Printice-Hall, New York
- Ünsal C (1997) Intelligent navigation of autonomous vehicles in an automated highway system: learning methods and interacting vehicles approach, Virginia Polytechnic Institute and State University