Joint Pricing and Power Allocation for Uplink Macrocell and Femtocell Cooperation

Tuan LeAnh, Nguyen H. Tran, S.M.Ahsan Kazmi, Thant Zin Oo and Choong Seon Hong*
Department of Computer Engineering, Kyung Hee University, 449-701, Korea
Email: {latuan,nguyenth,ahsankazmi,tzoo and cshong*}@khu.ac.kr

***Abstract*—In this paper, we study cooperation among mobile users for uplink in two-tiers heterogeneous wireless networks. In our cooperative model, a macrocell user equipment can relay its data via a femtocell user equipment when it cannot connect to its macro base station or any femtocell base stations directly. In this scenario, the macrocell user equipment tries to find the best relay user in a set of candidate relay femtocell user equipments to maximize its utility function. Additionally, the candidate relay femtocell user equipments give a pricing-based strategy per each power unit to the macrocell user equipment along with power level at relay femtocell user equipments which would be used for relaying data in order to maximize both the relay femto and macrocell user equipment's utility function. In static network environment, this problem is formulated as a Stackelberg game. Moreover, in stochastic network environment we find stochastic optimization in a long-term for both the utility functions by modeling the problem as a restless bandit problem. Simulation results illustrate the efficiency of our proposal.**

Keywords*—Heterogeneous Wireless Network, Femtocell Network, Cooperation, Staclkelberg game, Restless Bandit problem, Stochastic Optimization.

I. Introduction

Recently, the novel wireless communication paradigm has shifted to a future wireless network such as deployment of femtocell network [1], [2]. One of the paradigms is known as the HetNets with coexistence of two-tier which comprises of a macrocell underlaid on *femtocell base stations* (FBSs). The femtocell network is a solution to improve both spectrum efficiency and network capacity because they can act as an enabler for offloading traffic in heavily crowded cells [1],[2],[3].

The deployment of femtocell network poses a number of challenges which needs to be addressed to enhance the overall system performance: 1) How to share traffic-load among femtocells and macrocell, i.e Efficiently sharing traffic-load between a heavy-load cell and lightly loaded cell? 2) Providing high quality of data connection to mobile users which lie outside or at border coverage of the femtocells or macrocell base stations? One approach to address this problem is using handover technique combined with cooperative modeling, as in [3]. However, they only focus on avoiding interference in femtocell networks by using coalitional game approach.

In our work, we investigate a cooperative model for uplink mode of the macrocell and femtocells cooperation. The *macrocell user equipments* (MUEs) which are willing to handover to femtocells but cannot handover to the FBSs directly, will carry a handover process by the help of a *femtocell user equipment* (FUE). In cooperative models, some challenges faced are mentioned in [3]: 1) Modeling of cooperation among users belonging to different tiers? 2) What is the price for cooperation and when is cooperation beneficial? 3) How to provide incentive to encourage cooperation? In our framework we consider these problems, consequently.

Our paper consider two environment scenarios: static and stochastic network. In static network environment the decision of the MUE to select FUE for relaying data is modeled based on one-shot Stackelberg game approach at each time slot independently. The Stackelberg game [4], [5] captures a trading between a MUE and a candidate relay FUE. Firstly, in our work we investigate a joint pricing and power allocation scheme where the candidate relay FUEs give a price per each power unit to the MUE. This joint power-pricing will maximize the utility function value for the relay FUE and the MUE, simultaneously. The optimal relay selection in this scenario is determined by making decision of the MUE following a greedy scheme. The relay selection helps maximizing not only the MUE's utility but also maximizes the relay FUE's utility.

Secondly, we consider an optimal relay selection in stochastic network environment. The stochastic network environment parameters are mentioned with the channel gain states information, residual energy states of the candidate relay FUEs's battery and own traffic of each candidate relay FUE [6]. These parameters are observed based on history information and Markovian is used for modeling changes of all the above impacts. The stochastic network environment in our model is formulated as a restless bandit problem to predict the upcoming relay FUEs states in order to maximize a pair of expected utility values (the MUE and relay FUE system) in a long-term.

The rest of this paper is organized as follows. The system model is presented in section II. The Stackelberg game formation of our model is discussed in section III. The stochastic network environment is given in section IV. Numerical results are illustrated in section V. Finally, section VI provides conclusion and future works.

II. Problem formulation

A. System model

We consider an uplink connection Orthogonal Frequency Division Multiple Access (OFDMA) of two-tier HetNets system including a macrocell and a set of femtocells as shown in Fig. 1. The HetNets consists of a set of $\mathcal{N} = \{1...N\}$ FUEs and a set of $\mathcal{K} = \{0, 1...K\}$ base stations. We distinguish macro and femto base stations by denoting base station with

This work was supported by the ICT R&D program of MSIP/IITP, Republic of Korea. (2014-044-011-003, Open control based on distributed mobile core network).
*Dr. CS Hong is the corresponding author.

index k equals "0" for the MBS and $k \in \mathcal{K}\backslash\{0\}$ for the FBSs. Moreover, let n_k be the FUE that is connected to FBS k, $\forall n \in \mathcal{N},\ \forall k \in \mathcal{K}\backslash\{0\}$.

In our work, we consider a MUE that is out of the service range of its MBS and any nearby FBSs, this MUE cannot connect directly to any base station. In order to transmit data, this MUE needs assistance by the nearby FUEs which are idle and are willing to act as relays. In our scenario each FUE can only become a candidate relay and serve for one MUE. The MUE, which has to pay an incentive for relaying data to the FUE, is denoted by i. Let $\mathcal{N}^i_{n_k}$ be a set of candidate relay FUE n that associated with BS k for MUE i, $\forall n \in \mathcal{N},\ \forall k \in \mathcal{K}$.

Without loss of generality, the considered system is time-

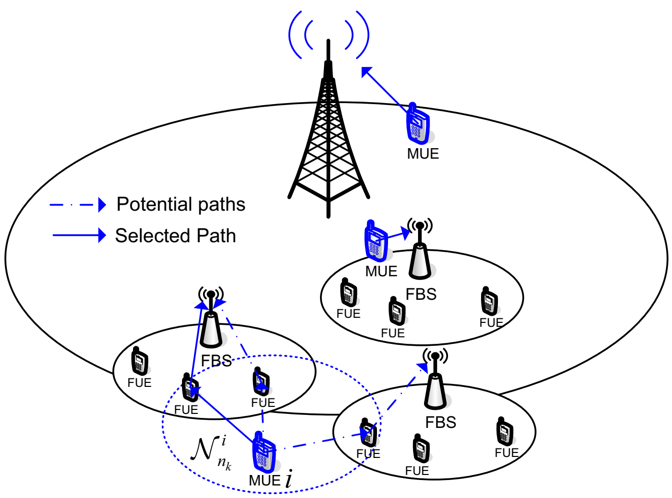

Fig. 1: System model.

slotted. Each relay selection is decided at the beginning of each time slot. Additionally, MUE i's original channel, which is registered to the MBS, is kept till MUE i's data is relayed via the relay FUE. In order to successfully transmit data, the subsequent section of data transmission model provide the details. Moreover, for stochastic environment we need to consider certain parameter discussed in subsection stochastic model which affect the decision process of the MUE for relay selection.

B. Data transmission model

The data transmission can be categorized as follows:

1) The direct transmission without relay: We consider the scenario when MUE i establishes a connection with the base station k directly [4], [7] as follows:

$$R^d_i = B_w \log\left(1 + \frac{P_i|h_{i,k}|^2}{\sigma^2}\right), \tag{1}$$

where B_w is the bandwidth of a channel that MUE i registered to the MBS, P_i is the maximum power level that can be allocated for MUE i's when the handover occurs $|h_{i,k}|^2$ is the channel gain between MUE i and the base station k, σ^2 is Gaussian noise at receiver.

2) The cooperative transmission via relay FUE: In this case the transmission is considered via relay FUE, the *amplify-and-forward* (AF) protocol is applied in the relay transmission with two stages [8]. We treat one input and two outputs complex Gaussian noise channel by using maximal ratio combining (with N equals 1) as follows:

$$R^{\text{relay}}_{i,n_k} = \frac{B_w}{2}\log\left(1 + SINR^{AF}_{i,n_k}\right),\ \forall n_k \in \mathcal{N}^i_{n_k} \tag{2}$$

where $SINR^{AF}_{i,n_k}$ is Signal to Interference plus Noise Ratio (SINR) of AF method and be computed as below:

$$\begin{aligned} SINR^{AF}_{i,n_k} = {} & \frac{P_i|h_{i,k}|^2}{\sigma^2} \\ & + \frac{P_i P_{n_k,i}|h_{i,n_k}|^2|h_{n_k,k}|^2}{\sigma^2\left(P_i|h_{i,n_k}|^2 + P_{n_k,i}|h_{n_k,k}|^2 + \sigma^2\right)}, \end{aligned} \tag{3}$$

where $P_{n_k,i}$ is power level of relay FUE n_k for MUE i. $|h_{i,n_k}|^2$ and $|h_{n_k,k}|^2$ are the channel gains from MUE i to the candidate relay FUE n_k and from the candidate relay FUE n_k to base station k, respectively.

In our scenario, FUE n_k can be selected for cooperating transmission iff $R^d_i < R^r_{i,n_k}$. In order to describe the relationship between the potential relay FUE n_k and MUE i, we denote $a_{i,n_k} = \{0, 1\}$ as the action for choosing relay FUE n_k from MUE i, $\mathcal{A} = \{a_{i,n_k}\}$ is a matrix of actions. Here "1" represents relay FUE n_k which is selected as a relay for MUE i, "0" represents the candidate relay FUE n_k which is not selected for relaying data by MUE i.

C. Stochastic Model

In stochastic network environments we analyze three parameters discussed below and study their affect on the performance of the system.

1) Channel gain model: The assumed stochastic network environment is considered with block-Rayleigh fading channels as a *Finite State Markov Chain* (FSMC) model [9]. The average channel gain, $\bar{\sigma}_{s,d}(t) = \mathbb{E}[h_{s,d}(t)]$ (**s**ource node to **d**estination node), is modeled as a random variable according to a L-states Markov chain, which has a finite state space denoted by $\mathcal{C} = \{C_0, C_1, .., C_l, ...C_{L-1}\}$. Here, $\mathbb{E}(.)$ represents expectation. Let $\phi_{C_l,C_{l'}(t)}$ be the probability that $\bar{\sigma}_{s,d}(t)$ transmits from state C_l to $C_{l'}$ at epoch t. The expression of channel state transition probability matrix is presented as follows:

$$\Phi_{s,d}(t) = [\phi_{C_l,C_{l'}}(t)]_{L\times L}, \tag{4}$$

where $\phi_{C_l,C_{l'}}(t)$=Pr$\{\bar{\sigma}_{s,d}(t+1) = C_{l'}|\bar{\sigma}_{s,d}(t) = C_l\}$, for C_l, $C_{l'} \in \mathcal{C}$.

2) The own traffic model of candidate relay users: We consider the own traffic state model as two-state Markov model to represent the utilized candidate relay FUE in epoch time t, denoted by $G_{n_k}(t) \in \mathcal{G}$ = {Busy, Idle} as in [6], [10], [11]. "Busy" state means that the FUE has its own data to transmit and relaying is not permitted in this state; while in the "Idle" state it can relay data. Hence, the candidate relay FUE's traffic state transition probability matrix can be written as follows:

$$\Theta(t) = [\theta_{g_{n_k}, g'_{n_k}}(t)]_{2\times 2}, \tag{5}$$

where $\theta_{g_{n_k}, g'_{n_k}}(t)$ = $\Pr\{G_{g_k^r}(t+1) = g'_{n_k} | G_{n_k}(t) = g_{n_k}\}$, for $g_{n_k}, g'_{n_k} \in \mathcal{G}$.

3) Energy model: Since the battery energy of the FUE decreases depending upon applications running on the FUE, we cannot exactly know the energy state at the next time slot [12]. Therefore, the residual battery energy can be modeled as a random variable $e_{n_k}(t)$ with two discrete levels, low and high energy level denoted by $\mathcal{E}$ = $\{\mathcal{E}_0, \mathcal{E}_1\}$. Here "$\mathcal{E}_0$" corresponds to FUE not acting as a relay while $\mathcal{E}_1$ corresponds to FUE acting as a relay. The transition model of the residual energy levels of each candidate relay FUE follows the Markov chain as in [12], [13]. We adopt this model and define the energy-state-transition probability matrix of relay FUE n_k as:

$$\Omega_{n_k}(t) = [\omega_{h_{n_k}, h'_{n_k}}(t)]_{2\times 2}, \tag{6}$$

where $\omega_{h_{n_k}, h'_{n_k}}(t)$ = $\Pr\{E_{h_{n_k}}(t+1) = h'_{n_k} | E_{n_k}(t) = h_{n_k}\}$, for $h_{n_k},\ h'_{n_k} \in \mathcal{E}$. The energy level is assumed to be reduced by a fixed amount after every data-transmission action [12], [14].

In real systems, the above parameters can be obtained in the aforementioned transition probability matrices from the history observation based on feedback information received at the end of each time slot.

III. Stackelberg Game Analysis and Static Formation in Trading Exchange

This section expresses the cooperative payment and power allocation among MUE i and the candidate relay FUEs. MUE i is formulated as a buyer (or follower) and the candidate relay FUEs as the sellers (or leaders). Each of relay user $n_k \in \mathcal{N}_{n_k}^i$ has an incentive to earn, the payment which not only covers their forwarding cost but also obtains as much profit as possible [15]. Hence, when the candidate relay FUE n_k becomes a relay of i, relay FUE n_k sets a price and a power level to maximize the utility function for payment represented as follows:

$$\max \;\; U_{n_k} = (\omega_{n_k,i} - \zeta_{n_k}) P_{n_k,i} \tag{7}$$

subject to:

$$0 \le \zeta_{n_k} \le \omega_{n_k,i}, \tag{8}$$

$$\text{variables} \;\; \{\omega_{n_k,i}\},$$

where ζ_{n_k} is the cost of relay FUE n_k for a power unit of relaying data, $\omega_{n_k,i}$ is the power unit price of MUE i that it pays to relay FUE n_k when relaying MUE i's data and $P_{n_k,i}$ is power level of relay FUE n_k for relaying data of i.

In order to compete with other relay FUEs, given power level demand from MUE i, relay FUE n_k gives a pricing-based strategy for relaying data of MUE i to maximize its utility function. At the MUE side, given pricing-based strategy from the relay FUE, the i gives a power level demand to the relay FUE to maximize MUE i's utility function as follows:

$$\max\!: \;\; U_i = R_{i,n_k}^{\text{relay}} - \omega_{n_k,i} P_{n_k,i} \tag{9}$$

subject to:

$$0 \le P_{n_k,i} \le P_{n_k}^{\max}, \tag{10}$$

$$P_i = P_i^{\max}, \tag{11}$$

$$\text{variables} \;\; \{\omega_{n_k,i} P_{n_k,i}\},$$

where $R_{n_k,i}^{\text{relay}}$ is data rate of MUE i via relay FUE n_k as computed in [3].

Through the backward induction computation, given price of the candidate relay FUE n_k and P_i^{max} of MUE i, MUE i requests an optimal power demand that contains the variable price ω_{n_k} from candidate relay FUE to maximize its utility function. This maximization problem can be solved by finding the root of the first derivation (by setting $\frac{\partial U_i}{\partial P_{n_k,i}} = 0$) as follows:

$$P_{n_k,i}^*(\omega_{n_k,i}) = \min\left(\left[\frac{-A\omega_{n_k,i} + \sqrt{B\omega_{n_k,i}^2 + C\omega_{n_k,i}}}{2\omega_{n_k,i} D}\right]^+, P_{n_k,i}^{\max}\right), \tag{12}$$

where the parameters A, B, C, D are determined in (13).

The optimum power level demand of relay FUE n_k is determined based on (12) that follow variable $\omega_{n_k,i}$. After that, relay FUE n_k takes an optimal price by optimizing problem (7) as: $\omega_{n_k,i}^* = \max\{\arg\max\left((\omega_{n_k,i} - \zeta_{n_k}) P_{n_k,i}^*\right), \zeta_{n_k}\}$. Due to $\frac{\partial^2 U_{n_k}(\omega_{n_k,i})}{\partial^2 \omega_{n_k,i}} < 0$, there exists an optimal point of price that maximizes (7). Consequently, the candidate relay FUE n_k adjusts an optimum pricing-based strategy $\omega_{n_k,i}^*$ to maximize (7) or take by first order of (7) to find the optimal solution and optimum utility function $U_{n_k}^*$. Hence, there exists a pair of optimum utility function values $\left(U_{n_k}^*, U_i^*\right)$.

IV. Stochastic Formation as the Restless Bandit Problem

In above section, we considered only static network environment. This section presents a stochastic optimization in stochastic network environment. We apply a restless bandit problem to formulate a stochastic relay selection problem as follows: we consider the candidate relay FUEs as the projects in restless bandit problem; each candidate relay FUE n_k can be in a state $i_{n_k}(t) \in \mathcal{S}_{n_k}$ in each time slot t = 1, 2, ...[12], [16]. According to their states, $M = 1$ out of n_k candidate relay FUEs is selected to work or set to be active (a_{n_k} = 1), and the remaining candidate relay FUEs are set to be passive (a_{n_k} = 0). The system reward $\mathcal{R}_{i_{n_k}(t)}^{a_{n_k}(t)}$ is earned when action $a_{n_k}(t)$ is taken, and their states change in a Markovian fashion, according to a transition probability matrix into state $j_{n_k}(t+1)$ with probability $p_{i_{n_k}, j_{n_k}}^a$. Rewards are discounted in time by a discount factor β. The candidate relay FUEs are selected overtime under an optimal policy $u^* \in \mathcal{U}$, where $\mathcal{U}$ is a set of all Markovian policies. Now we formulate and discuss our solution for the procedure of relay FUE selection as follows.

$$A = \frac{\left(\sigma^2 + P_i|h_{i,n_k}|^2\right)\left(2B_w^{-1}\sigma^2 + P_i\left(|h_{i,n_k}|^2 + |h_{i,k}|^2\right)\right)}{|h_{n_k,k}|^2\left(\sigma^2 + P_i|h_{i,k}|^2\right)}, \quad B = \left(\frac{P_i|h_{i,n_k}|^2\left(\sigma^2 + P_i|h_{i,n_k}|^2\right)}{|h_{n_k,k}|^2\left(\sigma^2 + P_i|h_{i,k}|^2\right)}\right)^2,$$
$$C = \frac{2B_w^{-1}P_i|h_{i,n_k}|^2\left(\sigma^2 + P_i|h_{i,n_k}|^2\right)\left(2B_w^{-1}\sigma^2 + P_i\left(|h_{i,n_k}|^2 + |h_{i,k}|^2\right)\right)}{2\ln 2|h_{n_k,k}|^2\left(\sigma^2 + P_i|h_{i,k}|^2\right)^2}, \quad D = \frac{\sigma^2 + P_i\left(|h_{i,n_k}|^2 + |h_{i,k}|^2\right)}{\sigma^2 + P_i|h_{i,k}|^2}. \quad (13)$$

A. Action space for relay selections

After the state transition of the candidate relay FUE at the beginning of each slot t, MUE i needs to make an urgent decision for selection of one candidate relay FUE from the $\mathcal{N}_{n_k}^i$. The composite action in time slot t is denoted by $A_{n_k}^i(t) = \{a_{n_k}(t), P_{n_k,i}(t), \omega_{n_k,i}(t)\}$. The first scenario $a_{n_k}(t) = 1$ corresponds to n_k is selected and it establishes an optimum power level $P_{n_k,i}^*$ and price $\omega_{n_k,i}^*$. The second scenario, if $a_{n_k}(t) = 0$ means that relay n_k was not selected to relay then the established power level and price is equal to zero. In order to simplify for consideration, we denote the action $a_{n_k}(t)$ = 1 corresponding to $A_{n_k}(t)$ = 1 and the action $a_{n_k}(t)$ = 0 correspond to $A_{n_k}(t)$ = 0.

B. State and transition Probability

The state of candidate relay FUE $n_k \in \mathcal{N}_{n_k}^i$ in the t-th epoch is denoted as $i_{n_k}(t)$ that is characterized by $\bar{\sigma}_{i,n_k}(t)$, $\bar{\sigma}_{n_k,k}(t)$, $\bar{\sigma}_{i,k}(t)$, $G_{n_k}(t)$, $e_{n_k}(t)$ for $i - n_k$ channel state, $n_k - k$ channel state, $i - k$ channel state, candidate relay FUEs usage and candidate relay user's energy state, respectively. Consequently, the state of a candidate relay FUE is the combination as follow:

$$i_{n_k}(t) = [\bar{\sigma}_{i,n_k}(t), \bar{\sigma}_{n_k,k}(t), \bar{\sigma}_{i,k}(t), G_{n_k}(t), e_{n_k}(t)]. \quad (14)$$

In practical, the changes of each above sub-state are independent with each other. Hence, the candidate relay FUE states will change in a Markovian fashion, and the finite-state space of each candidate relay FUE n_k as $\mathcal{S}_{n_k}$, with the transition probability matrix as below:

$$\mathcal{P}_{n_k}(t) = [\phi_{i,n_k}(t), \phi_{n_k,k}(t), \phi_{i,k}(t), \theta_{n_k}(t), \omega_{n_k}(t)]_{[S\times S]}, \quad (15)$$

where $\phi_{i,n_k}(t)$, $\phi_{n_k,k}(t)$, $\phi_{i,k}(t)$ are defined as in (4); $\theta_{n_k}(t)$, $\omega_{n_k}(t)$ are defined in (5), (6), respectively; and $S = L^3 \times 2 \times H$. The element of $\mathcal{P}_{n_k}(t)$ is $p_{i_{n_k},j_{n_k}}(t)$ which denotes the transition probability that the state of relay user n_k changes form i_{n_k} to j_{n_k}, where $i_{n_k}, j_{n_k} \in \mathcal{S}_{n_k}(t)$.

C. Expected System Reward

We aim to optimize the system reward in the long-term that corresponds to the restless bandit problem. We formulate the system reward to be pair of MUE i and candidate relay user utility functions which is effected by system states as in (15). Therefore, we define the immediate reward as follows:

$$\mathcal{R}_{i_{n_k}(t)}^{A_{n_k}(t)} = \{U_{n_k}(A_{n_k}(t), i_{n_k}(t)); U_i(A_{n_k}(t), i_{n_k}(t))\}, \quad (16)$$

where $U_i(A_{n_k}(t), i_{n_k}(t))$, $U_{n_k}(A_{n_k}(t), i_{n_k}(t))$, respectively denote the utility function of MUE i and selected candidate relay FUE when the system takes action $A_{n_k}(t)$ at state $i_{n_k}(t)$ in time slot t.

For a stochastic process, a maximum immediate value does not mean the maximum expected long-term accumulate value. Solving the optimal policy for the infinite-horizon problems requires the discount factor $0 < \beta < 1$ to ensure that the expected reward is bounded and converged [9], [16], [17]. We assume that the duration of the whole communication is long enough and that T is approximately infinite. Our goal is to find the optimal total expected discounted reward for the whole communication period which corresponds to the optimal candidate relay selected policy. This optimum value is defined as below:

$$\mathcal{Z}^* = \max_{u\in\mathcal{U}} E_u\left[\sum_{t=0}^{T-1}\left(\mathcal{R}_{i_1(t)}^{A_1(t)} + \mathcal{R}_{i_2(t)}^{A_2(t)} + ... + \mathcal{R}_{i_{n_k}(t)}^{A_{n_k}(t)}\right)\beta^t\right]. \quad (17)$$

D. Solution to the Restless Bandit Problem

To solve the restless bandit problem, a hierarchy of increasingly stronger LP relaxations is developed based on the classical result on LP formulations of Markov decision chains (MDC) [16]. In order to formulate the restless bandit problem as a linear program, we introduce some performance measures as follows: $x_{i_{n_k}}^{A_{n_k}(t)}(u) = E_u[\sum_{t=0}^{T-1}(I_{i_{n_k}}^{A_{n_k}(t)}(t)\beta^t)]$ represents the expected discounted time that relay user n_k in state i_{n_k} given (active, passive) state of relay at time t; $I_{i_{n_k}}^{A_{n_k}(t)}(t) = 1$ if action $A_{n_k}(t) = 1$ and corresponds to $\{a_{n_k} = 1, P_{n_k,i}(t) = P_{n_k,i}^*(t), \omega_{n_k,i}(t) = \omega_{n_k,i}^*(t)\}$ is given at epoch t; Otherwise, $I_{i_{n_k}}^{A_{n_k}(t)}(t) = 0$ if action of relay $A_{n_k}(t) = 1$ and corresponds to $\{a_{n_k} = 0, P_{n_k,i}(t) = 0, \omega_{n_k,i}(t) = 0\}$. We denote X which is performance region spanned by performance vector $x = (x_{i_{n_k}}^{A_{n_k}}(u))_{i_{n_k}\in\mathcal{S}_{n_k} A_{n_k}\in\mathcal{A}}$ under all admissible $u \in \mathcal{U}$ policies .

Our problem can be formulated by the following linear program:

$$(\text{LP}) \quad Z^* = \max_{x\in X} \sum_{n_k\in\mathcal{N}_{n_k}^i} \sum_{i_{n_k}\in S_{n_k}} \sum_{A_{n_k}\in\mathcal{A}} \mathcal{R}_{i_{n_k}}^{A_{n_k}} x_{i_{n_k}}^{A_{n_k}}. \quad (18)$$

In order to solve (18), with the relaxation of polytope X that yields polynomial-size relaxation of the LP, we construct a primal and dual problem as in [16]. Denote by $\{\bar{x}_{i_{n_k}}^{A_{n_k}}\}$ and $\{\bar{\lambda}_{i_{n_k}}\bar{\lambda}\}$ the optimal primal and dual solution pair to the first-order relaxation (LP) and its dual (D^1). Let $\{\bar{\gamma}_{i_{n_k}}^1\}$, $\{\bar{\gamma}_{i_{n_k}}^0\}$ represent the rate of decrease in the objective-value of linear program (LP) per unit increase in the value of the variable $x_{i_{n_k}^1}$ and $x_{i_{n_k}^0}$, respectively, i.e, $\{\bar{\gamma}_{i_{n_k}}^0\} = \{\bar{\lambda}_{i_{n_k}} - \beta\sum_{j_{n_k}\in\mathcal{S}_{n_k}} p_{i_{n_k},j_{n_k}}^0 \bar{\lambda}_{j_{n_k}} - \mathcal{R}_{i_{n_k}}^0\}$ and $\{\bar{\gamma}_{i_{n_k}}^1\} =$

$\{\bar{\lambda}_{i_{n_k}} - \beta \sum_{j_{n_k} \in \mathcal{S}_{n_k}} p^1_{i_{n_k}, j_{n_k}} \bar{\lambda}_{j_{n_k}} - \mathcal{R}^1_{i_{n_k}}\}$, which must be non-negative. Based on this, white index of $\mathcal{R}_{n_k}$ in state i_{n_k} is defined as:

$$\delta_{i_{n_k}} = \bar{\gamma}^1_{i_{n_k}} - \bar{\gamma}^0_{i_{n_k}}. \quad (19)$$

Base on this parameter, each MUE broadcasts message to collect some candidate relay FUEs. Base on observing its history, each candidate relay FUE determines and computes i_{n_k}, $\mathcal{P}_{n_k}(t)$, $\mathcal{R}^{A_{n_k}}_{i_{n_k}}$, $P^*_{n_k,i}(i_{n_k}, A_{n_k}), \omega^*_{n_k,i}(i_{n_k}, A_{n_k})$ and sends to MUE i.Then, each FUE offline calculates the index $\delta_{i_{n_k}}$ according to (19). This index is stored in MUE's table. At epoch t, the MUE looks up the index-table to find out the corresponding a relay FUE's index with smallest value $\delta_{i_{n_k}}$.

V. Simulation results and discussions

This section presents the simulation results of our preposition. We use the system model as shown in Fig. 1. The radius of MBS and FBS are fixed to 1000m and 20m, respectively. The number of candidate relay FUEs of MUE i are assumed to be three. The power transmission of P_i = 100 mW, $P^{\max}_{n_k,i}$ = 100mW, B_w =1, $\sigma = 10^{-10}$ and the setup price $\zeta_{n_k} = 0.5$ are equal for all candidate relay FUEs. The expected channel gain states are divided into three states $\bar{\sigma}^{\text{good}}_{s,d}$, $\bar{\sigma}^{\text{normal}}_{s,d}$ and $\bar{\sigma}^{\text{bad}}_{s,d}$ and given by 0.8, 0.2 and 0.01, respectively. The case where the expected channel gain state between i and FBSs is in good state, it corresponds to direct transmission mode. The state of residual energy level is divided into two states high and low (available and not available for relaying data). The states of traffic model are on-off (the candidate relay FUE has own data or not). Consequently, there are 108 states for each available candidate relay FUE (IV-B).

The rewards, which are pairs of utility functions $\{U^*_{n_k}; U^*_i\}$, are determined in IV-C that depends on each pair (action, state) of candidate relay FUEs. Simultaneously, the optimal price and optimal power for relaying data are computed corresponding to maximum value pair of utility function values. In order to compute the rewards, the candidate relay FUEs will be computed and updated based on history computation, i.e, in considering static network environment, the MUE exchanges information with the candidate relay FUEs and give a pricing-based strategy to maximize its utility function as shown in Fig. 2.

We run the simulation with T = 1500 timeslots. In epoch t, the candidate relay FUE is selected based on white index table following (19). Depending on initial states of Primal-Dual Priority-Index Heuristic, the optimum expected utility function of MUE i and relay FUEs system are computed and represented in Fig. 3. Moreover, the expected utility functions value also depend on discount factor β, i.e $\beta = 0.3, 0.5, 0.8$.

Next, in order to recognize the dependence of the utility functions value on state of candidate relay, we set the transition probability value of the FUEs traffic state and energy level as follows: $\theta_{idle \to idle} = \theta_{busy \to idle} = \omega_{low \to high} = \omega_{high \to high} = p$. Initially from 108 states, consequently we set values p from 0.1 to 1 and only consider the expected utility value of MUE i. The results are shown in Fig. 4, when p value increases for of all candidate relay FUEs, the expected utility value

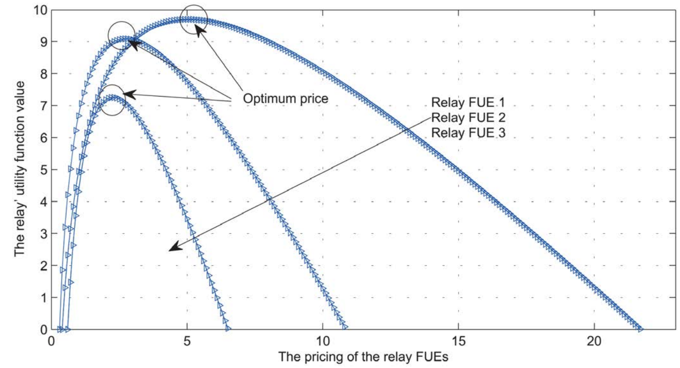

Fig. 2: The utility function of FUEs with pricing-based for each unit-power for relaying data.

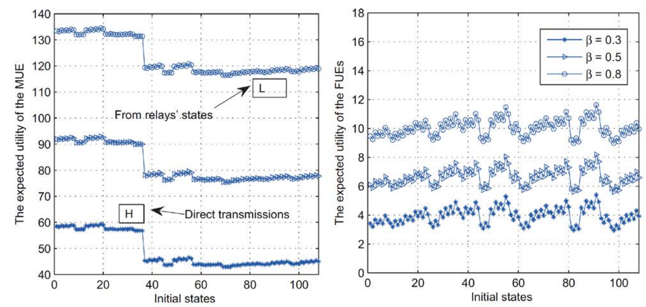

Fig. 3: The maximum expected utility of MUE i(left) and relay FUEs (right) depend on discount factor β and initial states.

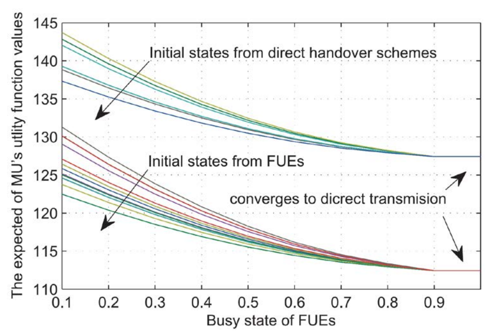

Fig. 4: The efficient of relay FUEs' traffic state to the MUE's expected reward with discount factor $\beta = 0.8$

.

will decrease. The expected utility values converges to direct transmission mode when p is equal 1 because there is no candidate FUE in the set of candidate relay FUEs which can support for relaying data.

VI. Conclusion

In this paper, we investigated a trading cooperative model in uplink HetNets. The Stackelberg game is formulated to maximize the utility functions of both the MUE and relay FUEs in static network model with one-shot. Moreover, we investigated the cooperative model in stochastic network environment. We applied a restless bandit problem to maximize total expected utility functions in a long-term. It can be inferred from the results that our proposal outperforms other schemes in terms of relay selection. In future work, we will consider cooperation among multiple MUEs and multiple relay FUEs with self-learning and self-optimizing.

REFERENCES

[1] D. Knisely, T. Yoshizawa, and F. Favichia, "Standardization of femtocells in 3gpp," *Communications Magazine, IEEE*, vol. 47, no. 9, pp. 68–75, 2009.

[2] S. Ortiz, "The wireless industry begins to embrace femtocells," *Computer*, vol. 41, no. 7, pp. 14–17, 2008.

[3] F. Pantisano, M. Bennis, W. Saad, and M. Debbah, "Spectrum leasing as an incentive towards uplink macrocell and femtocell cooperation," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 3, pp. 617–630, 2012.

[4] D. Liu, Y. Chen, T. Zhang, K. K. Chai, J. Loo, and A. Vinel, "Stackelberg game based cooperative user relay assisted load balancing in cellular networks," *Communications Letters, IEEE*, vol. 17, no. 2, pp. 424–427, 2013.

[5] P. Hande, M. Chiang, R. Calderbank, and J. Zhang, "Pricing under constraints in access networks: Revenue maximization and congestion management," in *INFOCOM, 2010 Proceedings IEEE*. IEEE, 2010, pp. 1–9.

[6] R. Zhang and L. Cai, "Markov modeling for data block transmission of ofdm systems over fading channels," in *Communications, 2009. ICC'09. IEEE International Conference on*. IEEE, 2009, pp. 1–5.

[7] Y. Fu, L. Yang, and W.-P. Zhu, "A nearly optimal amplify-and-forward relaying scheme for two-hop mimo multi-relay networks," *Communications Letters, IEEE*, vol. 14, no. 3, pp. 229–231, 2010.

[8] Z. Han and H. V. Poor, "Coalition games with cooperative transmission: a cure for the curse of boundary nodes in selfish packet-forwarding wireless networks," *Communications, IEEE Transactions on*, vol. 57, no. 1, pp. 203–213, 2009.

[9] H. S. Wang and N. Moayeri, "Finite-state markov channel-a useful model for radio communication channels," *Vehicular Technology, IEEE Transactions on*, vol. 44, no. 1, pp. 163–171, 1995.

[10] Y. Xie, J. Hu, Y. Xiang, S. Yu, S. Tang, and Y. Wang, "Modeling oscillation behavior of network traffic by nested hidden markov model with variable state-duration," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 24, no. 9, pp. 1807–1817, 2013.

[11] L. Muscariello, M. Mellia, M. Meo, M. Ajmone Marsan, and R. Lo Cigno, "Markov models of internet traffic and a new hierarchical mmpp model," *Computer Communications*, vol. 28, no. 16, pp. 1835–1851, 2005.

[12] Y. Wei, F. R. Yu, and M. Song, "Distributed optimal relay selection in wireless cooperative networks with finite-state markov channels," *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 5, pp. 2149–2158, 2010.

[13] P. Hu, Z. Zhou, Q. Liu, and F. Li, "The hmm-based modeling for the energy level prediction in wireless sensor networks," in *2007 2nd IEEE Conference on Industrial Electronics and Applications*, 2007, pp. 2253–2258.

[14] Y. Chen, Q. Zhao, V. Krishnamurthy, and D. Djonin, "Transmission scheduling for sensor network lifetime maximization: A shortest path bandit formulation," in *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, vol. 4. IEEE, 2006, pp. IV–IV.

[15] Z. Han, *Game theory in wireless and communication networks: theory, models, and applications*. Cambridge University Press, 2012.

[16] D. Bertsimas and J. Niño-Mora, "Restless bandits, linear programming relaxations, and a primal-dual index heuristic," *Operations Research*, vol. 48, no. 1, pp. 80–90, 2000.

[17] Y. Zhao, R. Adve, and T. J. Lim, "Improving amplify-and-forward relay networks: optimal power allocation versus selection," in *Information Theory, 2006 IEEE International Symposium on*. IEEE, 2006, pp. 1234–1238.