Stackelberg Game and Restless Bandits Problem for Towards Uplink Macrocell and Femtocell Cooperation

Tuan LeAnh, Cuong T.Do, Tai Manh Ho,S.M.Ahsan Kazmi, Saeed Ullah, Choong Seon Hong
Department of Computer Engineering, Kyung Hee University, 449-701, Korea
Email: {latuan,dtcuong, hmtai,ahsankazmi,saeed,cshong}@khu.ac.kr

Abstract

In this paper, we consider the cooperation among users in heterogeneous wireless networks (HetNets). Some users in HetNets will play a role as the relays for a macro user when the macro user cannot connect to the macro base stations or femto base stations directly. In this scenario, the macro user tries to find a best relay user in a set of candidate relays to maximize its utility function. In return, each relay user gives a pricing-based strategy to the macro user to maximize the relay user's utility function. This problem is formulated as a Stackelberg game in static environment. Moreover, with stochastic environment we model as a restless bandit problem to maximize the utility functions of the relay users and the macro user in a long-term. Simulation results illustrate the efficiency of our proposal.

I. Introduction

Recently, the new wireless communication paradigm has shifted to a future wireless network such as deployment of femtocell network [1]. One of the paradigms is known as the HetNet with coexistence of two-tiers which comprises of the macrocell underlaid on femtocell base stations (FBSs). FBSs are expected to lie at the heart of emerging wireless systems [2], [3]. When the macro users (MUEs) lie outside of theirs Macro Base Station (MBS) coverage, the MUEs will find other ways such as handover to the femtocell networks. In such scenarios the MUEs cannot handover via the FBSs of femtocell networks, they perform detecting idle-users (femtocell users or other macro users) for relaying data that satisfy their data rate requirements. Not like the other methods to relay data, we propose the relay selection scheme that captures a trading among the macro user (MUEs) and relay users based on Stackelberg game approach [4],[5]. The MUEs are willing to relay data are modelled as buyers and the candidate relays are considered as sellers. Moreover, because the available for relaying data of the candidate relay users depend on some characteristics as the relay user power level, are dynamic on-off, hence we consider our problem as a restless bandit problem to maximize a pair of utilities (the MUE and relay user) 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 consideration in stochastic environments is given in section IV. Next, some numerical results are showed in section V. Finally, section VI provides conclusion and our future works.

II. System model and formation problem

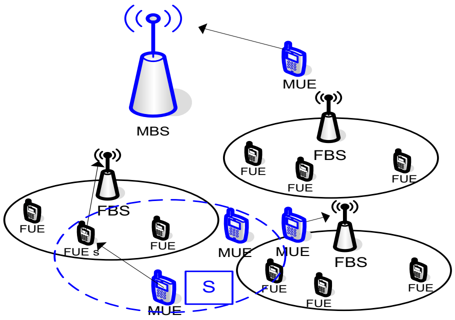

Figure 1: system model

We consider the uplink direction of an Orthogonal Frequency Division Multiple Access (OFDMA) macrocell network in which the set of K={1,⋯K} FBSs, FUEs and MUEs are deployed as showed in figure 1, respectively. We assume that, there is a mobile MUE m moves to a position where the MUE m cannot connect to its MBS or any FBSs for data transmission, directly. The MUE has an incentive to find a relay user s ($s \in$ a set of candidate relay user $\mathrm{S} = \{1,\ldots,S\}$) that connects to its base station to forward data packets. We assume at each time t, a FUE relay user can serve only one MUE that asks the relay users in S to relay data. Additionally, the MUE's original channel which is registered to MBS, is kept while the MUE's data meanwhile the MUE's data relays via another relay user.

In order to transmit via a relay user, the amplify-and-forward protocol procedure is applied based on using maximal ratio combining [6] as following:

$$R_m^{r(s,k)} = B_W \log\left(1 + \frac{P_m |h_{mk}|^2}{\sigma^2} + \frac{P_m P_{msk} |h_{ms}|^2 |h_{sk}|^2}{\sigma^2 (P_m |h_{ms}|^2 + P_{msk} |h_{sk}|^2 + \sigma^2)}\right) \quad (1)$$

where $R_m^{r(s,k)}$ is the data rate that MUE m chooses relay user $s \in S$ which belongs to base station k (k=0, mean that user s belongs to MBS). P_m, P_{msk} are power of MUE m and user relay user s, respectively.

In our problem, each relay user belongs to set S is only allowed to relay data when it has no data of its own to transmit or the power level of remaining battery is allowed to relay. Therefore, we model the relay user state with two-state Markov model to represent the relay user in timeslot t by X(t) = {0(not allow to relay),1(allow to relay)}. The utilized state transmission probability matrix can be written as follows:

$$i_s(t) = [\theta_{ij}(t)]_{2\times 2} \qquad (2)$$

Where $\theta_{ij}(t)$ =Pr{X(t+1) = j|X(t)=i}, for i, j∈X.

In real systems, the value in the aforementioned transition probability matrices can be obtained from the history observation of the relay users.

III. Stackelberg game analysis for trading model.

This section expresses the cooperative payment among the MUE m and FUEs in the set S. The MUE m is formulated as a buyer and the relay users as sellers. Each relay user is incentive to earn the payment which not only covers their forwarding cost but also obtains as much profit as possible. The utility function for payment is represented as follows [7]:

$$U_{msk}^s = (\omega_{msk} - \zeta_{nk}) P_{msk}, \qquad (3)$$

where ω_{msk} is unit power price of MUE m charges for relay users s. ζ_{sk} is cost of relay user s for an unit power of relaying data. Let P_{msk} be the power allocation level of relay user s to maximize utility function of MUE m. In order to compete with other relay users, the relay user s gives a strategy price ω_{msk} to maximize its utility function. At the MUE m side, the utility of MUE m when the MUE m chooses relay user s is defined as follow:

$$U_{msk}^m = R_{msk}^m - \omega_{msk} P_{msk} \qquad (4)$$

Where R_{msk}^m is data rate of the MUE m via relay user s, $s \in k$. The utility function of MUE m is a concave function with variable power P_s. First, the MUE m tries to find power level at relay side to maximize its utility function, then user relay user s decide a price to maximize its utility. Let P_{mnk}^* and ω_{mnk}^* are optimum value of (3), (4). By first order of U_{msk}^m followed by P_{msk}, from (1), (3), (4), we have as follows:

$$P_{mnk}^*(\omega_{msk}) = \min\left(\left[\frac{-A\omega_{msk} + \sqrt{B\omega_{msk}^2 + C\omega_{msk}}}{2\omega_{msk} D}\right]^+, P_s^{\max}\right), \quad (5)$$

where:

$$A = \frac{(\sigma^2 + P_m |h_{ms}|^2)(2\sigma^2 + P_m(|h_{ms}|^2 + |h_{mk}|^2))}{|h_{sk}|^2 (\sigma^2 + P_m |h_{mk}|^2)}, \quad B = \left(\frac{P_m |h_{ms}|^2 (\sigma^2 + P_m |h_{ms}|^2)}{|h_{sk}|^2 (\sigma^2 + P_m |h_{mk}|^2)}\right)^2$$

$$C = \frac{2P_m |h_{ms}|^2 \left(\sigma^2 + P_m |h_{ms}|^2\right)\left(\sigma^2 + P_m(|h_{ms}|^2 + |h_{mk}|^2)\right)}{\ln 10 |h_{sk}|^2 (\sigma^2 + P_m |h_{mk}|^2)^2}, \quad D = \frac{\sigma^2 + P_m(|h_{ms}|^2 + |h_{mk}|^2)}{\sigma^2 + P_m |h_{mk}|^2}.$$

Then, the relay user s will take an optimum price as $\omega_{msk}^* = \arg\max\left((\omega_{msk} - \zeta_{nk}) P_{mnk}^*\right)$ by first order by ω_{msk} variable.

IV. Restless Bandit problem for stochastic optimization

In this section, by dividing the considering time into discrete time and under stochastic optimization, we can maximize the expected reward $\mathcal{R}\,(P_{mnk}^*, \omega_{msk}^*)$ in a long term. In restless bandit problem, we consider the state space and transmission probability of candidate relay user s which has been mentioned in section II. We define the expected reward by optimizing relay user s selection as:

$$\mathcal{R}_{i_s(t)}^{a_x(t)} = \left\{P_{mnk}^*(i_s(t), a_s(t)); \omega_{msk}^*(i_s(t), a_s(t))\right\}, \quad (6)$$

where $a_s(t) \in \{0,1\}$, where "0" is idle state and "1" is active state of relay user s in the respective time slot.

For a stochastic process, a maximum immediate value is not equivalent to the maximum expected long-term accumulate value [8]. The notation by $u \in \mathcal{U}$ and $0 < \beta < 1$ are the admissible policy and the discount factor, respectively. The objective of the restless bandit problem is to find an optimal policy u^* that maximize the total expected discounted reward during the whole data transmission period:

$$Z(u^*) = \max_{u \in \mathcal{U}} \mathrm{E}_u \left[\sum_{t=0}^{T-1} \sum_{s=1}^{S} \mathcal{R}_{i_s(t)}^{a_x(t)} . \beta^t \right] \text{ (7).}$$

This restless bandit problem can be sold by the primal-dual index heuristic algorithm based on the first order relaxation, which has been demonstrated to have less complexity and very close performance in comparison with the optimal one [8]. The optimal reduced cost co-efficents can be represented as:

$$\gamma_{i_s}^0 = \lambda_{i_s} - \beta \sum_{j_s \in \Omega} \pi_{i_s j_s}^0 \lambda_{j_s} - R_{i_n}^0 \quad (8), \qquad \gamma_{i_s}^1 = \lambda_{i_s} - \beta \sum_{j_s \in \Omega} \pi_{i_s j_s}^1 \lambda_{j_s} + \lambda - R_{i_n}^1 \quad (9)$$

where Ω is state space of all relay users in set S, with state space is equal $2^{|S|}$. These values $\gamma_{i_s}^0$, $\gamma_{i_s}^1$ ≥ 0, and the mean rates of decrease in the objective value of the primal problem per unit increase in the optimum value in (primal, dual) of the variable $X_{i_s}^0$, $X_{i_s}^1$ respectively [8]. Let δ_{i_s} denotes the index for the s-th relay user when it is in state $i_s \in \Omega$, which is represented as $\delta_{i_n} = \gamma_{i_s}^1 - \gamma_{i_s}^0$. In each epoch t, the MUE m computes the indices of all the relay user in S according to their states and the relay user with the smallest indices is set active or selected.

IV. Simulation results

We utilize system model as figure 1 for simulation. Number of users in set S is equal 6 relay users. Moreover, assume that the position of all users are fixed in time T. The radius of FBS and MBS are 20m and 1000m respectively. The orthogonal channel has B_w=180 kHz, The probability on-off of relay users 1st to 6st are 0.2-0.8, 0.3-0.7, 0.4-0.6, 0.5-0.5, 0.1-0.9, 0.25-0.75, respectively. P_m = 50mW, P_{sk}(max) = 50mW, P_{s0}(max) = 100mW.

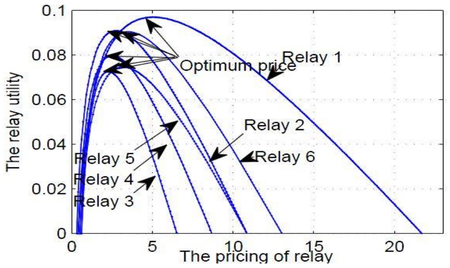

Figure 2: The optimum price of the relay users.

The optimum price for each relay user is determined from (2) and showed in figure 2. From (7), (8), (9) we compute the primal-dual index, we archived expected reward of both the MUE utility value and relay users utility value as showed in figure 3.

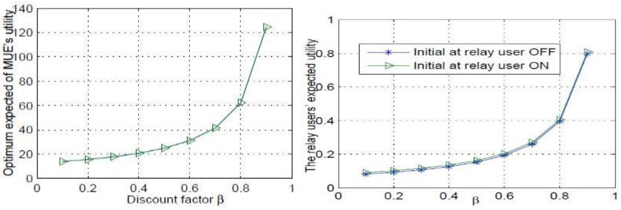

Figure 3: The maximum expected of the MUE and relay users' utility

V. Conclusions and future works

In this paper, we investigated trading model for cooperation among a MUE and relay users in which the MUE can relay via other users in macrocell or femtocell for data transmission. The Stackelberg game was formulated to maximize utility function of the MUE and relay users. Moreover, we considered our model in stochastic environment under restless bandit problem to maximize expected reward in a long-term. In future works, we will consider to the cooperation among multiple users and multiple relays in stochastic environment with self-learning and self-optimizing of users.

Acknowledment

This research was supported by the MSIP (Ministry of Science, ICT&Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (NIPA-2014(H0301-14-1020)) supervised by the NIPA (National IT Industry Promotion Agency). *Dr. CS Hong is the corresponding author

Reference

[1] Francesco Pantisano, Mehdi Bennis, Walid Saad, and Merouane Debbah, "Spectrum Leasing as an Incentive Towards Uplink Macrocess and Femtocell Cooperation". IEEE JSAC Arpil 2012.

[2] D.Knisely, T.Yoshizawa, and F.Favichia, "Standardization of femtocells in 3GPP," IEEE Commun. Sept.2009.

[3] S. Ortiz, "The wireless industry begins to embrance femtocells," Computer, vol.41, no.7.pp.14-17, Jul,2008.

[4] J.Zhang and Q.Zhang, "Stackelberg game for uility-based cooperative cognitive radio networks", IEEE Trans.Mob.Comput., vol.8, no.7, July 2009.

[5] Lingjie Duan, Jianwei Huang, Biying Shou "Investment and Pricing with Spectrum Uncertainy: A Cognitive Operator's Perspective", IEEE Transactions on Mobile Computing, Vol.10, No.11, November 2011.

[6] Youhua Fu; Luxi Yang; Wei-Ping Zhu, "A nearly optimal amplifyand-forward relaying scheme for two-hop MIMO multi-relay networks," Communications Letters, March 2010

[7] Zhu Han, Dusit Nivato, Walid Saad, Tamer Basar, Are Hirungnes, "Game Theory in wireless and Communication Networks: Theory, Models, and Applications" Chapter 12 page: 402-409.

[8] D.Bertsimas and J.Nino-Mora, "Restless bandits, linear programming relaxations, and a primal dual index heuristic" Oper.Res., vol.48, no.1, pp.80-90, Jan./Feb.2000.